USER PERSPECTIVE IN JUDGMENTAL ADJUSTMENTS: NESTED ADJUSTMENTS AND EXPLANATIONS

A Ph.D. Dissertation

by

M. SİNAN GÖNÜL

Department of Management

Bilkent University

Ankara

July 2007

To my dearest family;
My father; Zeki,
My mother; Nurşin
and My brother Ayhan

USER PERSPECTIVE IN JUDGMENTAL ADJUSTMENTS: NESTED ADJUSTMENTS AND EXPLANATIONS

The Institute of Economics and Social Sciences

of

Bilkent University

by

M. SİNAN GÖNÜL

In Partial Fulfilment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in

THE DEPARTMENT OF MANAGEMENT

BİLKENT UNIVERSITY

ANKARA

July 2007

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Management.

Professor Dilek Önkal
Supervisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Management.

Associate Professor Ümit Özlale
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Management.

Assistant Professor Zahide Karakitapoğlu-Aygün
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Management.

Assistant Professor Ayşe Kocabıyıkoğlu
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Management.

Professor Serdar Sayan
Examining Committee Member

Approval of the Institute of Economics and Social Sciences

Professor Erdal Erel
Director

ABSTRACT

USER PERSPECTIVE IN JUDGMENTAL ADJUSTMENTS:
NESTED ADJUSTMENTS AND EXPLANATIONS

Gönül, M. Sinan

Ph.D., Department of Management

Supervisor: Prof. Dilek Önkal

July 2007

The purpose of this thesis is to investigate the judgmental adjustment behavior of forecast users on externally provided predictions. "Nested judgmental adjustments" are defined as a series of revisions on a set of given forecasts. These adjustments are commonly used in practice to integrate judgment into forecasting processes. Explanations accompanying predictions may also influence forecast acceptance and adjustment in organizations. To study nested judgmental adjustments, explanations and user perspective, this research reports the results of a survey and three experiments. The survey is conducted with forecasting practitioners to enhance our understanding of the reasons and motivations behind judgmental adjustments, as well as to examine expectations of forecast users and perceptions of forecast quality. In addition, experimental studies are carried out to investigate the effects of structural characteristics of explanations and the

presence of original forecasts on adjustment behavior. Results are discussed and future research directions are given.

Keywords: Judgmental Forecasting, Judgmental Adjustments, Forecast Explanations

ÖZET

YARGISAL DÜZELTMELERDE
TAHMİNLERİ KULLANANLARIN BAKIŞ AÇISI:
İÇİÇE DÜZELTMELER VE AÇIKLAMALAR

Gönül, M. Sinan
Doktora, İşletme Bölümü
Tez Yöneticisi: Prof. Dr. Dilek Önkal
Temmuz 2007

Bu çalışmada temel amacımız tahmin kullanıcılarının sunulan öngörülere uyguladıkları yargısal düzeltmeleri irdelemektir. "İçiçe yargısal düzeltmeler" belirli bir tahmin üzerinde birbiri ardına uygulanan düzeltmeler olarak tanımlanabilir. Bu düzeltmeler, şirketlerin tahmin sürecinde yaygın olarak kullanılırlar ve kullanıcıların tahminlere kendi yargı ve düşüncelerini eklemelerine olanak sağlarlar. Tahminlerin kabul ve düzeltilmelerinde önemli etkisi olabilecek bir başka araç da tahminlerle ilgili açıklamalardır. Bütün bu kavramların etkilerinin incelenebilmesi için bu tezde bir anket ve üç deneysel çalışmadan oluşan bir araştırma anlatılmaktadır. Anket çalışması gerçek şirket çalışanlarına uygulanmıştır. Anketin temel amacı yargısal düzeltmelerin arkasındaki sebepleri ve motivasyonları irdelemek ve kullanıcıların beklentileri ile tahmin kalitesi arasındaki ilişkileri araştırmaktır. Ek olarak, açıklamaların yapısal özelliklerinin ve orijinal tahminlere ulaşımın düzeltme sürecine olan etkilerini araştırmamıza

olanak veren üç deneysel çalışma da yapılmıştır. Bu çalışmalardan çıkan sonuçlar tartışılmış ve gelecekteki araştırmalar için yeni fikirler sunulmuştur.

