# Standards and Practices for Forecasting

J. Scott Armstrong
University of Pennsylvania, armstrong@wharton.upenn.edu

Reprinted with permission. Published in *Principles of Forecasting: A Handbook for Researchers and Practitioners*, J. Scott Armstrong (ed.): Norwell, MA: Kluwer Academic Publishers, 2001.

*Principles of Forecasting: A Handbook for Researchers and Practitioners,*
J. Scott Armstrong (ed.): Norwell, MA: Kluwer Academic Publishers, 2001

# Standards and Practices for Forecasting

J. Scott Armstrong
The Wharton School, University of Pennsylvania

**ABSTRACT**

One hundred and thirty-nine principles are used to summarize knowledge about forecasting. They cover formulating a problem, obtaining information about it, selecting and applying methods, evaluating methods, and using forecasts. Each principle is described along with its purpose, the conditions under which it is relevant, and the strength and sources of evidence. A checklist of principles is provided to assist in auditing the forecasting process. An audit can help one to find ways to improve the forecasting process and to avoid legal liability for poor forecasting.

---

When managers receive forecasts, they often cannot judge their quality. Instead of focusing on the forecasts, however, they can decide whether the *forecasting process* was reasonable for the situation. By examining forecasting processes and improving them, managers may increase accuracy and reduce costs.

One can examine the forecasting processes by systematically judging it against the 139 forecasting principles presented. These principles, organized into 16 categories, cover formulating problems, obtaining information, implementing methods, evaluating methods, and using forecasts.

Why do you need 139 principles? *You will not need all of them in any one situation*. Nearly all of the principles are conditional on the characteristics of the situation. It would be misleading to write a book on "The Five Principles Used by Successful Forecasters." They could never be appropriate for all the different situations that can arise.

The principles were drawn primarily from the papers in *Principles of Forecasting*. They include the major principles, but ignore some that are specific only to a certain forecasting method. To help ensure that the principles are correct, this paper was subjected to extensive peer review over a period of three years. The paper was also posted in full text on the Forecasting Principles website with a plea for comments. There were over 36,000 visitors to the site during the three years and helpful suggestions were received. Twenty experts provided careful reviews, and suggestions were obtained when versions of the paper were presented at five academic conferences.

I describe the strength of evidence for each principle and provide sources of empirical evidence. Many of the forecasting principles are based on expert opinion. I use the term "common sense" when it is difficult to imagine that things could be otherwise. "Received wisdom" indicates that the vast majority of experts agree.

Forecasters often ignore common sense and received wisdom. This observation was reinforced when I presented early versions of the principles to practitioners at the International Association of Business Forecasters in Philadelphia in 1997 and to academics at the "Judgmental Inputs to the Forecasting Process" conference at University College London in 1998. At both meetings, respondents to a questionnaire agreed with a vast majority of the principles, but they reported that few of these principles were followed in practice.

## FORMULATING THE PROBLEM

### *1. Setting Objectives*

Specify the objectives in the situation, then consider what decisions relate to reaching those objectives. The issues in this section can help to decide whether it is worthwhile to use formal procedures to make forecasts.

**1.1 Describe decisions that might be affected by the forecasts.**

**Description:** Analysts should examine how decisions might vary depending on the forecast.

**Purpose:** To improve the use of forecasts by relating them to decision making.

**Conditions:** Forecasts are needed only when they may affect decision making. If there are no decisions to be made, then there is no economic justification to do forecasting. Or, if the decision has already been made and cannot be changed, there is no need to make forecasts. Ignore this principle if the forecasts are strictly for entertainment, as with election-night forecasts.

**Strength of evidence:** Common sense.

**Source of evidence:** See Fischhoff (2001) for related evidence.

**1.2 Prior to forecasting, agree on actions to take assuming different possible forecasts.**

**Description.** One approach is to ask decision makers to describe what forecasts will change their decisions. Another is to present alternative possible forecasts and ask what decisions they would make. For example, “If the forecast is less than 100, we cancel the project. If it is between 100 and 149, we get more information. If it is 150 or more, we continue.” Griffith and Wellman (1979) showed that independent quantitative forecasts of bed needs, obtained without prior agreement about how to use them, were ignored by hospital administrators when the forecasts were not to their liking.

**Purpose:** To improve the use of forecasting.

**Conditions:** Forecasts are needed only when they may affect decision making.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**1.3 Make sure forecasts are independent of politics.**

**Description.** Separate the forecasting process from the planning process. One possibility is to have one group do the forecasting and another do the planning. Separating these functions could lead to different reports such as ones showing forecasts for alternative plans. This principle is sensible and important, yet it is often ignored. This is not surprising. Consider, for example, that you received a forecast that the passage of right-to-carry gun laws in the U.S. would have beneficial effects, such as reduced deaths. Would you consider that forecast in deciding how to vote on this issue?

**Purpose:** To improve the use of forecasts by reducing bias.

**Conditions:** Impartial forecasts are especially important when they imply major changes.

**Strength of evidence:** Common sense.

**Source of evidence:** Fildes and Hastings (1994), Griffith and Wellman (1979), Harvey (2001), Larwood and Whittaker (1977), and Sanders and Ritzman (2001).

**1.4 Consider whether the events or series can be forecasted.**

**Description.** Would using formal forecasting procedures produce better forecasts than the current procedure or a naive benchmark? For example, short-term forecasts of the stock market do not improve accuracy (unless they are based on inside information).

**Purpose:** To reduce costs by avoiding useless forecasting efforts.

**Conditions:** When prior research shows that an area is unlikely to benefit, avoid formal forecasting. Use it, however, when formal forecasting produces accuracy equivalent to the current method but at a lower cost.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Much evidence shows that forecasters cannot beat the stock market with respect to accuracy. This goes as far back as Cowles (1933) and has continued ever since.

**1.5 Obtain decision makers' agreement on methods.**

**Description.** Describe how the forecasts are to be made, and do so in intuitive terms. Do the decision makers agree that they make sense? It may help to propose using a forecasting method on an experimental basis.

**Purpose:** Agreement can improve the use of forecasts. Acceptance of the forecasts is more likely if decision makers believe the procedures are relevant.

**Conditions:** The decision makers' acceptance of forecasting methods is important when they control the use of the forecasts.

**Strength of evidence:** Some empirical evidence.

**Source of evidence:** Research on organizational behavior supports this principle, and implementing this principle proved useful in a laboratory experiment (Armstrong 1982).

## *2. Structuring the Problem*

The problem should be structured so the analyst can use knowledge effectively and so that the results are useful for decision making.

**2.1 Identify possible outcomes prior to making forecasts.**

**Description.** Brainstorming about possible outcomes helps in structuring the approach. For example, experts might be asked to brainstorm the possible outcomes from the imposition of an affirmative action plan in a workplace.

**Purpose:** To improve accuracy.

**Conditions:** Determining possible outcomes is especially important for situations in which the outcomes are not obvious or in which a failure to consider a possible outcome might bias the forecast.

**Strength of evidence:** Indirect evidence.

**Source of evidence:** Tiegen (1983) shows how the specification of outcomes can affect predictions. For example, as new outcomes are added to a situation, forecasters often provide probabilities that exceed 100 percent. Arkes (2001) summarizes evidence relevant to this issue.

**2.2 Tailor the level of data aggregation (or segmentation) to the decisions.**

**Description.** Decision makers should help to determine the need for forecasts specified by time, geography, or other factors. One can make forecasts, however, for various components that can then be aggregated or disaggregated to fit the decision needs. Thus, the analyst can focus on the level of aggregation that yields the most accurate forecasts.

**Purpose:** To improve the use of forecasts by tailoring them to decisions.

**Conditions:** Sufficient data must exist to enable different levels of aggregation.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**2.3 Decompose the problem into parts.**

**Description.** Use a bottom-up approach. That is, forecast each component, then combine them.

**Purpose:** To improve forecast accuracy by improving reliability. Also, by decomposing the problem, you can more effectively use alternative sources of information and different forecasting methods.

**Conditions:** It is helpful to decompose the problem in situations involving high uncertainty and extreme (very large or very small) numbers. Additive breakdowns are preferable to multiplicative ones if the components' errors are highly correlated. Disaggregation will not improve accuracy if the components cannot be measured reliably. Decomposition by multiplicative elements can improve accuracy when you can forecast each of them more accurately than the target value.

**Strength of evidence:** Received wisdom and strong empirical evidence.

**Source of evidence:** Armstrong (1985) cites many studies. Evidence is also provided by Armstrong, Adya and Collopy (2001), Harvey (2001), and MacGregor (2001).

**2.4 Decompose time series by causal forces.**

**Description:** Causal forces represent the expected directional effects of the factors that affect a series. They can be classified into the following categories: growth, decay, opposing, regressing, and supporting. Decompose by force, make extrapolations of the components, then synthesize the overall forecast.

**Purpose:** To improve forecast accuracy.

**Conditions:** Decompose by causal forces when time series are affected by factors that have conflicting effects on the trends and when they can be decomposed according to the type of causal force. This procedure can also be used for judgmental forecasts.

**Strength of evidence:** Weak empirical evidence.

**Source of evidence:** Burns and Pearl (1981) were unsuccessful in an attempt to use causal reasoning in helping experts decompose a forecasting problem. Armstrong, Adya and Collopy (2001) found that such decomposition improved accuracy in extrapolation.

**2.5 Structure problems to deal with important interactions among causal variables.**

**Description:** Interactions imply that the relationship of $X_1$ to Y is related to the level of $X_2$.

**Purposes:** To improve forecast accuracy; to assess effects of policy variables.

**Conditions:** When interactions have important effects, you should account for them in the analysis. Though decomposition requires large samples, it provides a simple way to handle interactions.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Little research has been done on this issue. However, in a study of sales at 2,717 gas stations, Armstrong and Andress (1970) found that decomposition to handle interactions substantially improved accuracy in comparison with forecasts from a regression model.

**2.6 Structure problems that involve causal chains.**

**Description:** Given a series of effects such as X causes Y, which then causes Z, simultaneous equations have not led to improved accuracy. Instead, construct a series of linked models. That is, develop a model using X to predict Y. Then use the predictions for Y, in a model to predict Z.

**Purpose:** To improve accuracy.

**Conditions:** Use causal chains when they have important effects, their relationships are well known, and the timing can be accurately forecast.

**Strength of evidence:** Received wisdom and some empirical evidence.

**Source of evidence:** Allen and Fildes (2001); Armstrong (1985, pp. 199-200).

**2.7 Decompose time series by level and trend.**

**Description:** The separate examination of level and trend is one of the oldest and more enduring principles, and it is widely used in practice.

**Purpose:** To improve forecast accuracy.

**Conditions:** Decomposition is useful when there are significant trends that can be assessed by different methods. For example, judgmental methods are especially useful for incorporating recent information into estimates of levels.

**Strength of evidence:** Received wisdom and some empirical evidence.

**Source of evidence:** Armstrong (1985, pp. 235-238) summarizes evidence from eight studies.

## OBTAINING INFORMATION

This section examines the identification, collection, and preparation of data to be used in forecasting.

### *3. Identify Data Sources*

Identify data that might be useful in making forecasts. While this should be guided by theory, you may need creativity in seeking alternative types of data.

**3.1 Use theory to guide the search for information on explanatory variables.**

**Description:** Theory and prior research can help in the selection of data on explanatory variables. For example, in sales forecasting, a common model is to predict sales based on market size, ability to purchase, and need. The search for information could then be limited to these variables. Operational measures are then needed – such as income, availability, and price – to measure “ability to purchase.”

**Purpose:** To improve forecast accuracy.

**Conditions:** To follow this principle, analysts must have good prior knowledge. That knowledge can be based on experience or research studies.

**Strength of evidence:** Received wisdom with little empirical testing. Received wisdom has been questioned by practitioners who violate this principle in the belief that more information is always better. Some researchers have ignored this principle in favor of data mining, which assumes that the data will reveal causal patterns.

**Source of evidence:** Armstrong (1985, pp. 52-57) describes studies that show how one can get absurd results by ignoring theory. It also describes a study in which a theory-based econometric model was more accurate than a model based only on statistical criteria.