Anahtar Kelimeler: Yargısal Tahminler, Yargısal Düzeltmeler, Tahmin Açıklamaları

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest and most sincere gratitudes to my supervisor, Dr. Dilek Önkal for her continuous support and encouragement, for her generous time and devotion and most importantly for her excellent guidance that skilfully shaped me into the academic I'm today. Her insights, perspective and experience have been invaluable indeed. I would also like to extend my gratitudes to my core committee members, Dr. Ümit Özlale and Dr. Zahide Karakitapoğlu-Aygün for their assistance, constructive criticisms, precious advice and support during various stages of my PhD study.

I would also like to thank Dr. Michael Lawrence, Dr. Paul Goodwin, Dr. Mary Thomson and Dr. Aslıhan Altay-Salih for their valuable contributions to various parts of my dissertation. I'm certainly obliged to thank Dr. Andrew Pollock, Dr. Alex Macaulay, Dr. Serdar Sayan and Dr. Ayşe Kocabıyıkoğlu for their helpful insights and comments. I'm also grateful to Beth Doğan for her help in improving the language of my thesis.

It has been a pleasure and privilege to be a part of the Faculty of Business Administration, Bilkent University for eight years, from the first days of my graduate studies. I have greatly benefited from and been nurtured by the wonderful academic environment there. I would like to thank to every faculty

member, past and present graduate assistants, secretaries and members of staff for providing such an atmosphere.

I will always remember the year I have been a part of Glasgow Caledonian University. The perspectives and experience I have obtained there will have important repercussions for the rest of my life. I'm extremely grateful to the faculty members, PhD assistants, staff members and my friends there for making me feel at home away from home.

My special gratitudes to Dr. Muammer Ermiş in Electrical and Electronics Engineering Department in METU, for being the first to inform and encourage me to pursue a PhD, when I was just a second year undergraduate student.

I can never thank enough to my dear friends who were always there for me. They never hesitated to listen, share, comfort and provide support when I needed the most. My sincere thanks also go to those who will always hold individual places in my heart and memories.

Finally I am forever indebted to my dearest parents, Zeki/Nurşin and my dearest brother, Ayhan for their never ending love, caring, encouragement and unconditional support for whatever I have chosen to do. They have always been and will always be beacons of light for my life. Without them this thesis would not have been completed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

One of the most prominent struggles in decision making today is the one that is fought against uncertainty. Human beings have never been blessed with the ability to know for certain what the future will bring. Thus, a decision maker should be able to come up with clever plans and decisions to compensate for this absence of information. In other words, he has to effectively manage the uncertainty about the future. However, just as wars cannot be won without weapons, a decision maker should be equipped with the proper tools if he wants to prevail in this struggle against uncertainty. These important weapons or tools in the service of decision makers are known as forecasts.

Forecasts provide intelligent predictions about the future; thereby serving as valuable aids for the management of uncertainty. Decision makers routinely seek a variety of forecasts in many different fields in order to support their decision-making processes. Forecasts have a wide range of usage from predicting, for example, the stock prices in a week, or the product sales in a sector, to the following day's weather.

Forecasting as a discipline emerged by putting various statistical and mathematical methods into use to generate predictions about future events. As it has evolved, forecasting research has expanded its tool box and its areas of application.

In addition to mathematical and statistical methods, the practice of forecasting also employs approaches that depend heavily on human judgment. Forecasting is not and cannot be devoid of the human mind and cognition. In fact, every step of forecasting involves judgment and requires some sort of judgmental input. Judgment influences the forecasting process from the very beginning, starting with model formulation and variable selection. It then affects forecast generation: instead of using formal statistical methods, forecasts can also be generated based on human judgment alone. Other ways in which judgment permeates the forecasting process include adjusting statistical forecasts using judgment and combining judgmentally produced forecasts with statistically produced ones. All these different types of forecasts are collectively known as *judgmental forecasts*.

As the foregoing discussion suggests, judgmental forecasts serve as preliminary agents for the incorporation of human knowledge, intuition, experience and opinions into the process of prediction formation. Thus, they are widely used and appreciated in a variety of fields and settings in business and industry (Menzter and Cox, 1984; Sparkes and McHugh, 1984; Dalrymple, 1987;

Batchelor and Dua, 1990; Winklhofer et al., 1996; Sanders and Manrodt, 1994; 2003).

Regardless of the nature of forecasts (i.e., whether judgmental or statistical), every forecasting process involves two perspectives. The first perspective is the provider perspective. This perspective is related to the generation and supply of forecasts and it generally involves people who are formally educated and experienced in forecasting theory. In this perspective, the focus is on issues such as gathering and preparation of data, selection and implementation of forecasting methods and, eventually, generation and testing of forecasts.

The second perspective is from the point of view of the decision makers and managers who actually demand, obtain and utilize the forecasts. In struggling with their various responsibilities, decision makers or managers usually cannot spend the time and effort needed to generate forecasts. Instead, they simply acquire already generated forecasts and use them in their decision making process. Their focus is on the issues of acceptance and ease of utilization of those forecasts.

Clearly, there are important distinctions between the two perspectives. That is, what the users of forecasts think to be important in forecasts may be very different from what the providers of these forecasts believe or think is important. Even the most basic definitions, such as what a ‘good’ forecast might be, may differ among these perspectives. If this distinction is not properly handled, the forecasts generated by the suppliers may not satisfy the needs of the forecast users,

and this will lead to the eventual rejection of those forecasts. What benefit will a forecast bring if it is not employed by actual users? A forecast may be generated by a superior method and prove to be highly accurate, but if it is not acquired or used, all those qualities will not be worth a cent. Clearly, the crucial thing is not the generation of accurate forecasts, but the acceptance of those forecasts by decision makers and their utilization in a real business setting.

Although the separate natures of forecast providers and users have been recognized since the 1970s, the provider perspective has always been the favorite theme of the forecasting literature. The majority of research in this field is focused on the development and testing of new methods and tools and the creation of new criteria, all of which are important subjects for forecast providers. Unfortunately, however, the research on user perspective has never shared this popularity. Only recently have studies specifically addressing users begun to appear in the literature (e.g., Yates et al., 1996; Ackert et al., 1997; Önkal and Bolger, 2004; Price and Stone, 2004).

To compensate for this imbalance in the forecasting academia, the main perspective of this thesis will be the users of forecasts. It is of critical importance to determine what the users of forecasts expect and demand from the forecasts. It is also extremely important to learn what the users' criteria for successful forecasts are. Moreover, what they mean by the "quality of forecasts", and what constitutes that quality, should be explored. Only in this way can new tools and methods that will successfully meet, or even exceed, the needs of forecast users be developed.

Imagine a decision maker who has to make an important decision. As an aid, he has obtained a forecast from a professional forecasting company. He receives the forecast and inspects it. Now, what can he do? The decision maker either accepts the forecast the way it is, or applies an adjustment on it based upon his knowledge and experience. This judgmental adjustment will be directly proportional to the extent of his acceptance. If the forecast is broadly acceptable, the adjustments will be smaller. However, if the forecast is largely unacceptable, he either applies excessive adjustments on it or simply rejects it to the waste bin.

This anecdote is aimed toward explaining an important concept of user perspective, namely, judgmental adjustment of the provided forecasts. The user perspective pertains more to judgmental adjustments than to the judgmental generation of forecasts. Quite expectedly, the presence and popularity of judgmental adjustments among managers and decision makers have been acknowledged on many occasions (Mathews and Diamantopoulos, 1986, 1989, 1990; Diamantopoulos and Mathews, 1989; Sanders and Manrodt, 1994; 2003, Önkal and Gönül, 2005).

The concept of judgmental adjustment is essential in order to have a complete understanding of user perspective. Without proper investigation of the acceptance and adjustment process, this understanding cannot be achieved. There are important gaps and opportunities for exploration in the literature on this process. For example, the reasons and motivations of decision makers that lie behind their judgmental adjustment of the provided forecasts are largely unknown.

Similarly, the situations that lead forecast users to the complete acceptance of those forecasts are not known either.

Research on judgmental adjustments has largely focused on single adjustments. In a real organization, the assumption that there will be single adjustments on acquired forecasts is not realistic; an acquired forecast could undergo the adjustment process more than once in different departments or levels before finally being utilized. In this manner, adjustments of adjustments, or *nested adjustments* of forecasts, all become principal elements of this complex process. The exploration of these issues constitutes one primary focus of this thesis. Indeed, the properties and effects of nested judgmental adjustments of forecasts have never before been investigated or even mentioned by the forecasting academia; therefore, they will form one of the unique contributions of this thesis to the current body of knowledge.