**3.2 Ensure that the data match the forecasting situation.**

**Description:** Data about past behavior in that situation are often the best predictors of future behavior.

**Purpose:** To improve forecast accuracy.

**Conditions:** This principle applies to all conditions, but especially when it is not obvious which data you should use to match the situation.

**Strength of evidence:** Strong empirical support from research in personnel selection.

**Source of evidence:** Armstrong (2001a,f) and Morwitz (2001) summarize some of the evidence. For example, studies have shown that personnel selection tests should be similar to the job requirements.

**3.3 Avoid biased data sources.**

**Description:** Avoid data collected by persons or organizations that are obviously biased to particular viewpoints, perhaps because they are rewarded for certain outcomes. Thus, for extrapolating crime rates, victim surveys would be preferable to police records. Identify biases before analyzing the data, especially when people are emotional about the outcomes, as in forecasting the effects of environmental hazards. Consider this forecast made by the biologist Paul Ehrlich, on the first Earth Day on April 22, 1970: “Population will inevitably and completely outstrip whatever small increases in food supply we make.”

**Purpose:** To improve accuracy.

**Conditions:** Follow this principle when you can identify biased sources and alternative sources of data are available.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**3.4 Use diverse sources of data.**

**Description:** Find alternative ways of measuring the same thing. If unbiased sources are not available, find sources with differing (and hopefully compensating) biases. For example, exports of a product from country A to country B should equal imports of that product to country B from country A. If the alternative sources do not agree, consider combining estimates from each source.

**Purpose:** To improve forecast accuracy.

**Conditions:** Use diverse sources when biases are likely to occur.

**Strength of evidence:** Received wisdom with some empirical support.

**Source of evidence:** Armstrong (1985, p. 236) mentions two studies.

**3.5 Obtain information from similar (analogous) series or cases. Such information may help to estimate trends.**

**Description:** Trends in analogous time series may provide useful information for trends in the series of interest. For example, the trendline for sales of all luxury cars might help to estimate the projected trend for a specific brand of luxury car.

**Purpose:** To improve forecast accuracy.

**Conditions:** You must be able to identify similar data. Analogous data are especially important when the series of interest has few observations or high variability.

**Strength of evidence:** Received wisdom with little empirical support.

**Source of evidence:** Duncan, Gore and Szczpula (2001) provides some evidence for time series. Claycamp and Liddy (1969) provide evidence from their study on sales forecasts for new products.

## *4. Collecting Data*

Once you identify a source, collect relevant, valid, and reliable data.

**4.1 Use unbiased and systematic procedures to collect data.**

**Description:** Data-collection procedures should be as free of bias as possible; the experts should have nothing to gain from biasing the data, and they should not be committed to a certain viewpoint.

**Purpose:** To improve forecast accuracy.

**Conditions:** Use this principle only when you have alternative sources of data. It is especially important when using judgmental methods and when bias is likely, as in forecasting the effects of deregulation, capital punishment, welfare reform, or charter schools.

**Strength of evidence:** Strong empirical evidence.

**Source of evidence:** Armstrong, Brodie and Parsons (2001) summarize research showing that a researcher's hypothesis can bias aspects of the research process. Armstrong (1985, pp. 108-111) reviews studies showing that bias causes serious forecast errors. For example, Rosenthal and Fode (1963) showed how the collection of data in experiments on rats was influenced by an experimenter's hypothesis.

**4.2 Ensure that information is reliable and that measurement error is low.**

**Description:** This applies, most importantly, to the dependent variable.

**Purpose:** To improve forecast accuracy.

**Conditions:** This principle is important when measurement is difficult, such as for events that occur infrequently (e.g., terrorist attacks or cancer due to environmental hazards). When policy models are involved, this applies also to explanatory variables.

**Strength of evidence:** Common sense.

**Source of evidence:** Rowe and Wright (2001), Stewart (2001), and Webby, O'Connor and Lawrence (2001) provide evidence on the importance of reliable information. However, the effects on forecasting accuracy are often small. Armstrong (1985, pp. 221-222) cites three studies showing that revised (and presumably more accurate) economic data yielded only small gains in accuracy.

**4.3 Ensure that the information is valid.**

**Description:** Does the information have face validity (i.e., do impartial experts agree that the information is relevant)? Does the information have construct validity (e.g., do alternative measures of the same variable agree with one another)?

**Purpose:** To improve forecast accuracy.

**Conditions:** This applies to all problems, but it is most important in situations where validity is low. Suppose that one needs information on the effectiveness of an educational system and a decision is made to examine trends in schools' teacher ratings. As it turns out, teacher ratings do not provide valid evidence of learning (Armstrong 1998). A similar problem occurs in predicting the success of managers, because advancement in the organization has little relationship to managers' effectiveness (Luthans, Hodgetts and Rosenkrantz 1988).

**Strength of evidence:** Received wisdom.

**Source of evidence:** Violations of this principle have detrimental effects (Armstrong 2001c).

**4.4 Obtain all of the important data**

**Description:** For time series, use all available time periods unless a strong a priori case can be made that a discontinuity has occurred. Obtain information about special events in the series.

**Purpose:** To improve forecast accuracy.

**Conditions:** This is especially applicable when large changes are involved.

**Strength of evidence:** Strong empirical evidence.

**Source of evidence:** Allen and Fildes (2001), Armstrong (2001f), Dorn (1950), and Makridakis (1996) present evidence for the importance of this principle. Simon (1985) showed how the principle is sometimes ignored, such as in the U.S. oil crisis that occurred in the early 1980s.

**4.5 Avoid the collection of irrelevant data.**

**Description:** Instead of casting a wide net for data, collect only data that are *clearly* relevant. Irrelevant data may confuse experts when making judgmental forecasts and introduce spurious relationships into quantitative models.

**Purpose:** To improve accuracy and reduce costs.

**Conditions:** This applies to all types of forecasting except extrapolation methods.

**Strength of evidence:** Strong empirical evidence.

**Source of evidence:** Armstrong (1985, p. 104) summarized results from four studies. Whitecotton, Sanders and Morris (1998) found that irrelevant data harmed accuracy. Gaeth and Shanteau (1984) showed that experiential training led judges to ignore irrelevant data; warnings alone did not help.

**4.6 Obtain the most recent data.**

**Description**: Even if the recent data are preliminary, they are likely to contain useful information.

**Purpose:** To improve accuracy.

**Conditions:** Recency is especially relevant to time-series data and to situations when there has been much recent change.

**Strength of evidence:** Common sense.

**Source of evidence:** Ash and Smyth (1973) and Joutz and Stekler (1998) provide limited evidence.

## *5. Preparing Data*

Prepare data for the forecasting processes.

**5.1 Clean the data.**

**Description**: Adjust for mistakes, changing definitions, missing values, and inflation. Keep a log to record adjustments. Armstrong (2001f) discusses this topic.

**Purpose:** To improve forecast accuracy.

**Conditions:** Clean the data when you can identify reasons for the revisions.

**Strength of evidence:** Common sense.

**Source of evidence:** Chatfield (1995, Chapter 6) provides indirect evidence.

**5.2 Use transformations as required by expectations.**

**Description:** A transformation should ensure that the data correspond with accepted theory and with domain experts' expectations as to proper relationships. For example, in forecasting economic behavior, you should typically expect constant elasticities (constant percentage relationships) so a log-log transformation may be called for. Comparing the historical fit to the data is considered to be an ineffective way to decide on transformations.

**Purpose:** To ensure that the forecasting procedure is valid in new situations.

**Conditions:** Transformations are especially important when large changes are expected. It is assumed that measurement errors are small. Otherwise, you might better use a conservative (e.g., additive) model.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Complex transformations, even when well supported by researchers' arguments, have not been shown to produce accurate forecasts. Little research has been done on this topic although Armstrong (1985, p. 202) reports on four studies. Also see Meade and Islam (2001).

**5.3 Adjust intermittent series.**

**Description:** Aggregate data across time, space, or decision units to avoid zeros. Consider forecasting the time to the next positive value.

**Purpose:** To improve forecast accuracy and to better assess uncertainty.

**Conditions:** Applies to time series and when only non-negative integer values are sensible.

**Strength of evidence:** Received wisdom and one empirical study.

**Source of evidence:** Willemain et al. (1994) provide evidence using tests on artificial and actual data.

**5.4 Adjust for unsystematic past events.**

**Description:** Use statistical techniques and/or domain knowledge to make adjustments for unsystematic past events. For example, a hurricane might have harmed sales.

**Purpose:** To improve forecast accuracy.

**Conditions:** You need to identify the timing and impact of the event with reasonable accuracy. Adjustments are especially important if the events are recent.

**Strength of evidence:** Some empirical evidence.

**Source of evidence:** Armstrong (2001f), Armstrong, Adya and Collopy (2001), and Duncan, Gorr and Szczypula (2001) provide evidence.

**5.5 Adjust for systematic events.**

**Description:** Adjust for systematic events (e.g., seasonality, holidays, and trading days for time series).

**Purpose:** To improve forecast accuracy.

**Conditions:** For time series, use seasonal adjustments only when seasonal changes are expected. This requires domain knowledge about causes of seasonality. (For example, photographic film sales vary by time of year in some locations because of tourism.)

**Strength of evidence:** Received wisdom.

**Source of evidence:** Makridakis et al. (1982) present evidence showing that seasonal adjustments generally reduce forecast errors.

**5.6 Use multiplicative seasonal factors for trended series when you can obtain good estimates for seasonal factors.**

**Description:** Multiplicative seasonal factors can represent behavior for much socioeconomic data.

**Purpose:** To improve forecast accuracy.

**Conditions:** You should use multiplicative adjustments when: (1) the seasonal pattern is well known and stable, (2) measurement errors are small, (3) ample data are available, (4) data are ratio scaled, and (5) the data show a strong trend. Lacking these conditions, multiplicative factors can be risky, so additive trends might be appropriate.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**5.7 Damp seasonal factors for uncertainty.**

**Description:** Seasonal factors can introduce errors in situations involving high uncertainty.

**Purpose:** To improve accuracy.

**Conditions:** Damp seasonal factors when estimates of seasonal factors are uncertain and when the seasonal pattern is likely to change.

**Strength of evidence:** Weak.

**Source of evidence:** Armstrong (2001f) cites indirect evidence.

**5.8 Use graphical displays for data.**

**Description:** When judgment is involved, graphical displays may allow experts to better assess patterns, to identify mistakes, and to locate unusual events. However, experts might also be misled by graphs if they try to extend patterns from the past.

**Purpose:** To improve accuracy.

**Conditions:** Graphical displays are useful when analysts have good domain knowledge and when there are clear patterns in the data. Experts should be trained so that they do not try to match time patterns when making judgmental forecasts in uncertain situations.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Studies have shown only minor benefits for graphs (Harvey 2001, Stewart 2001, and Webby, O'Connor and Lawrence 2001), but none of the studies were conducted in situations where forecasters had good domain knowledge.

## IMPLEMENTING FORECASTING METHODS

This section examines the selection and implementation of judgmental and quantitative methods. These tasks become more complex when policy decisions are involved. In some situations judgmental and quantitative methods should be integrated or their forecasts should be combined.

### *6. Selecting Methods*

Select the most appropriate methods for making the forecasts. You can expect that more than one forecasting method will be useful for most forecasting problems.

**6.1 List all the important selection criteria before evaluating methods.**

**Description:** Accuracy is only one of many criteria, as described in Armstrong (2001c). The relevant criteria should be specified at the start of the evaluation process.

**Purpose:** To select the most appropriate forecasting methods.

**Conditions:** This applies only when more than one feasible method exists. It is important when there are many criteria.

**Strength of evidence:** Received wisdom and indirect evidence.

**Source of evidence:** This principle was inferred from research on how people evaluate alternatives. References to this research are provided in Armstrong (2001c).

**6.2 Ask unbiased experts to rate potential methods.**

**Description:** Experts in forecasting and domain experts may be able to determine which forecasting methods are most useful for the task at hand. Armstrong (2001d) describes procedures for rating methods.

**Purpose:** To select the most appropriate methods.

**Conditions:** More than one feasible method exists, and a number of criteria are important.

**Strength of evidence:** Received wisdom, although in practice, forecasters seldom use formal ratings.

**Source of evidence:** None.

**6.3 Use structured rather than unstructured forecasting methods.**

**Description:** Structured methods are those consisting of systematic and detailed steps that can be described and replicated.

**Purpose:** To select the most appropriate forecasting methods.

**Conditions:** Structured methods are useful when accuracy is a key criterion and where the situation is complex.

**Strength of evidence:** Strong empirical evidence from a number of areas.

**Source of evidence:** Armstrong (1985) summarizes evidence that structured methods provide more accurate forecasts.

**6.4 Use quantitative methods rather than qualitative methods.**

**Description:** Quantitative methods tend to be less biased, and they make more efficient use of data.

**Purpose:** To improve forecast accuracy.

**Conditions:** Quantitative methods are appropriate when relevant data are available and they are especially useful in forecasting large changes, as in long-range economic forecasting.

**Strength of evidence:** Strong empirical evidence. This principle seems counterintuitive to many people.

**Source of evidence:** Allen and Fildes (2001), Armstrong (2001d), and Stewart (2001) summarize extensive evidence.

**6.5 Use causal methods rather than naive methods if feasible.**

**Description:** It is generally desirable to consider factors that cause changes in the variable of interest.

**Purpose:** To select the most appropriate methods.

**Conditions:** Use causal methods given (1) knowledge of causal relationships, (2) data on the causal variables, (3) expectations of large changes, (4) accurate forecasts of the causal variables, and (5) the cost of forecasting is small relative to its potential benefits. Furthermore, causal methods are important when one must forecast the effects of policy changes.

**Strength of evidence:** Strong empirical evidence.

**Source of evidence:** Armstrong (2001d) and Allen and Fildes (2001) summarize evidence showing that causal methods are more accurate than other methods under the above conditions.

**6.6 Select simple methods unless empirical evidence calls for a more complex approach.**

**Description:** Use few variables and simple relationships.

**Purpose:** To improve accuracy, aid understanding, reduce mistakes, and reduce costs of forecasting.

**Conditions:** While research shows this principle to be widely applicable, complex methods have proven useful in situations where there is extensive knowledge about relationships.

**Strength of evidence:** Strong empirical evidence. This principle seems counterintuitive when the situation is complex.

**Source of evidence:** Armstrong (1985) summarizes evidence showing that while some complexity may improve accuracy, seldom does one need highly complex methods. In some studies, complexity harmed accuracy.

**6.7 Match the forecasting method(s) to the situation.**

**Description:** Select methods that are appropriate given the criteria, the availability and type of data, prior knowledge, presence of conflict, amount of change expected, and value of forecast accuracy.

**Purpose:** To select the most appropriate forecasting method.

**Conditions:** When alternative methods are feasible and there is much uncertainty.

**Strength of evidence:** Strong empirical evidence.

**Source of evidence:** Armstrong (2001d) summarizes the evidence.

**6.8 Compare track records of various forecasting methods.**

**Description:** The comparisons should be in similar situations. This analysis can be expensive and time consuming. (Section 13 covers how to evaluate alternative forecasting methods.)

**Purpose:** To improve forecast accuracy.

**Conditions:** To compare methods, you need many ex ante forecasts made in similar situations.

**Strength of evidence:** Received wisdom and weak empirical evidence.

**Source of evidence:** Armstrong (2001d) summarizes evidence from three studies.

**6.9 Assess acceptability and understandability of methods to users.**

**Description:** Ask users what information they need in order to accept a proposed method.

**Purpose:** To increase the likelihood that decision makers accept forecasts and use them properly.

**Conditions:** When you need management support to implement forecasts. This typically applies to important forecasts in situations subject to large changes.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Yokum and Armstrong's (1995) survey of practitioners found understandability of a forecasting method to be an important criterion.

**6.10 Examine the value of alternative forecasting methods.**

**Description:** Examine whether the costs are low relative to potential benefits. Forecasters seldom do this, primarily because of the difficulty of assessing benefits. One approach to assessing benefits is described in the practitioners' section at hops.wharton.upenn.edu/forecast.

**Purpose:** To ensure that forecasting is cost effective.

**Conditions:** This principle is unnecessary when potential savings are obviously large relative to the costs of forecasting.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

### *7. Implementing Methods: General*

Some principles are common to implementing all forecasting methods.

**7.1 Keep forecasting methods simple.**

**Description:** Complex methods may include errors that propagate through the system or mistakes that are difficult to detect. Select simple methods initially (Principle 6.6). Then use Occam's Razor; that is, use simple procedures unless you can clearly demonstrate that you must add complexity.

**Purpose:** To improve the accuracy and use of forecasts.

**Conditions:** Simple methods are important when many people participate in the forecasting process and when the users want to know how the forecasts are made. They are also important when uncertainty is high and few data are available.

**Strength of evidence:** Strong empirical evidence. Many analysts find this principle to be counterintuitive.

**Source of evidence:** This principle is based on evidence reviewed by Allen and Fildes (2001), Armstrong (1985), Duncan, Gorr and Szczypula (2001), and Wittink and Bergestuen (2001).

**7.2 The forecasting method should provide a realistic representation of the situation.**

**Description:** Realism may call for adding some complexity, as forecasters sometimes do when developing econometric models. For example, to predict how someone will perform on a job, have them perform tasks that are representative of those in the job.

**Purpose:** To improve forecast accuracy.

**Conditions:** Often the matching of the method to the situation will be obvious, but this principle is most important when the match is not obvious. It is important when the situation is complex, as often happens for situations involving conflict among groups.

**Strength of evidence:** Received wisdom and some evidence.

**Source of evidence:** Armstrong (2001 a, c).

**7.3 Be conservative in situations of high uncertainty or instability.**

**Description:** To the extent that uncertainties and instabilities occur in the data or in expectations about the future, reduce changes in the time-series forecasts. For cross-sectional data, make sure that forecasts do not deviate much from an appropriate base rate.

**Purpose:** To improve forecast accuracy.

**Conditions:** This applies when the data contain much measurement error, high variation about the trend line has occurred or is expected, instabilities have occurred or are expected, or the forecast goes outside the range of the historical data.

**Strength of evidence:** Common sense and some empirical evidence.

**Source of evidence:** Armstrong (2001f) and Armstrong, Adya and Collopy (2001) summarize relevant evidence with respect to forecasting time series.

**7.4 Do not forecast cycles.**

**Description:** Cycles generally refer to systematic fluctuations in annual data, but they can also occur in other data such as hourly electric power demands. Seasonal variations are treated separately.

**Purpose:** To improve forecast accuracy.

**Conditions:** This applies unless you know (e.g., based on contractual relationships or on physical or biological laws) that cycles will occur and have good knowledge about timing.

**Strength of evidence:** Much research has been devoted to showing the value of annual cycles, but it has produced no favorable evidence.

**Source of evidence:** Armstrong (2001f).

**7.5 Adjust for events expected in the future.**

**Description:** Use domain knowledge about planned changes (e.g., a large price reduction as part of a sale for a product) or environmental changes (e.g., a snowstorm).

**Purpose:** To improve forecast accuracy.

**Conditions:** You must be able to identify the timing and impact of the event with reasonable accuracy.

**Strength of evidence:** Received wisdom and indirect evidence.

**Source of evidence:** Armstrong, Adya and Collopy (2001) and Sanders and Ritzman (2001) summarize indirect evidence.

**7.6 Pool similar types of data.**

**Description:** Use similar types of data to estimate key elements of a forecasting model such as seasonal factors, base rates, trends, or relationships. One way to identify similar data is to look for data that are subject to the same causal forces. For example, in forecasting growth rates in school enrollments at a certain school, use data from similar schools. To forecast sales of products, use sales of similar products sold to similar customers.

**Purpose:** To improve accuracy by improving the reliability of parameter estimates.

**Conditions:** Pooling is especially important when data for time series are highly variable, are intermittent, consist of small samples, or contain outliers.

**Strength of evidence:** Received wisdom. This principle is intuitively pleasing, and it is probably used in many organizations. Weak empirical evidence.

**Source of evidence:** Duncan, Gorr and Szczypula (2001) summarize limited evidence.

**7.7 Ensure consistency with forecasts of related series and related time periods.**

**Description:** If the plan depends upon a set of forecasts, these forecasts should be consistent with one another. This is a basic principle behind input-output forecasting. Some series are systematically related to others (e.g., cars need four wheels, so forecasts of cars and wheels should be related). Or, if the quarterly forecasts indicate that sales will go down over the next four quarters while an annual forecast shows an increase, the differences must be reconciled.

**Purpose:** To improve the accuracy and use of forecasts.

**Conditions:** Consistency is important when plans depend upon forecasts for related items and when the data are unreliable.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

### *8. Implementing Judgmental Methods*

Some principles for forecasting concern only judgmental methods. In general, you need to ask the right people the right questions at the right time.

**8.1 Pretest the questions you intend to use to elicit judgmental forecasts.**

**Description:** Prior to data collection, questions should be tested on a sample of potential respondents to ensure that they are understood and that they relate to the objectives of the problem.

**Purpose:** To improve accuracy.

**Conditions:** Applies to data collection for any type of judgmental forecasts unless good questions were used previously and unless it is important to have consistency across time.

**Strength of evidence:** Received wisdom.

**Source of evidence:** This principle is based on standard procedures for survey research.

**8.2 Frame questions in alternative ways.**

**Description:** The way the question is framed can affect the forecast. Sometimes even small changes in wording lead to substantial changes in responses. Consider alternatives such as asking for forecasts of unit changes and of percentage changes. Provide experts with different background data and summarize using graphs and tables.

**Purpose:** To improve accuracy by compensating for possible biases in the wording.

**Conditions:** Important when response errors are likely to be substantial and alternative framing is sensible.

**Strength of evidence:** Received wisdom and substantial evidence.

**Source of evidence:** This principle is based on standard procedures for survey research. Morwitz (2001) describes evidence on the effects of alternative wording for intentions questions. Armstrong (1985, pp. 96-108) summarized evidence related to judgmental forecasting.

**8.3 Ask experts to justify their forecasts in writing.**

**Description:** Experts should provide written support showing the reasons supporting their forecasts.

**Purpose:** To improve accuracy and learning.

**Conditions:** For expert opinion studies.

**Strength of evidence:** Some empirical support.

**Source of evidence:** Although this is a common practice, its effects on accuracy are speculative. Arkes (2001) summarizes evidence showing that justification improves calibration; from this, one might infer some gains in forecast accuracy.

**8.4 Provide numerical scales with several categories for experts' answers.**

**Description:** In general, use as many categories as seems reasonable. For example, to assess purchase intentions, use 11-point scales.

**Purpose:** To improve forecast accuracy.

**Conditions:** Use many scale points whenever it does not look odd to do so.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Morwitz (2001) summarizes some mixed evidence.

**8.5 Obtain forecasts from heterogeneous experts.**

**Description:** Experts should vary in their information and in the way they approach the problem.

**Purpose:** To improve accuracy by incorporating more information.

**Conditions:** Use for opinion studies, especially when experts might be subject to different biases.

**Strength of evidence:** Received wisdom. This principle is obvious but it is often violated.

**Source of evidence:** Rowe and Wright (2001) summarize supporting evidence. Also see Batchelor and Dua (1995).

**8.6 Obtain intentions or expectations from representative samples.**

**Description:** For example, to determine whether people will purchase cars, ask a representative sample of potential car buyers.

**Purpose:** To improve accuracy when generalizing to the entire population.

**Conditions:** This principle applies to expectations and intentions studies and, to some extent, to role playing, but *not* to surveys of expert opinion. This principle is especially important when the target population contains segments that differ substantially with respect to the behavior being forecasted.

**Strength of evidence:** Common sense and anecdotal evidence.

**Source of evidence:** Failure to follow this principle is commonly thought to have caused the incorrect forecast of the outcome of the Roosevelt-Landon presidential election in 1936. (According to Squire, 1988, however, non-response bias was a more important source of error in that election poll.)

**8.7 Obtain forecasts from enough respondents.**

**Description:** Larger samples are always preferred in term of accuracy, but costs must be considered.

**Purpose:** To improve forecast accuracy.

**Conditions:** This applies to expert opinion studies and to intentions studies. You need only a few experts (between 5 and 20), but many participants for intentions studies, often 500 or more.