Like every process, the forecast acceptance and adjustment process is not perfect and needs to be improved. There is evidence that adjustments that are conducted may not always produce beneficial results (e.g. Carbone et al., 1983; Sanders and Ritzman, 2001). On these occasions, the application of adjustments may actually lead to a worsening in the performance of the forecasts. It seems that the forecast acceptance and adjustment process could benefit from additional tools provided as decision aids. One preliminary and crucial decision aid is the accompanying explanations to the forecasts.

Explanations are very important vessels for communication. Owing to this fact, explanations could be provided along with forecasts to improve the adjustment and acceptance process of forecasts. These explanations could describe the data, procedure or the line of reasoning behind the forecasts. They might even provide information about the underlying theory. Indeed, it has been suggested that if the provided advice or forecast is accompanied by a relevant explanation, its acceptance is improved (Lawrence et al., 2001).

Aside from this study, it is impossible to locate any other study conducted specifically on this topic. Given its vast potential for improving the process, the provision of explanations on the adjustment and acceptance of forecasts constitutes a highly fertile research opportunity. The most evident characteristics and properties of explanations must be explored to determine their impact and influence on the process. This is the other primary goal of this thesis.

For the exploration of the aforementioned issues, a research design composed of two parts was carried out. The first part of the research was done in the field, and involved interviews with professionals accompanied by a survey study. The second part of the research was composed of three experiments for the controlled investigation of the relationships among the concepts. Financial forecasts were chosen as the setting owing to their popularity (Önkal-Atay, 1998).

The organization of this dissertation is as follows: in the next chapter, a literature review on judgmental forecasts is provided along with their implications for current research. In the first part of the literature review, information on the

most prominent forecasting formats is provided. The second part explains the factors affecting judgmental forecasts. The third part of the review is dedicated to the distinction between the provider and user perspectives and the relevant studies. Judgmental adjustment of forecasts is the theme of the fourth part, while the review on explanations constitutes the last part of the review section. Primary research questions are defined in the third chapter. The fourth chapter reports the methodology and findings of the survey. Following these, the design and results of the experimental studies are presented in the fifth chapter. The final chapter of the thesis is devoted to a general discussion, conclusions and directions for future research.

CHAPTER 2

LITERATURE REVIEW ON JUDGMENTAL FORECASTS

2.1. Commonly Used Formats in Judgmental Forecasts

Two commonly used formats that are used to express judgmental forecasts (in fact, all forecasts) are point forecasts and interval forecasts. These two formats have different properties and characteristics, respond to different manipulations and are suitable for different task structures. Each of them has its own advantages as well as disadvantages. In this respect, it seems important to start this review with an introduction to these formats in order to have a better understanding of judgmental forecasts and the underlying research.

2.1.1. Point Forecast Format

One of the most prevalent forecasting formats in the literature is the point forecast format (Önkal-Atay, 1998; Önkal-Atay et al., 2002). Point forecasts are solitary numbers about the forecasted event, like “the price of this stock will be 2.5 YTL next week” or “tomorrow’s temperature will be 15˚C”. The forecast format is easy

for providers to express and for users to understand; it is this property that makes it highly sought and used in forecasting practice and research. However, its main disadvantage is the fact that it provides no clues about the uncertainty of the forecasted variable, and thus, it may convey a false sense of certainty regarding this variable.

One famous example in the forecasting literature of a study conducted via point forecasts is the M-competition series (Makridakis et al., 1982, 1993; Makridakis and Hibon, 2000; Koning et al., 2005). The first Makridakis-competition, or M-competition, had the aim of comparing the post-sample accuracies of many statistical forecasting methods applied to real-life data (Makridakis et al., 1982). For that purpose, a total of 1,001 real-life macro, micro, industrial and demographic data were utilized. Expert forecasters applied various statistical forecasting methods on monthly, quarterly and yearly data for a variety of horizons. The major finding of the analysis was the fact that simple models were just as accurate as more complex and sophisticated ones. Moreover, there was no single best method, and the accuracies of all the methods depended on the length of the forecasting horizon and the accuracy measure used. Combinations of single methods were also found to perform better than individual models.

The major criticisms of the M-competition were that no contextual information was available to the forecasters about the forecasted series and that only statistical methods were compared. The judgmental approach to forecasting was not even mentioned. In real life, no forecasting practice is devoid of

contextual information and judgmental inputs. The forecasters were thus unable to apply the full extent of their skills.

Therefore, the M2-competition was organized as a follow-up to the M-competition. It was conducted in 1987 with the purpose of overcoming the criticisms directed towards the M-competition. In this new challenge, real-life, real-time data were used. The study compared post-sample accuracies of point forecasts generated both by statistical methods and by expert forecasters. The expert forecasters participating in the study generated monthly forecasts for the following year. As time advanced and the actual values of the forecasted variables became clear, the forecasters learnt the outcomes of their forecasts. Afterwards, new forecasts for the upcoming months were requested. This process continued until 1990.

This analysis showed very similar results to the M-competition in that simple methods generated point forecasts as accurate as the more sophisticated models (Makridakis et al., 1993). Moreover, the forecasts generated by experts, in general, did not improve in accuracy over mechanical methods. This result was observed even though the M2-competition allowed expert forecasters to access contextual information and also allowed continual updating and revising of the forecasts as new information became available.

The last competition in the series, the M3-competition, was very similar in structure to the first competition. However, this time, a total of 3,003 series were used and the statistical methods involved new approaches such as neural networks

(Makridakis and Hibon, 2000, Koning et al., 2005). The results obtained were in line with those obtained in the previous competitions. Again, no relationship was found between the sophistication of a method and its accuracy. Similar to the previous competitions, the forecasts generated by a combination of different methods resulted in improved performance when compared against individual forecasts.

Overall, the M-competition series were paramount in demonstrating that for point forecasts, unsophisticated methods may produce forecasts as accurate as sophisticated ones, although sophisticated methodologies have a much better fit to historical data. The fame of the M-competitions also comes from the fact that the vast reservoir of time-series data of the competition were made available for the use of researchers in the areas of both statistical and judgmental forecasting. Makridakis and Hibon (2000) mention that more than 600 researchers have utilized the data of these competitions.

2.1.2. Interval Forecast Format

The second widely used forecasting format is interval forecasts. In this format, a forecaster provides a probabilistic range of values for a forecasted event. For example: “there is 90% probability that the price of this stock will be between 2 and 3 YTL next week” or “we are 70% certain that tomorrow’s temperature will be between 12˚C and 18˚C”. Many researchers in the field argue that a forecast’s presentation should involve not only the estimate itself but also the uncertainty of

that estimate as well (e.g., Fischhoff, 1988). In this respect, interval forecasts are more advantageous than point forecasts: as the previous examples suggest, interval forecasts provide both estimates about the future event and the uncertainty of those estimates.

Quite expectedly, then, interval forecasts are perceived to be more useful than point forecasts (Önkal and Bolger, 2004). Moreover, in an accuracy study, Johnson (1982) demonstrated that subjects receiving confidence intervals in a Bayesian revision task performed much better than other subjects receiving different formats when there was high uncertainty in the environment.

Although judgmental interval forecasts seem to have advantages over point forecasts, they also have some weak spots. The most prominent disadvantage is known as the overconfidence effect. Judgmental forecasters are generally found to be overconfident in their estimations, so that the judgmental interval forecasts they provide are excessively narrow (O'Connor and Lawrence, 1989; Arkes, 2001; Chatfield, 2001). These excessively narrow interval forecasts lead to the persistent finding that the realized values fall into predicted intervals less often than they should. For example, if a forecaster had predicted a 90% interval, the realized values would be observed in that interval much less than 90% of the time. That is to say, judgmental prediction intervals are not well-calibrated (i.e., accurate).

This effect is not pertinent only to judgmentally generated interval forecasts. The overconfidence effect has also been shown in statistically generated

interval forecasts when compared against actual data. Makridakis et al. (1987) generated statistical interval forecasts on the time-series data of the M-competition. The post-sample accuracy results showed that the number of actual values falling outside the statistical interval forecasts was much higher than theoretically anticipated. Likewise, in the O'Connor and Lawrence (1989) study, judgmentally generated 50% and 75% interval forecasts were compared against statistically generated interval forecasts; although the widths of the judgmentally generated forecasts were narrower than the widths of the statistical ones, the calibrations of both kinds of intervals were similar, owing to the fact that the statistically generated forecasts were also much narrower than they should be.