**Strength of evidence:** Received wisdom.

**Source of evidence:** The benefits of large sample sizes can be overestimated. Lau (1994) showed that sampling error was small relative to other types of errors in predicting the outcomes of political elections from polls.

**8.8 Obtain multiple forecasts of an event from each expert.**

**Description:** Ask experts to make forecasts and then repeat the process some days later. This is an important aspect of the Delphi technique, as described in Rowe and Wright (2001).

**Purpose:** To improve forecast accuracy.

**Conditions:** In studies of expert opinion, multiple forecasts are especially useful if the experts gain access to additional information after making their first forecast.

**Strength of evidence:** Some empirical evidence.

**Source of evidence:** Rowe and Wright (2001) found evidence from four studies showing that additional rounds from Delphi panels improved accuracy.

## *9. Implementing Quantitative Methods*

**9.1 Tailor the forecasting model to the horizon**.

**Description:** Short-term models should place a heavy emphasis on the latest observations and long-term models should rely on long-term trends.

**Purpose:** To improve forecast accuracy.

**Conditions:** It is important to select a method appropriate to the horizon, especially if the forecast covers a long forecast horizon.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Armstrong, Adya and Collopy (2001) and Armstrong (2001f) summarize the limited evidence.

**9.2 Match the model to the underlying phenomena.**

**Description:** This issue typically requires domain knowledge. It calls for addressing such questions as: To what extent should the process represent the actual situation? For example, is a time-series process best represented by additive or multiplicative relationships?

**Purpose:** To improve forecast accuracy and gain acceptance of forecasts.

**Conditions:** This applies only when there is knowledge about the phenomena. It is important to identify features of the data.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Armstrong, Adya and Collopy (2001) present evidence on the value of using domain knowledge to select extrapolation methods.

**9.3 Do not use “fit” to develop the model.**

**Description:** The ability to fit (explain) historical data is a poor basis for selecting variables, specifying relationships, or selecting functional forms. The dangers of improper use of fit statistics often outweigh their benefits. Instead use domain knowledge and theory to specify the model.

**Purpose:** To improve forecast accuracy.

**Conditions:** This principle is especially important for time-series data in which spurious relationships are common. If you cannot assess forecast validity outside of the data used to estimate the model, you might (cautiously) use fit as a last resort. However, fit can be useful for cross-sectional data.

**Strength of evidence:** Strong empirical evidence that refutes received wisdom.

**Source of evidence:** There is little empirical evidence supporting the use of fit in time-series forecasting. Armstrong (2001c) summarizes evidence from many studies.

**9.4 Weight the most relevant data more heavily.**

**Description:** For time series, the most recent data are typically, though not always, most relevant and thus deserving a heavier weight. For cross-sectional data, domain expertise may be needed to identify cases that are most relevant to the forecast situation.

**Purpose:** To improve forecast accuracy.

**Conditions:** It is important to use the most recent data when large changes have occurred or are expected. Also, the measurement errors should be small and forecast horizons should be short.

**Strength of evidence:** Received wisdom and strong empirical evidence.

**Source of evidence:** Armstrong (2001f) summarizes evidence.

**9.5 Update models frequently.**

**Description:** Revise parameters as information is obtained. In particular, ensure that the estimate of the level in a time series is up to date.

**Purpose:** To improve forecast accuracy.

**Conditions:** This principle is important when there are large recent changes, when relationships are likely to change, and when relationships are subject to much uncertainty.

**Strength of evidence:** Received wisdom with some evidence.

**Source of evidence:** Armstrong (2001f) summarizes evidence that updating parameters in extrapolation models improves accuracy.

## *10. Implementing Methods: Quantitative Models with Explanatory Variables*

Explanatory (or causal) models show how policies (e.g., different prices, different advertising campaigns, or new laws) affect forecasts. The primary methods for policy analysis are judgmental bootstrapping, conjoint analysis, expert systems, and econometric methods. Use a policy variable in the model when (1) there is a strong causal relationship, (2) it is possible to estimate the relationship, (3) the policy variable will change substantially over the forecast horizon, and (4) it is possible to accurately forecast (or control) changes in the policy variable. Condition 4 can be omitted if one is developing contingency plans; even if one cannot forecast the changes, it would be useful to forecast what would happen *if* a variable changed.

**10.1 Rely on theory and domain expertise to select causal (or explanatory) variables.**

**Description:** Avoid irrelevant variables. Do not use statistical significance in selecting key variables; specifically, do not use stepwise regression. Mosteller and Tukey's (1977, pp. 270-271) advice was to choose variables that are reasonably presentable and will avoid hilarious newspaper columns or the appearance of injustice.

**Purpose:** To improve forecast accuracy.

**Conditions:** This principle assumes that you have information about the expected relationships over the forecast horizon. It can be based on domain knowledge or on previous studies.

**Strength of evidence:** Strong empirical support. This principle has been challenged (with little success) in the past two decades by researchers who use data-mining techniques.

**Source of evidence:** Allen and Fildes (2001), Armstrong (1970), Dawes and Corrigan (1974), Makridakis (1996), and McCloskey and Ziliak (1996) summarize empirical evidence.

**10.2 Use all important variables.**

**Description:** Presumably, you would obtain a list of important variables from domain experts and from prior research. This principle must be balanced against the principle of simplicity (#7.1). Armstrong (1985, p. 198) summarizes research from three studies showing that econometric models with few variables (2 or 3) are likely to be adequate. However, you can incorporate additional information by decomposing the problem, integrating information from alternative sources, using domain knowledge to estimate relationships, or combining forecasts.

**Purpose:** To improve the use of forecasts.

**Conditions:** Applies to cases that involve policy analysis and to situations involving large changes.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Allen and Fildes (2001), Armstrong (2001b), and Wittink and Bergestuen (2001) summarize evidence.

**10.3 Rely on theory and domain expertise when specifying directions of relationships.**

**Description:** Academics and practitioners often violate this principle by their willingness to let the data "speak for themselves." Data-mining techniques are popular but ill-suited for forecasting.

**Purpose:** To improve forecast accuracy.

**Conditions:** This principle assumes that information exists about the expected relationships over the forecast horizon. It can be based on domain knowledge or previous studies.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Empirical evidence is summarized in Allen and Fildes (2001), Armstrong (1985), and Dawes and Corrigan (1974).

**10.4 Use theory and domain expertise to estimate or limit the magnitude of relationships.**

**Description:** Sometimes there are physical limits to relationships between the dependent and explanatory variables. Other times there are well-established relationships based on prior research.

**Purpose:** To improve forecast accuracy.

**Conditions:** This principle assumes that knowledge exists about the magnitude of relationships over the forecast horizon and that you are aware of possible limits.

**Strength of evidence:** Some empirical evidence.

**Source of evidence:** Dawes and Corrigan (1974) provide indirect support and Allen and Fildes (2001) review other studies.

**10.5 Use different types of data to measure a relationship.**

**Description:** Obtain data of different types such as cross-sectional data, time-series data, and longitudinal data. For example, estimates of income elasticity could be obtained from data from households, states, or countries.

**Purpose:** To improve forecast accuracy by using more information.

**Conditions:** It is useful to make alternative estimates when there is uncertainty about the magnitudes of the relationships and when large changes are expected in the causal variables.

**Strength of evidence:** Received wisdom and weak empirical support.

**Source of evidence:** Armstrong (1985, pp. 205-217) showed that combining alternative estimates improved forecast accuracy in four of the five studies found.

**10.6 Prepare forecasts for at least two alternative environments.**

**Description:** Prepare forecasts for different assumptions about the uncontrollable elements. For example, what would be the forecast if the environment became unfavorable? How would this compare with forecasts from the most likely environment?

**Purpose:** To improve the use of forecasts by helping decision makers to develop contingency plans for alternative environments.

**Conditions:** This principle is important for situations with potentially large environmental changes.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**10.7 Forecast for alternate interventions.**

**Description:** This pertains to controllable elements, the what-if-we-did-x issues. Anecdotal evidence indicates that people using unaided judgment and traditional group meetings are poor at comparing alternatives.

**Purpose:** To improve decisions by using forecasts to make systematic, accurate, and consistent comparisons of alternate strategies.

**Conditions:** This applies in situations where forecasts can guide decisions about which policy to pursue. It is especially important to do this when the future policies differ substantially from past policies.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**10.8 Apply the same principles to forecasts of explanatory variables.**

**Description**: The forecast accuracy of an explanatory model relies upon being able to make reasonably accurate forecasts of the explanatory variables. To do this, apply the same principles as noted in this section.

**Purpose**: To improve forecast accuracy.

**Conditions**: This principle is applicable when the explanatory variables are expected to change substantially and where there are reasonably good estimates of the relationships. Early studies indicated that this principle was not important in many situations. In a review of 13 studies that compared the accuracy of unconditional forecasts with conditional forecasts (where the explanatory variables are known), the conditional forecasts were more accurate in only two studies and less accurate in 10 studies, with one tie (Armstrong, 1985, pp. 241-242). All but one of these studies involved short-term forecasts. However, more recent and more extensive evidence, summarized by Allen and Fildes (2001), shows that conditional forecasts are more accurate. There is also a well-established finding that forecast errors increase as the forecast horizon increases, partly because it becomes more difficult to forecast the causal variables.

**Strength of Evidence**: Weak.

**Source of Evidence**: Received wisdom.

**10.9 Shrink the forecasts of change if there is high uncertainty for predictions of the explanatory variables.**

**Description**: One should be cautious in forecasting change when it is difficult to forecast or control explanatory variables. One way to compensate is to shrink the forecasts of the changes in the explanatory variables. Shrinking can also be achieved by reducing the magnitude (absolute value) of the estimated relationship. Regression models shrink the estimates to adjust for uncertainty in the calibration data, but they ignore uncertainty about the forecasts of explanatory variables.

**Purpose**: To improve forecast accuracy.

**Conditions**: This is important when there is uncertainty associated with future values of the explanatory variables and when large changes are expected.

**Strength of Evidence**: Weak.

**Source of Evidence**: Armstrong (1985, pp. 240) reported gains in accuracy due to shrinkage. This principle is consistent with the use of damped trends, which has produced accurate extrapolations in situations involving uncertainty.

## *11. Integrating Judgmental and Quantitative Methods*

Judgmental information can be combined with quantitative methods in many ways.

**11.1 Use structured procedures to integrate judgmental and quantitative methods.**

**Description:** Use prespecified rules to integrate judgment and quantitative approaches. In practice, analysts often violate this principle.

**Purpose:** To improve accuracy.

**Conditions:** This principle is relevant when you have useful information that is not incorporated in the quantitative method. Whether to integrate will depend on the data, types of methods, and experts' information.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Armstrong and Collopy (1998) describe various procedures for integrating information and summarize evidence from 47 empirical studies.

**11.2 Use structured judgment as inputs to quantitative models.**

**Description:** Use judgment as inputs to a model rather than revising the models' forecasts.

**Purpose:** To improve accuracy.

**Conditions:** This principle is important when the model would not otherwise include judgmental knowledge. The use of this information as an input rather than to revise the forecasts is especially important when forecasts are subject to biases, as for example, in forecasts on the effects of new governmental social programs.

**Strength of evidence:** There is some empirical support and it challenges received wisdom.

**Source of evidence:** Armstrong and Collopy (1998) infer this principle from extensive research on integration.

**11.3 Use prespecified domain knowledge in selecting, weighting, and modifying quantitative methods.**

**Description:** Decide how to select and weight forecasting methods prior to making the forecasts.

**Purpose:** To improve forecast accuracy.

**Conditions:** Relevant when some domain knowledge has not been included and when there is little potential for bias by the forecaster.

**Strength of evidence:** Some empirical support.

**Source of evidence:** Armstrong, Adya and Collopy (2001) provide evidence from a series of studies.

**11.4 Limit subjective adjustments of quantitative forecasts.**

**Description:** Subjective adjustments should be limited to situations in which you have domain knowledge that is independent of the model.

**Purpose:** To improve forecast accuracy.

**Conditions:** Subjective adjustments are most appropriate for short-term forecasts and when unbiased experts have additional information. This is likely to apply to levels, but it could also apply to trends.

**Strength of evidence:** Received wisdom. For example, financial auditors are skeptical of forecasts that incorporate large subjective revisions according to Danos and Imhoff (1983). In addition, there is strong empirical support.

**Source of evidence:** Goodwin and Fildes (1998), McNees (1990), Sanders and Ritzman (2001), and Webby et al. (2001).

**11.5 Use judgmental bootstrapping instead of expert forecasts.**

**Description:** Use a model that infers the experts' rules.

**Purpose:** To improve forecast accuracy and consistency and to reduce costs of forecasting.

**Conditions:** Use bootstrapping when good data are not available for the dependent variable, when many expert forecasts are needed, or when comparing forecasts for alternative policies.

**Strength of evidence:** This principle is counterintuitive and seldom used. Strong empirical evidence refutes received wisdom.

**Source of evidence:** Armstrong (2001b) provides strong evidence showing that judgmental bootstrapping is more accurate than experts' forecasts in eight studies, less accurate in one, and there were two ties.

## *12. Combining Forecasts*

By combining forecasts, you can incorporate more information than you could with one forecast. Combining also reduces risk due to effects of bias associated with a single method. Armstrong (2001e) summarizes evidence from 30 empirical comparisons. Combining always reduced the error from the typical method. The average error reduction was 12.5 percent.

**12.1 Combine forecasts from approaches that differ.**

**Description:** Use forecasts drawn from different methods or data.

**Purpose:** To improve forecast accuracy.

**Conditions:** The situation must permit the use of more than one reasonable forecasting method. Combining independent methods produces greater benefits than combining similar ones, but even similar methods produce gains in accuracy.

**Strength of evidence:** Received wisdom. Some empirical support.

**Source of evidence:** Armstrong (2001e) summarizes evidence from two studies.

**12.2 Use many approaches (or forecasters), preferably at least five.**

**Description:** The gain from adding more than five approaches decreases rapidly while costs increase.

**Purpose:** To improve accuracy.

**Conditions:** The situation must permit a range of reasonable approaches from which to choose.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Armstrong (2001e) summarizes evidence from six studies.

**12.3 Use formal procedures to combine forecasts.**

**Description:** Specify the combining procedures before preparing the forecasts.

**Purpose:** To improve forecast accuracy.

**Conditions:** Formal procedures are important when some outcomes may be undesirable to the forecaster.

**Strength of evidence:** Some empirical evidence.

**Source of evidence:** Armstrong (2001e) summarizes evidence from five studies.

**12.4 Start with equal weights.**

**Description:** Equal weighting of forecasts is best in many situations.

**Purpose:** To improve forecast accuracy.

**Conditions:** Starting with equal weights is important when you are uncertain about the situation (low domain knowledge) or about the best forecasting method.

**Strength of evidence:** Some empirical evidence.

**Source of evidence:** Armstrong (2001e) provides support based on three studies.

**12.5 Use trimmed means.**

**Description:** Discard the highest and lowest forecasts, and then average the remaining forecasts.

**Purpose:** To avoid large errors.

**Conditions:** To use trimming, you should have at least five reasonable approaches. Trimmed means are especially important when large forecast errors are likely.

**Strength of evidence:** Weak empirical support.

**Source of evidence:** Armstrong (2001e) summarizes three studies that provide indirect support.

**12.6 Use track records to vary the weights on component forecasts.**

**Description:** Evidence on comparative accuracy of methods can be obtained in a given situation. For example, earlier periods of time series can be used for assessing ex ante forecast validity in a hold-out sample. Do not weight forecasts by the inverse of the variance of their errors because this method is unreliable.

**Purpose:** To improve forecast accuracy.

**Conditions:** Substantial evidence is needed as to the relative accuracy of the methods.

**Strength of evidence:** Some empirical support.

**Source of evidence:** Armstrong (2001e) summarizes four studies, all providing support.

**12.7 Use domain knowledge to vary weights on component forecasts.**

**Description:** Ask domain experts, preferably a number of them, to assign weights to component forecasts.

**Purpose:** To improve forecast accuracy.

**Conditions:** The experts must have good domain knowledge and they should not be subject to obvious biases.

**Strength of evidence:** Weak empirical support.

**Source of evidence:** Armstrong, Adya and Collopy (2001) summarize some evidence.

**12.8 Combine forecasts when there is uncertainty about which method is best.**

**Description:** Combining helps as long as each component has some predictive validity.

**Purpose:** To improve forecast accuracy.

**Conditions:** Combining improves accuracy (in comparison with typical forecasts) under nearly all conditions.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Armstrong's (2001e) review found little evidence**:** one study provided indirect support.

**12.9 Combine forecasts when you are uncertain about the situation.**

**Description:** When there is much uncertainty in the situation to be forecast, combining is of potentially greater value.

**Purpose:** To improve forecast accuracy.

**Conditions:** Combine forecasts when there is uncertainty about what happened in the past and what might happen in the future.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Armstrong (2001e) found supporting evidence in all six studies relevant to this principle.

**12.10 Combine forecasts when it is important to avoid large errors.**

**Description:** The accuracy from equally weighted combined forecasts can be no more than average error of the components.

**Purpose:** To reduce the likelihood of large forecast errors.

**Conditions:** Applies in situations where large errors have extreme costs such as those leading to war, bankruptcy or death.

**Strength of evidence:** Common sense. Equally weighted combined forecasts cannot be less accurate than the typical component.

**Source of evidence:** Not relevant.

## EVALUATION OF FORECASTING METHODS

When many forecasts are needed, you should compare alternative methods. The comparison should include accuracy and other criteria. Among these other criteria, it is of particular importance to properly assess uncertainty.

### *13. Evaluating Methods*

The principles for evaluating forecasting methods are based on generally accepted scientific procedures.

**13.1 Compare reasonable methods.**

**Description:** Use at least two methods, preferably including the current procedure as one of these. Exclude methods that unbiased experts would consider unsuitable for the situation.

**Purposes:** To select the best method and improve methods**.**

**Conditions:** Whenever biases can affect the evaluation (which is often). Knowledge of alternative approaches is helpful.

**Strength of evidence:** Some empirical support.

**Source of evidence:** Armstrong, Brodie and Parsons (2001) provide evidence showing that this principle reduces a researcher's biases.

**13.2 Use objective tests of assumptions.**

**Description:** Use quantitative approaches to test assumptions.

**Purposes:** To select the best method and to improve methods.

**Conditions:** Tests are relevant for important assumptions, assuming that you can obtain objective assessments. This is relevant only for cases where you are uncertain about the validity of the assumptions.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**13.3 Design test situations to match the forecasting problem.**

**Description:** Test forecasting methods by simulating their use in making actual forecasts. For example, to assess how accurate a model is for five-year-ahead forecasts, test it for five-year-ahead out-of-sample (ex ante) forecasts. (This is related to Principle 6.7.)

**Purposes:** To select the best method and improve methods.

**Conditions:** You need knowledge of alternative approaches and the situation.

**Strength of evidence:** Some empirical support.

**Source of evidence:** Armstrong (2001c) summarizes evidence, much of which comes from studies of personnel selection.

**13.4 Describe conditions associated with the forecasting problem.**

**Description:** Ideally, these conditions will be similar to those in other situations, allowing for a comparison of the present situation with others.

**Purpose:** To apply appropriate methods for new situations.

**Conditions:** Whenever you need to generalize to new situations.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**13.5 Tailor the analysis to the decision.**

**Description:** Often the proper analytic procedure will be obvious, but not always.

**Purpose:** To ensure proper use of forecasts.

**Conditions:** This is relevant when it is not immediately obvious how to compare forecasting methods. Armstrong (2001c) describes situations in which it is not clear how to analyze the information.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**13.6 Describe potential biases of forecasters.**

**Description:** Describe biases that might affect forecasters or their methods.

**Purpose:** To select the best method and to improve methods.

**Conditions:** Adjust for biases especially when the forecasting process relies on judgment.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Armstrong (2001c) summarizes evidence from studies of government revenue forecasts, political polls, and government deregulation.

**13.7 Assess the reliability and validity of the data.**

**Description:** Provide quantitative assessments of validity and reliability.

**Purpose:** To improve forecast accuracy.

**Conditions:** It is important to assess data quality when forecasting the effects of alternative policies. Armstrong (2001c) discusses a study that concluded that increases in the minimum wage would help unskilled workers. However, the study had serious problems with the reliability of the data and a reanalysis with corrected data reversed the findings.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**13.8 Provide easy access to the data.**

**Description:** If the data are easily available, replications can be done**.** Given the evidence on the difficulty of replicating findings in management science (Armstrong 2001c), the principle is important.

**Purpose:** To reliably assess the accuracy of alternative methods.

**Conditions:** Full access to data is particularly important when forecasts might be affected by biases. Sometimes, reanalysis of data yields different results. Websites can now make full disclosure of data inexpensive. For example, data from the M-Competitions are available on the Forecasting Principles website.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**13.9 Provide full disclosure of methods.**

**Description:** Detailed descriptions of the methods can allow others to audit forecasting methods and to replicate them. Whereas full disclosure used to be expensive due to limited space in journals, it can now be accomplished by putting methodological details on websites.

**Purposes:** To select the best method and improve methods.

**Conditions:** Full disclosure is most important when the methods require judgmental inputs or when the methods are new to the situation.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Armstrong (2001c) provides evidence on the value of this principle.

**13.10 Test assumptions for validity.**

**Description:** Provide quantitative assessments of the validity of the assumption. This includes face, construct, and predictive validity.

**Purpose:** To assess the accuracy of forecasts.

**Conditions:** This is important when comparing the effects of proposed alternative policies.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**13.11 Test the client=s understanding of the methods.**

**Description:** A method that is easy to understand might be preferable even if it reduces accuracy. In practice, the clients often do not understand the methods. This principle is related to Principle 1.5 (obtain agreement on methods).

**Purposes:** To select the most appropriate forecasting method and to increase the likelihood that it will be used properly. For example, the client should be able to identify when the methods need to be revised.

**Conditions:** It is important to understand the methods if key aspects of the problem are likely to change.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**13.12 Use direct replications of evaluations to identify mistakes.**

**Description:** By redoing evaluations, one can check for mistakes. Researchers have replicated the M-Competition studies and have identified some mistakes. (However, these mistakes did not alter the conclusions.)

**Purpose:** To check for mistakes in comparisons of methods.

**Conditions:** Replication is especially useful for complex methods and when forecasts might be affected by biases.

**Strength of evidence:** Weak evidence.

**Source of evidence:** Armstrong (2001c) reviews four studies showing that mistakes occur often in forecasting.

**13.13 Replicate forecast evaluations to assess their reliability.**

**Description:** Replications provide the best way to assess reliability. However, replications are seldom used in management science (Hubbard & Vetter 1996).

**Purpose:** To obtain reliable comparisons of alternative forecasting methods.

**Conditions:** Replication is especially important when the data are likely to be unreliable, biases are likely, and when forecast errors can have serious consequences.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**13.14 Use extensions of evaluations to better generalize about what methods are best for what situations.**

**Description:** This involves replications that contain variations in important elements of the situation or method.

**Purpose:** To ensure use of the proper forecasting methods.

**Conditions:** Extensions are important when you expect to use the forecasting procedure for a wide range of problems.

**Strength of evidence:** Some indirect empirical support.

**Source of evidence:** Hubbard and Vetter (1996), in their review of published extensions in accounting, economics, finance, management, and marketing, found that 46 percent of the findings differed from those in the original study.

**13.15 Conduct extensions of evaluations in realistic situations.**

**Description:** When evaluating alternative forecasting methods, do so in situations that provide realistic representations of the forecasting problem.

**Purpose:** To ensure use of the proper forecasting methods.

**Conditions:** This is important when a situation involves large changes and when forecast errors have serious consequences.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**13.16 Compare forecasts generated by different methods.**

**Description:** Comparisons of forecasts from different methods can be used to examine forecast accuracy and to assess uncertainty. Armstrong (2001c) discusses this issue.

**Purpose:** To ensure use of the proper forecasting methods.

**Conditions:** This principle applies when the situation permits the use of multiple methods. It is especially useful when methods differ substantially.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**13.17 Examine all important criteria.**

**Description:** Yokum and Armstrong (1995) describe various criteria along with ratings of their importance by decision makers, practitioners, educators, and researchers.