Another important characteristic of judgmental interval forecasts is that they are generally asymmetric with respect to point forecasts. Statistical interval forecasts are conventionally calculated to fall in a symmetric range around the point forecast in order to promote the idea that the event has an equal chance of falling above or below the point value. However, in judgmental forecasts this symmetry is not preserved even though the forecasters are trained in probability and statistics. The most probable reason for this is that the subjects do not believe the forecasted event has an equal chance of falling above or below the point value. Their experience, knowledge and expectations allow them to create asymmetrical judgmental intervals.

The research on the asymmetry of judgmental intervals is in its infancy and holds great promise for the future. One of the preliminary studies conducted on this issue was carried out by O'Connor et al. (2001). The researchers

investigated the causes and the extent of asymmetry in judgmental confidence intervals. They stated that asymmetry is most likely introduced due to either non-time-series information (i.e., contextual information) about the forecasted event or the trend and noise characteristics in series. The researchers found that forecasters seem to give up some accuracy in order to include additional information, leading to asymmetry. They also found that trend influences asymmetry in judgmental intervals, and that there is a relation between the most recent actual data, point forecasts and asymmetrical interval forecasts, in that forecasters have a tendency to bias their point forecasts in one direction, while prejudicing their judgmental intervals in the opposite direction to balance their judgment.

The issue of giving up some accuracy in order to include additional information in judgmental intervals was suggested previously by Yaniv and Foster (1995). This phenomenon was named the "accuracy-informativeness trade-off". This trade-off states that in order to increase the accuracy of interval forecasts, the width of the forecasts has to be increased. However, at the same time, wide intervals will be less informative for users. For example, it may be stated that "the price of this stock, which is 5YTL today, will be between 0YTL and 1,000YTL next week". In this case, perfect accuracy will be achieved, but the interval forecast that is provided will not be informative at all. Therefore, a trade-off must be made between being accurate and being informative.

Yaniv and Foster (1995), in their study, investigated this trade-off in a subjective probability-assessment setting and tried to find an empirical model for this trade-off. In a series of experiments, they concluded that the trade-off could

actually be observed in generating judgmental intervals, and that an additive model was appropriate in describing the relationship between informativeness and accuracy.

In a later study, Yaniv and Foster (1997) manipulated the elicitation methods the subjects used in providing interval judgments on general knowledge questions. The elicitation methods were grain scales, 95% confidence intervals and point values coupled with plus and minus bounds. The grain scales were graphical units with differing intervals so that the subjects drew their interval judgments on these units. The plus and minus bounds on the point value provided a range assessment just like a confidence interval. In the end, it was found that the accuracy-informativeness trade-off was present, regardless of the elicitation method used. Moreover, there were not different trade-offs for different elicitation methods. Yaniv and Foster also stated that there appears to be a direct relation between interval width and error. It is as if the interval width indicates the magnitude of the error.

In addition to the point and interval forecasts, probabilistic forecasts are another widely used forecasting format. Probabilistic forecasts are generally composed of two components: an estimate about an event and a probability assessment regarding the occurrence of that event. Some examples of this forecasting format are “there is 90% probability that the price of this stock will increase by more than 5% next week” or “there is 85% probability that tomorrow’s temperature will decline”.

Similar to interval forecasts, probabilistic forecasts are advantageous in the sense that they not only provide an estimate about the forecasted event but also a sense of uncertainty of the estimate in the form of a probability assessment. Moreover, the communication of this type of forecast to the users is simple and easy to understand. As can be expected, probabilistic predictions are widely used in financial and economic forecasting (Önkal-Atay, 1998) and numerous studies have been based on this forecasting format (Lichtenstein et al., 1982; Keren, 1991, 1997; Yates et al., 1991; Ayton and Wright, 1994; McClelland and Bolger, 1994; Önkal and Muradoğlu, 1994, 1996; Bolger and Wright, 1994; Whitecotton, 1996; Ayton and McClelland, 1997; Wilkie-Thomson et al., 1997; Önkal et al., 2003; Andersson et al., 2005).

Up to this point, we have tried to identify the differences and similarities among different forecasting formats used in judgmental forecasting. It is now time to explain some of the main factors affecting judgmental forecasts.