**Purposes:** To improve acceptance of the proposed methods and to ensure that they meet the needs of the decision makers.

**Conditions:** Good knowledge of the problem is needed in order to evaluate all important criteria (e.g., accuracy, ability to assess uncertainty, cost). The importance of criteria varies by conditions (e.g., long term vs. short term) and by methods (e.g., extrapolation vs. econometric methods). This principle is especially important when biases are likely.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**13.18 Specify criteria for evaluating methods prior to analyzing data.**

**Description:** List the criteria in order of importance before analyzing the data.

**Purpose:** To help in selecting proper forecasting methods.

**Conditions:** This is important when different methods yield substantially different forecasts, when judgmental inputs are important, or when biases may have a strong influence.

**Strength of evidence:** Some empirical support.

**Source of evidence:** Armstrong (2001c) summarizes evidence on the need to prespecify criteria.

**13.19 Assess face validity.**

**Description:** Face validity involves asking whether the evaluation study makes sense to independent unbiased experts. Assess face validity in a structured way (e.g., by using questionnaires) to obtain expert opinions.

**Purpose:** To ensure the use of the proper methods and to gain acceptance of the forecasts.

**Conditions:** Face validity is important when large changes are expected.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**13.20 Use error measures that adjust for scale in the data.**

**Description:** Ensure that the comparisons among methods are not distorted by one series having larger numbers than other series and thus being weighted more heavily. One can use error measures that are expressed as percentages to adjust for scale.

**Purpose:** To help ensure the use of the proper forecasting methods.

**Conditions:** When comparing across different situations (e.g., across different time series), you need error measures that are not unduly influenced by a small number of series. This is important when dealing with heterogeneous time series.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**13.21 Ensure error measures are valid.**

**Description:** Error measures should relate to the decision being made, such as to determine which is the most accurate method.

**Purpose:** To help ensure use of the proper forecasting methods.

**Conditions:** In general, evaluation studies should be concerned with the validity of the error measures.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**13.22 Use error measures that are not sensitive to the degree of difficulty in forecasting.**

**Description:** This principle prevents the evaluation from being dominated by a few series that have very large forecast errors. Apply this principle when some series are subject to large changes. Ohlin and Duncan (1949) identified the need for this principle. Relative absolute errors (RAE) compensate somewhat for differences in the difficulty of forecasting series (Armstrong & Collopy 1992).

**Purpose:** To properly assess the relative accuracy of different methods.

**Conditions:** This principle applies only when generalizing across time series that vary in their forecasting difficulty.

**Strength of evidence:** Common sense.

**Source of evidence:** None

**13.23 Avoid biased error measures.**

**Description:** Do not use an error measure favoring forecasts that are systematically high (or low). Armstrong (2001c) describes this issue and how to resolve it.

**Purpose:** To properly assess relative accuracy.

**Conditions:** This applies when one needs to assess forecasts that cover a wide range of values and is especially relevant for non-negative time series.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**13.24 Avoid error measures that are highly sensitive to outliers.**

**Description:** Armstrong (2001c) describes error measures that offer protection against the effects of outliers.

**Purpose:** To properly assess relative accuracy.

**Conditions:** This principle is only needed when outliers are likely. However, if it is the outliers that are of concern, such as hurricanes or floods, ignore this principle.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**13.25 Use multiple measures of accuracy.**

**Description:** Armstrong (2001c) describes a variety of error measures.

**Purpose:** To properly assess the relative accuracy of alternative forecasting methods.

**Conditions:** Use multiple measures when there is uncertainty about the best error measure.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Armstrong (2001c) shows how evaluations of alternative forecasting methods can differ depending upon the error measure chosen.

**13.26 Use out-of-sample (ex ante) error measures.**

**Description:** Conditional (ex post) error are not closely related to ex ante errors.

**Purpose:** To properly assess the relative accuracy of forecasting methods.

**Conditions:** Ex ante error measures are especially important for time series that include moderate to large changes.

**Strength of evidence:** Strong empirical evidence supports this principle, which conflicts with common practice and with recommendations by statisticians.

**Source of evidence:** Armstrong (2001c) summarizes evidence from six studies.

**13.27 Use ex post error measures to evaluate the effects of policy variables.**

**Description:** Assuming that changes in the explanatory variables were correctly forecast, how well does the model predict the effects of policy changes?

**Purpose:** To determine how effectively methods can forecast the outcomes of policy changes (e.g., to examine the effects of different price levels for a product).

**Conditions:** Ex post tests are important when decision makers want to access the outcomes of alternative policies, such as when using econometric models. In addition, ex post tests help improve econometric models by showing the sources of error.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**13.28 Do not use R-square (either standard or adjusted) to compare forecasting models.**

**Description:** R-square ignores bias and it has little relationship to decision-making.

**Purpose:** To avoid improper evaluation of the accuracy of methods.

**Conditions:** R-square is a misleading measure for comparing time series models although it may have some relevance for cross-sectional data.

**Strength of evidence:** This principle is in conflict with received wisdom and there is some empirical evidence.

**Source of evidence:** Armstrong (2001c) describes the problems associated with the use of R-square.

**13.29 Use statistical significance only to compare the accuracy of *reasonable* methods.**

**Description:** Little is learned by rejecting an unreasonable null hypothesis. When comparing accuracy, adjust the significance level for the number of models that are compared when more than two models are involved.

**Purpose:** To avoid improper evaluation of the accuracy of forecasting methods.

**Conditions:** Statistical significance can be misleading in forecasting time series because of autocorrelation or outliers. It can be useful, however, in making comparisons of reasonable methods when one has only a small sample of forecasts.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Armstrong (2001c) describes studies showing the dangers of using statistical significance.

**13.30 Do not use root mean square errors (RMSE) to make comparisons among forecasting methods.**

**Description:** The RMSE is an unreliable measure for comparing forecasting methods.

**Purpose:** To avoid improper evaluation of the accuracy of methods.

**Conditions:** The RMSE is not needed in forecasting. More appropriate procedures exist. Using root mean squares can be especially misleading when you are dealing with heterogeneous time series.

**Strength of evidence:** There is strong empirical support, and it conflicts with received wisdom.

**Source of evidence:** Armstrong and Fildes (1995) summarize evidence on this issue.

**13.31 Base comparisons of methods on large samples of forecasts.**

**Description:** For time series, use many series, horizons, and origins. Try to obtain forecasting cases that are somewhat independent of one another. To the extent that they are not independent, use larger samples of forecasts. Armstrong (2001c) discusses how to expand the sample of forecasts.

**Purpose:** To assess the accuracy of alternative forecasting methods.

**Conditions:** Relevant primarily for time series. It must be possible to obtain many forecasts from similar situations.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Armstrong (2001c) summarizes evidence on the need for large samples.

**13.32 Conduct explicit cost-benefit analyses.**

**Description:** Examine the costs and benefits of each forecasting method.

**Purpose:** To select the most appropriate forecasting method.

**Conditions:** This is relevant when the cost of forecasting may exceed the potential benefits.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

## *14. Assessing Uncertainty*

In addition to forecasting the most likely outcomes, it is important to assess the confidence one should have in the forecast. To do this, use prediction intervals (PI). Armstrong (1988) reviewed the literature and found little research on estimating uncertainty, but the situation has improved in more recent years.

**14.1 Estimate prediction intervals (PIs).**

**Description:** Decision makers can often make better forecasts if they are aware of the risks. To assess risks, you could estimate, say, 95 percent for each forecast horizon in a time series. Rush and Page (1979), in a study on forecasting natural resources, showed that PIs are often ignored. Dalrymple's (1987) survey concluded that only about ten percent of firms "usually" use PIs for sales forecasts. Tull (1967), in a study of new product forecasting at 16 companies, found that twelve considered only point forecasts and four considered "optimistic" and "pessimistic" forecasts; none used PIs in this situation, which involved much uncertainty.

**Purpose:** To improve the use of forecasts.

**Conditions:** PIs are needed when decisions are affected by uncertainty, which means nearly always.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**14.2 Use objective procedures to estimate explicit prediction intervals.**

**Description:** Chatfield (2001) describes how to develop objective PIs. Judgmental estimates may be of some value, but they are likely to be biased.

**Purpose:** To improve assessment of PIs (i.e., to improve calibration).

**Conditions:** You need many comparisons of forecasts and actuals to estimate PIs. Judgmental estimates are appropriate when the forecaster receives excellent feedback, as in weather forecasting, or when there is knowledge of events that might affect the forecast.

**Strength of evidence:** Weak empirical evidence.

**Source of evidence:** This principle is inferred from findings related to point forecasts. Subjective procedures are often poorly calibrated (Arkes 2001), but this also occurs for quantitative PIs.

**14.3 Develop prediction intervals by using empirical estimates based on realistic representations of forecasting situations.**

**Description:** Makridakis and Winkler (1989) showed that the fit of a model to the calibration sample is a poor way to establish PIs in some situations. The preferred way to construct PIs is to use earlier holdout data to simulate the forecasting situation and to summarize ex ante errors for each forecast horizon. Forecasting software packages can aid this process.

**Purpose:** To improve the assessment of PIs.

**Conditions:** A number of observations are needed to develop reliable PIs. Furthermore, one should impose the constraint that uncertainty increases as the forecast horizon lengthens.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**14.4 Use transformations when needed to estimate symmetric prediction intervals.**

**Description:** To effectively estimate PIs, it is often important to transform the predicted and actual values to logs (in many cases, the distribution of errors will be asymmetric in the original units). This principle, while commonly recommended, is seldom used by academics or practitioners.

**Purpose:** To improve the assessment of PIs.

**Conditions:** Transformations are of particular concern for time series that have positive values only, large variations, trended data, heteroscedastic errors, or limits (such as with percentages).

**Strength of evidence:** Received wisdom.

**Source of evidence:** Armstrong and Collopy (2001) and Bolger and Harvey (1995) provide supporting evidence.

**14.5 Ensure consistency over the forecast horizon.**

**Description:** PIs should increase smoothly across the horizon.

**Purpose:** To improve the assessment of PIs.

**Conditions:** Adjustments are most important when one has few observed errors for some forecast horizons.

**Strength of evidence:** Common sense and some evidence.

**Source of evidence:** Smith and Sincich (1991), in their study of population forecasts, found that the MAPE grows linearly with the horizon. Makridakis et al. (1982, tables 12-14) show relatively consistent increases of MAPE over the forecast horizons for monthly, quarterly, and annual data.

**14.6 Describe reasons why the forecasts might be wrong.**

**Description:** Use the devil's advocate procedure with groups to elicit reasons why the forecasts might be wrong. In that procedure, one person tells other group members what is wrong with the forecasts, while the others defend their position. It is best if the analysis is written and avoids emotional attacks.

**Purpose:** To improve the assessment of PIs.

**Conditions:** Applies to judgmental and quantitative forecasting. This is important when estimates of uncertainty are subjective and when the forecasters are likely to be overconfident.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Arkes (2001) summarizes evidence from 16 studies. The procedure reduces, but does not eliminate, overconfidence.

**14.7 When assessing PIs, list possible outcomes and assess their likelihoods.**

**Description:** One reason for this principle is that just thinking about a possible outcome leads people to increase their estimates of its likelihood. Thus, it is desirable to examine a range of possible outcomes.

**Purpose:** To improve the assessment of PIs.

**Conditions:** Listing alternative outcomes is particularly important in assessing PIs for policy changes or large environmental changes.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Arkes (2001) describes six studies.

**14.8 Obtain good feedback about forecast accuracy and the reasons why errors occurred.**

**Description:** Feedback should be explicit, systematic, and frequent.

**Purpose:** To improve the assessment of PIs.

**Conditions:** Feedback is especially important for judgmental forecasting.

**Strength of evidence:** Strong.

**Source of evidence:** Arkes (2001) reviewed evidence from five studies.

**14.9 Combine prediction intervals from alternative forecasting methods.**

**Description:** Combine estimates obtained from methods such as judgment and from extrapolation, being careful to ensure that all the PIs are unconditional intervals.