2.2. Factors Affecting Judgmental Forecasts

It seems clear, then, that the accuracy of judgmental forecasts is conditional on many factors. Contemporary research on judgmental forecasting has identified four crucially important factors (see Goodwin and Wright, 1993, 1994; Webby and O'Connor, 1996; Önkal-Atay, 1998; Önkal-Atay et al., 2002; Lawrence et al., 2006, for comprehensive reviews on judgmental forecasting). These factors are the presence or absence of contextual information, the time-series characteristics

of the forecasted event, the heuristics and biases of the human mind, and lastly, the presentation format of the data.

2.2.1. Contextual Information

The first factor influencing judgmental forecasts is the availability of contextual information along with time-series data. This type of information generally consists of non-time-series information in the form of news, pieces of knowledge, or rumors regarding the forecasted event. Contextual information can be related to the past behavior of the time-series data or it can be about the future of the time-series. It can also consist of information about an event whose effects were not incorporated into the forecast at the time of generation. Whatever and however it is, contextual information provides a much richer understanding about the forecasted event than simple time-series data. In this sense, the decision-making process and the judgmental-forecast generation process may greatly benefit from available contextual information. Moreover, by using contextual information, forecasters can incorporate the effects of some special cases, sporadic events and external influences into their judgmental forecasts. Önkal-Atay et al. (2002) suggest that the presence of contextual information can improve the accuracy of judgmental forecasts if it is incorporated appropriately.

Edmundson et al. (1988) have successfully demonstrated that judgmental forecasts constructed with contextual information perform better in terms of accuracy against forecasts generated without any contextual information. Sanders

and Ritzman (1992) achieved similar results. They found that contextually aided forecasts were superior both to forecasts generated with technical knowledge (i.e., knowledge of data-analysis and forecasting procedures) and to forecasts generated by statistical methods. In both of these studies, the correctness of the contextual information was taken for granted. Remus et al. (1995 and 1998) have relaxed this assumption by investigating the effects of the reliability and the correctness of contextual information.

In a judgmental point-forecast generation task, Remus et al. (1995) have utilized artificial time series containing structural breaks. In this setting, the contextual information was related to the presence of this break. One group received perfect contextual information (i.e., that the series would undergo a structural change) just before the introduction of this change, while the second group received imperfect information (i.e., that there was a 50% chance of a structural change). The last group served as the control group, and therefore received no contextual information. The analysis of the results indicated that as the reliability of the information increased, the accuracy of the generated forecasts also increased; namely, the group with perfect information performed better than the group with imperfect information, and both of these groups performed better than the control group. On the other hand, Remus et al. noted that this contextual information was not properly used since the forecasts generated by both of the contextual-information groups were less accurate than those generated by statistical methods based on exponential smoothing.

In the later study of Remus et al. (1998), the correctness of contextual information was manipulated. As in the previous study, the time-series used in the experiment was artificially constructed to incorporate a structural change. Correct contextual information would hint to the subjects about the correct direction of change in the trend, while incorrect information would talk about a downward change in the slope if there would actually be an upward change, or vice versa. The results showed that the correctness of the contextual information had a profound influence on the accuracy of judgmental forecasts, in that forecasts generated using correct information were more accurate than either forecasts generated using incorrect information or those generated using no information. Furthermore, participants receiving incorrect information at any point in the experiment generated more erroneous forecasts near the data points where the structural change was introduced. From both of these studies it is logical to conclude that if contextual information is correct and reliable, judgmental forecasters will benefit from it, and the generated forecasts will attain higher accuracies.

All the above studies provide evidence for the positive effects of contextual information. However, it is also possible to locate some studies where the presence of contextual information had negative effects on the accuracy of judgmental forecasts. In one of these studies, Davis et al. (1994) investigated the effects of redundant contextual information in a stock-earnings forecasting task. The subjects were presented with either baseline information about the firm or received additional contextual information in addition to baseline information.

The additional information consisted of two types: one group received non-redundant information that supported the baseline information, while the other group received redundant information that could have easily been derived from the baseline information. The results of the study state that in terms of accuracy, the group receiving only baseline information outperformed both the redundant and non-redundant information groups. However, in terms of confidence, the non-redundant information group felt more confident compared to the redundant information group, and the redundant information group felt more confident than the baseline group. Therefore, contextual information provided in addition to baseline information seems to have increased the forecasters' confidence, but decreased their accuracy.