**Purpose:** To improve the assessment of PIs.

**Conditions:** When risk has a strong influence on decision making.

**Strength of evidence:** Weak empirical evidence.

**Source of evidence:** Armstrong (2001e) summarizes the limited evidence.

**14.10 Use safety factors to adjust for overconfidence in the PIs.**

**Description:** Arkes (2001) and Chatfield (2001) show that judgmental and quantitative PIs are often too narrow.

**Purpose:** To improve the assessment of PIs so as to better manage risk.

**Conditions:** Safety factors are important when assessments of PIs are poor or when substantial changes are expected in the future.

**Strength of evidence:** Speculation.

**Source of evidence:** None.

**14.11 Conduct experiments to evaluate forecasts.**

**Description:** If the historical data do not vary, it may be possible to conduct experiments to assess the consistency of effects from changes in policy variables.

**Purpose:** To gain evidence that will help in estimating PIs.

**Conditions:** Experiments are often needed for forecasts of policy variables.

**Strength of evidence:** Received wisdom.

**Source of evidence:** None.

**14.12 Do not assess uncertainty in a traditional (unstructured) group meeting.**

**Description:** Groups are typically overconfident (their PIs are too narrow).

**Purpose:** To avoid poor assessments of PIs.

**Conditions:** This applies only to judgmental assessments of PIs.

**Strength of evidence:** Strong empirical support.

**Source of evidence:** Arkes (2001) summarizes results from six studies.

**14.13 Incorporate the uncertainty associated with the prediction of the explanatory variables in the prediction intervals.**

**Description**: Errors in forecasting the explanatory variable introduce uncertainty into regression-model forecasts. However, the standard least-squares PI ignores this uncertainty and consequently is too narrow. Ex ante PIs are better calibrated. They can be derived by simulation or, even better, estimated empirically from the out-of-sample forecast errors. The latter requires a reasonably long historical series.

**Purpose**: To better estimate PI s.

**Conditions**: This is not relevant when one uses ex ante empirical PIs to assess uncertainty, if the future values of the explanatory variables can be forecast as well as in the development of the empirical confidence intervals. Nor is it relevant in assessing ex post accuracy.

**Strength of Evidence**: Received wisdom and some evidence.

**Source of Evidence**: Tashman, Bakken and Buzas (2000).

## USING FORECASTS

### *15. Presenting Forecasts*

**15.1 Present forecasts and supporting data in a simple and understandable form.**

**Description:** Keep the presentation simple yet complete. For example, do not use insignificant digits because they imply false precision. Graphs are often easier to understand than tables. Present forecasts in units that are meaningful to the decision makers. For detailed suggestions, see Ehrenberg (1981), Bailar and Mosteller (1998), and Wilkinson et al. (1999).

**Purposes:** To improve decision makers' understanding of the forecasts and to reduce the likelihood of overconfidence.

**Conditions:** Clear presentations are especially important for forecasts on the effects of policy changes.

**Strength of evidence:** Common sense and some evidence.

**Source of evidence:** Wagenaar, Schreuder and Van der Heijden (1985), in a study on TV versus radio weather forecasts, found that the more elaborate TV forecasts did not increase the number of facts that could be recalled.

**15.2 Provide complete, simple, and clear explanations of methods.**

**Description:** Test your presentation about forecasting methods on a sample of clients.

**Purposes:** To improve the forecast's acceptability and use.

**Conditions:** Do this when the forecasts may be controversial and when large changes are forecasted. It is also important when the forecasting method is judgmental.

**Strength of evidence:** Common sense and one study.

**Source of evidence:** Weimann (1990) found that polling agencies that provided more complete explanations of their methods produced more accurate forecasts.

**15.3 Describe your assumptions.**

**Description:** Provide a written and detailed account of your assumptions. Fischhoff (2001) discussed the need to record assumptions.

**Purpose:** To help decision makers to assess the extent to which they can use the forecasts in other situations.

**Conditions:** A good understanding of assumptions is important in situations where assumptions may change over time.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

**15.4 Present prediction intervals.**

**Description:** PI s can help decision makers to understand how the forecasts might affect decisions, and can indicate the need for contingency plans.

**Purpose:** To help assess risk.

**Conditions:** When decisions depend on the risk involved, PIs should be an important part of the presentation, especially when there is high uncertainty. Decision makers should be willing to examine risk.

**Strength of evidence:** Common sense. In practice, however, many organizations resist using PI s.

**Source of evidence:** None.

**15.5 Present forecasts as scenarios.**

**Description:** Scenarios are stories of “what happened in the future.” They help decision makers take forecasts seriously. Decision makers often ignore forecasts that are unpleasant or unexpected, even if they are life-threatening. For example, Baker (1979) found that hurricane warnings sometimes do not affect behavior substantially. For warnings to be effective, people have to believe them and they must be able to respond effectively.

**Purpose:** To improve the use of forecasts by preparing decision makers for undesirable outcomes.

**Conditions:** When forecasts are surprising or unfavorable, ask decision makers to describe how they would act in situations implied by the forecasts.

**Strength of evidence:** Some empirical support.

**Source of evidence:** Gregory and Duran (2001) summarize research on using scenarios to increase the acceptability of forecasts.

## *16. Learning That Will Improve Forecasting Procedures*

Ideally, as forecasters gain experience in using forecasting procedures, the procedures should improve.

**16.1 Consider the use of adaptive forecasting models.**

**Description:** Adaptive models are those whose parameters are automatically revised in light of new information.

**Purpose:** To update the parameters of a model when the situation has changed. For example, causal forces might change from growth to decay.

**Conditions:** Adaptive models are expected to be important in rapidly changing environments where one has good domain knowledge. To develop adaptive methods, you should assess conditions.

**Strength of evidence:** Received wisdom. Weak empirical support.

**Source of evidence:** Many studies have been done on adaptive models. However, Armstrong (1985, p. 171) found only weak evidence to support this principle. These studies did not concern large changes nor do they consider the use of domain knowledge, so the tests do not cover the situations of primary interest.

**16.2 Seek feedback about forecasts.**

**Description:** Design procedures for soliciting feedback. Review the forecasting methods periodically and identify the reasons for large forecast errors.

**Purpose:** To improve forecasting procedures by learning how current procedures fell short.

**Conditions:** Especially relevant for judgmental forecasts.

**Strength of evidence:** Common sense.

**Source of evidence:** Arkes (2001) provides related evidence.

**16.3 Establish a formal review process for forecasting methods.**

**Description:** Include written forecasts and support. Obtain data to assess accuracy, other benefits, and costs. Prepare summaries showing the accuracy of forecasts and reasons for large errors. Monitor forecasts, adjustments, and accuracy. If forecasts are changed, keep records of when, why, and by whom. Assess unconditional (ex ante) and conditional (ex post) accuracy.

**Purpose:** Encourage learning to improve accuracy and calibration of PI s.

**Conditions:** Relevant only when learning can be translated into guidelines.

**Strength of evidence:** Received wisdom.

**Source of evidence:** Arkes (2001), Fischhoff (2001), and Harvey (2001) provide related discussions and some evidence.

**16.4 Establish a formal review process to ensure that forecasts are used properly.**

**Description:** Periodic assessments should be made to examine how the forecasts are being used.

**Purpose:** To improve the use of forecasts.

**Conditions:** This principle is important when forecasts are for large or unusual changes and when decision makers have strong prior views.

**Strength of evidence:** Common sense.

**Source of evidence:** None.

## AUDITING THE FORECASTING PROCEDURE

Managers should agree on an auditing process well in advance of the forecast review. They should then support the process. For example, they could provide forecasters with a checklist for standards and practices.

The Forecasting Standards Checklist at end the of this paper can help forecasters and decision makers examine the forecasting process. The checklist includes 16 areas with 139 principles. As a result, using it is no trivial matter. It includes a column to be checked for principles that are not applicable (NA) to the given situation. For example, some guidelines apply only to judgmental procedures. The checklist also includes a column labeled “?” to indicate principles that seem applicable, but for which information is lacking.

The checklist is intended to provide ideas on ways to improve the forecasting process. Forecasters should examine the checklist prior to developing forecasting procedures. To ensure objectivity, it may be sensible to have outsiders conduct an audit of the forecasting procedure. They could be forecasting experts from a different department in the same organization, or they could be specialists from outside of the organization.

To facilitate audits, forecasters should keep good records either in a notebook or in a computer log. People often argue about whether a forecasting method is reasonable, so a forecaster who keeps no record of the process of choosing a method might be accused of bias. A forecaster's notebook can protect the forecaster and the organization. Ideally, these records would be provided on a secure website.

**Legal Aspects**

The primary purposes of “Standards and Practices” are to help forecasters improve forecast accuracy, to better assess uncertainty, and to help decision makers to use the forecasts properly. They could also protect the forecaster. To my knowledge, no one has successfully sued a forecaster for making an incorrect forecast. However, some have successfully sued forecasters by showing that they did not adhere to best practice.

*Beecham vs. Yankelovich* illustrates some of the legal issues in forecasting. Beecham alleged that an inaccurate forecast for a new cold water detergent resulted in a $24 million loss. They claimed that Yankelovich used incorrect inputs to the forecasting models. In response, Yankelovich replied that Beecham failed to follow the marketing plan on which the forecasts were based because of changes in the advertising claims and reduced promotional expenses. This suit, which was settled out-of-court in September 1988, created much concern among research firms (*Adweek's Marketing Week*, December 7, 1987, pp. 1,4; *Business Week*, August 10, 1987, pp. 28).

An erroneous forecast of a severe drought in Yakima Valley in Washington caused farmers to undertake expensive actions with the cattle. The farmers took legal action against the U.S. Bureau of Reclamation to recover their losses. The government admitted making a mistake when they made an ill-advised subjective adjustment. However, it had been under no contractual agreement to provide a forecast. As a result, the court ruled against the farmers (*Schinmann vs. U.S*., 618 F. Supp. 1030, September 18, 1985), which was upheld by the appellate court (*Schinmann vs. U.S*., unpublished opinion, U.S. Court of Appeals for the Ninth Circuit).

In another case, 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 grounds that the National Weather Service was negligent in failing to repair a weather buoy that could 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, F.2d 199; 1st Cir., 1986). The issue was not the inaccurate forecast, but that the National Oceanic and Atmospheric Administration (NOAA) had failed to take reasonable steps to obtain accurate data. In addition, when it failed to obtain key information, NOAA did not notify users of this deficiency. The court ruled that there was no contractual requirement for the government to report on the process it used to make the forecast. Their ruling also implied that it is reasonable for forecasters to make tradeoffs between the cost of the forecast and its benefits. The reverse side of this is, if there is a contractual relationship, the forecaster should reveal the process.

In a British case, *Esso Petroleum vs. Mardon* (London, 1966 E. no. 2571), Mardon contracted with Esso to own and operate a gas station. A critical part of the negotiations was Esso’s forecast that the station would sell 200,000 gallons of gas annually by the third year. 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 forecast misrepresented the situation. Esso had originally forecast the 200,000 gallon figure under the assumption that the gas pumps would face the road. After a zoning hearing, it had to change the station’s design so that the pumps were not visible from the road. Despite this unfavorable change, Esso used the original 200,000 gallon forecast in drawing up its contract with Mardon. Mardon won; the court concluded that Esso misrepresented the facts in this situation.

These cases 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 involve uncertainty; making reasonable attempts to balance costs and benefits should provide forecasters with protection against lawsuits. Finally, forecasters can be held liable if it can be shown that they did not use reasonable practices to obtain forecasts, or if they intentionally used poor practices so as to bias the forecasts.

In addition to their use in legal cases, agreements on good standards of forecasting can 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, convention centers, and sports stadiums?

## SUMMARY

This chapter summarizes 139 forecasting principles that were drawn primarily from the *Principles of Forecasting Handbook*. The checklist at the end of this paper is designed to help forecasters and managers to systematically evaluate the forecasting processes they use. Given the complexity of forecasting, a structured evaluation procedure should help.

Future research is expected to produce new principles and to refine the existing ones. Additions and revisions to the principles will be provided on the principles website.

## REFERENCES

Allen, P. G. & R. Fildes (2001), "Econometric forecasting," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Arkes, H. R. (2001), "Overconfidence in judgmental forecasting," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Armstrong, J. S. (1970), "How to avoid exploratory research," *Journal of Advertising Research*, 10, 27-30. Full text at hops.wharton.upenn.edu/forecast

Armstrong, J. S. (1982), "Strategies for implementing change**:** An experiential approach," *Group and Organization Studies*, 7, 457-475. Full text at hops.wharton.upenn.edu/forecast

Armstrong, J. S. (1985), *Long-Range Forecasting**:** From Crystal Ball to Computer*, 2nd Edition. New York**:** John Wiley. Full text at hops.wharton.upenn.edu/forecast

Armstrong, J. S. (1988), "Research needs in forecasting," *International Journal of Forecasting*, 4, 449-465. Full text at hops.wharton.upenn.edu/forecast

Armstrong, J. S. (1998), "Are student ratings of instruction useful?" *American Psychologist*, 53, 1223-1224.

Armstrong, J. S. (2001a), "Role playing**:** A method to forecast decisions," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Armstrong, J. S. (2001b), "Judgmental bootstrapping**:** Inferring experts' rules for forecasting," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Armstrong, J. S. (2001c), "Evaluating forecasting methods," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Armstrong, J. S. (2001d), "Selecting forecasting methods," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Armstrong, J. S. (2001e), "Combining forecasts," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Armstrong, J. S. (2001f), "Extrapolation for time-series and cross-sectional data," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Armstrong, J. S., M. Adya & F. Collopy (2001), "Rule-based forecasting**:** Using judgment in time-series extrapolation," in J. S. Armstrong (ed.), *Principles of Forecasting*. Norwell, MA**:** Kluwer Academic Publishers.

Armstrong J.S. & J. Andress (1970), "Exploratory analysis of marketing data**:** Trees vs. regression," *Journal of Marketing Research*, 7, 487-492. Full text at hops.wharton.upenn.edu/forecast.

Armstrong, J. S., R. Brodie & A. Parsons (2001), "Hypotheses in marketing science**:** Literature review and publication audit," *Marketing Letters* (forthcoming).

Armstrong, J. S. & F. Collopy (1992), "Error measures for generalizing about forecasting methods**:** Empirical comparisons," *International Journal of Forecasting,* 8, 69-80. Full text at hops.wharton.upenn.edu/forecast.

Armstrong, J. S. & F. Collopy (1998), "Integration of statistical methods and judgment for time series forecasting**:** Principles from empirical research," in G. Wright & P. Goodwin (eds.), *Forecasting with Judgment.* Chichester, England**:** John Wiley, 263-293. Full text at hops.wharton.upenn.edu/forecast.

Armstrong, J. S. & F. Collopy (2001), "Prediction intervals for extrapolation of annual economic data**:** Evidence on asymmetry corrections," *Journal of Forecasting* (forthcoming).

Armstrong, J. S. & R. Fildes (1995), "On the selection of error measures for comparison among forecasting methods," *Journal of Forecasting,* 14, 67-71. Full text at hops.wharton.upenn.edu/forecast

Ash, J. C. K. & D. J. Smyth (1973), *Forecasting the United Kingdom Economy*. Farnborough, England.

Bailar, J. C. & F. Mosteller (1998), "Guidelines for statistical reporting in articles in medical journals," *Annals of Internal Medicine,* 108, 266-273.

Baker, E. J. (1979), "Predicting responses to hurricane warnings," *Mass Emergencies,* 4, 9-24.

Batchelor, R. & P. Dua (1995), "Forecaster diversity and the benefits of combining forecasts," *Management Science*, 41, 68-75.

Bolger, F. & N. Harvey (1995), "Judging the probability that the next point in an observed time series will be above or below a given value," *Journal of Forecasting*, 14, 567-607.

Burns, M. & J. Pearl (1981), "Causal and diagnostic inferences**:** A comparison of validity," *Organizational Behavior and Human Performance*, 28, 379-394.

Chatfield, C. (1995), *Problem Solving: A Statistician's Guide.* (2nd ed.) London**:** Chapman and Hall.

Chatfield, C. (2001), "Prediction intervals for time-series forecasting," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Claycamp, H. J. & L. E. Liddy (1969), "Prediction of new product performance**:** An analytical approach," *Journal of Market Research*, 6, 414-420.

Cowles, A. (1933), "Can stock market forecasters forecast?" *Econometrica*, 1, 309-324.

Dalrymple, D. J. (1987), "Sales forecasting practices," *International Journal of Forecasting*, 3, 379-391.

Danos, P. & E. A. Imhoff 1983), “Factors affecting auditor’s evaluations of forecasts,” *Journal of Accounting Research*, 21, 473-494.

Dawes, R. M. & B. Corrigan (1974), “Linear models in decision making,” *Psychological Bulletin*, 81, 95-106.

Dorn, H. F. (1950), “Pitfalls in population forecasts and projections,” *Journal of the American Statistical Association*, 45, 311-334.

Duncan, G. T., W. L. Gorr & J. Szczypula (2001), “Forecasting analogous time series,” in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA: Kluwer Academic Publishers.

Ehrenberg, A. S. C. (1981), “The problem of numeracy,” *The American Statistician*, 35, 27-70.

Fildes, R. & R. Hastings (1994), “The organization and improvement of market forecasting,” *Journal of the Operational Research Society*, 45, 1-16

Fischhoff, B. (2001), “Learning from experience: Coping with hindsight bias and ambiguity,” in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA: Kluwer Academic Publishers.

Gaeth, G. J. & J. Shanteau (1984), “Reducing the influence of irrelevant information on experienced decision makers,” *Organizational Behavior and Human Performance*, 33, 263-282.

Goodwin, P. & R. Fildes (1998), “Judgmental forecasts of time series affected by special events: Does providing a statistical forecast improve accuracy?” *Journal of Behavioral Decision Making*, 12, 37-53.

Gregory, W. L. & A. Duran (2001), “Scenarios and acceptance of forecasts,” in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA: Kluwer Academic Publishers.

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

Harvey, N. (2001), “Improving judgment in forecasting,” in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA: Kluwer Academic Publishers.

Hubbard, R. & D. E. Vetter (1996), “An empirical comparison of published replication research in accounting, economics, finance, management, and marketing” *Journal of Business Research*, 35, 153-164.

Joutz, F. L. & H. O. Stekler (1998), “Data revisions and forecasting,” *Applied Economics*, 30, 1011-1016.

Larwood, L. & W. Whittaker (1977), “Managerial myopia: Self-serving biases in organizational planning," *Journal of Applied Psychology*, 62, 194-198.

Lau, R. R. (1994), “An analysis of the accuracy of trial heat polls during the 1992 presidential election,” *Public Opinion Quarterly*, 58, 2-20.

Luthans, F., R. M. Hodgetts & S. A. Rosenkrantz (1988), *Real Managers*. Cambridge, MA: Ballinger.

MacGregor, D. G. (2001), “Decomposition for judgmental forecasting and estimation,” in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA: Kluwer Academic Publishers.

Makridakis, S. (1996), “Forecasting: Its role and value for planning and strategy,” *International Journal of Forecasting*, 12, 513-537.

Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen & R. Winkler (1982), “The accuracy of extrapolation (time series) methods: Results of a forecasting competition,” *Journal of Forecasting*, 1, 111-153.

Makridakis, S. & R. Winkler (1989), "Sampling distributions of post-sample forecasting errors," *Applied Statistics*, 38, 331-342.

McCloskey, D. N. & S. Ziliak (1996), "The standard error of regressions," *Journal of Economic Literature,* 34, 97-114.

McNees, S. K. (1990), "The role of judgment in macroeconomic forecasting accuracy," *International Journal of Forecasting*, 6, 287-299.

Meade, N. & T. Islam (2001), "Forecasting the diffusion of innovations: Implications for time-series extrapolation," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Morwitz, V. G. (2001), "Methods for forecasting from intentions data," in J. S. Armstrong (ed.), *Principles of Forecasting*. Norwell, MA**:** Kluwer Academic Publishers.

Mosteller, F. & J. W. Tukey (1977), *Data Analysis and Regression*. Reading. MA**:** Addison-Wesley.

Ohlin, L .E. & O. D. Duncan (1949), "The efficiency of prediction in criminology," *American Journal of Sociology*, 54, 441-452.

Rosenthal, R. & K. L. Fode (1963), "The effect of experimenter bias on the performance of the albino rat," *Behavioral Science*, 8, 183-189.

Rowe, G. & G. Wright (2001), "Expert opinions in forecasting**:** The role of the Delphi technique," in J. S. Armstrong (ed.)., *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Rush, H. & W. Page (1979), "Long-term metals forecasting**:** The track record**:** 1910-1964," *Futures*, 1, 321-337.

Sanders, N. & L. P. Ritzman (2001), "Judgmental adjustments of statistical forecasts," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Simon, J. (1985), "Forecasting the long-term trend of raw material availability," *International Journal of Forecasting,* 1, 85-109.

Smith, S. K. & T. Sincich (1991), "An empirical analysis of the effect of length of forecast horizon on population forecat errors," *Demography*, 28, 261-274.

Squire, P. (1988), "Why the 1936 *Literary Digest* poll failed?" *Public Opinion Quarterly*, 52, 125-133.

Stewart, T. R. (2001), "Improving reliability of judgmental forecasts," in J. S. Armstrong (ed.), *Principles of Forecasting.* Norwell, MA**:** Kluwer Academic Publishers.

Tashman, L. J., T. Bakken & J. Buzas (2000), "Effects of regressor forecast error on the variance of regression forecasts," *Journal of Forecasting*, (forthcoming).

Tiegen, K. H. (1983), "Studies in subjective probability III**:** The unimportance of alternatives," *Scandinavian Journal of Psychology*, 24, 97-105.

Tull, D. S. (1967), "The relationship of actual and predicted sales and profits in new-product introductions," *Journal of Business*, 40, 233-250.

Wagenaar, W. A., R. Schreuder & A. H. C. Van den Heijden (1985), "Do TV pictures help people to remember the weather forecast?" *Ergonomics*, 28, 756-772.

Webby, R., M. O'Connor & M. J. Lawrence (2001), "Judgmental time series forecasting using domain knowledge," in J. S. Armstrong (ed.), *Principles of Forecasting*. Norwell, MA**:** Kluwer Academic Publishers.

Weimann, G. (1990), "The obsession to forecast**:** Pre-election polls in the Israeli press," *Public Opinion Quarterly,* 54, 396-408.

Whitecotton, S. M., D. E. Sanders & K. B. Norris (1998), "Improving predictive accuracy with a combination of human intuition and mechanical decision aids," *Organizational Behavior and Human Decision Processes*, 76, 325-348.

Wilkinson, L. & the Task Force on Statistical Inference (1999), "Statistical method in psychology journals," *American Psychologist*, 54, 594-604.

Willemain, T. R., C. N. Smart, J. H. Shockor & P. A. DeSautels (1994), "Forecasting intermittent demand in manufacturing: A comparative evaluation of Croston's method," *International Journal of Forecasting*, 10, 529-538.

Wittink, D. & T. Bergestuen (2001), "Forecasting with conjoint analysis," in J. S. Armstrong (ed.), *Principles of Forecasting*. Norwell, MA**:** Kluwer Academic Publishers.

Yokum, T. & J. S. Armstrong (1995), "Beyond accuracy**:** Comparison of criteria used to select forecasting methods," *International Journal of Forecasting*, 11, 591-597. Full text at hops.wharton.upenn.edu/forecast.

**Acknowledgments:** Monica Adya, Dennis A. Ahlburg, Hal R. Arkes, Christopher Chatfield, Fred Collopy, James E. Cox Jr., Stephen A. DeLurgio, Robert Fildes, John Guerard, Wilpen L. Gorr, Nigel Harvey, James H. Hoover, Hans Levenbach, David G. Loomis, Stephen K. McNees, Marcus O'Connor, J. Keith Ord, Roy L. Pearson, Leonard J. Tashman, Dick R. Wittink and others provided suggestions on the principles, conditions, and evidence. Useful suggestions were obtained when earlier versions of this paper were presented at the International Association of Business Forecasters meeting in Philadelphia (May 1997), the International Symposium on Forecasting in Barbados (June 1997), the Judgmental Inputs to Forecasting Conference at University College, London (November 1998), the Federal Forecasters Conference in Washington, D.C. (June 1999), and the World Future Society Conference (July 1999). Raphael Austin, Zack Michaelson, Ling Qiu, and Mariam Rafi made editorial revisions.