Davis et al. mentioned that the most likely explanation for these findings is that additional information to baseline information, whatever its nature, was received as an overload by the subjects (1994: 236). Due to this information-overload effect, the human cognitive system was unable to register all the information, which led to a degradation in performance.

More evidence of the improper utilization of information comes through the work of Lim and O'Connor (1996b). This study tested the performance of forecasters in a forecasting task involving an interactive decision support system (DSS). The DSS was designed so that the forecasters were able to access a multitude of information upon request. The results of the experiment showed that the forecasters acquired information too inefficiently to make any positive contribution to their performance. They even showed a tendency to select less-

reliable information. It seems that people have some problems discerning reliable information and aggregating information coming from a variety of sources in an efficient manner.

As a bottom line for this section, it can be argued that contextual information together with time-series data can either improve or diminish the accuracy of judgmental forecasts. To obtain positive effects, a prerequisite is that contextual information should be correct and reliable. Given this condition, improvement seems to occur only if the information received does not cause an overload in the human cognitive system and if forecasters can incorporate that piece of information appropriately into their forecasts. This requires forecasters to have an appropriate level of domain knowledge (Webby et al., 2001; Lawrence et al., 2006). Having domain knowledge means that through training and experience, a forecaster becomes able to distinguish beneficial pieces of information from redundant or detrimental ones. At the same time, he will develop an understanding of what and how large the effects of the selected pieces of information would be on the forecasts he is producing.

2.2.2. Time-Series Characteristics

The second factor affecting judgmental forecasts is the characteristics of the underlying time-series. The major characteristics of time-series data are trend, seasonality and randomness/noise. Among these characteristics, trend was reported to be the one most frequently researched (Önkal-Atay et al., 2002).

The research on the effects of trend reports a frequent finding: judgmental forecasts of trended series generally produced dampened estimates for these trends when compared against statistically generated ones. That is, upward-sloping series are generally underforecasted (i.e., the judgmental forecasts are generally less than they should be) and downward-sloping series are generally overforecasted (i.e., the judgmental forecasts are generally more than what they should be) (Eggleton, 1982; Lawrence and Makridakis, 1989; Sanders, 1992).

The study conducted by Lawrence and Makridakis (1989) was important in determining the effects of trend on the accuracy of judgmentally generated point and interval forecasts. The researchers clearly demonstrated that trend influenced both point and interval forecasts, and that there was a dampening of the trend, so that underforecasting was observed for series with an upward trend and overforecasting was reported for series with a downward trend. Another important result of the study was that the tendency to dampen the downward series was greater than the tendency to dampen the upward series.

O'Connor et al. (1997) tried to further explore these results. They conducted an experiment with the purpose of investigating the effects of direction of trend on the accuracy of judgmental point forecasts. The results obtained confirmed Lawrence and Makridakis' study in that people produced different forecasts for series with different trend directions. The participants in the experiment generated worse forecasts for downward series when compared against flat and upward series. O'Connor et al. attributed this result to the

anticipation people generally have that downward series are more likely to reverse themselves than upward series, and because of this anticipation asymmetries with respect to trend occur in forecast generation.

However, this dampening effect observed in judgmental forecasts may not be bad news. Webby and O'Connor (1996) argue that for real-life series, dampening of the trend may produce better forecasts than statistical forecasts which maintain the historical trend at the same level. For extrapolation, these statistical forecasts assume that the historical trend will continue in the same manner into the future, while in real life it is very difficult, if not impossible, to make that assumption.

Besides direction, another aspect of trend is strength. In a recent study, Thomson et al. (2003) investigated the effects of trend strength on the generation of probabilistic judgmental forecasts in currency-exchange rate forecasting. The trend strength was found to have a substantial influence on every aspect of judgmental forecasting performance. The authors stated that the stronger the trend became, the easier it was for forecasters to notice and forecast that trend, since it would become more visible and evident. Moreover, the hard-easy effect frequently encountered in subjective probability research was observed (Lichtenstein et al., 1982; Keren, 1991, 1997; Ayton and Wright, 1994; McClelland and Bolger, 1994; Ayton and McClelland, 1997). Participants showed overconfidence in forecasting difficult (or weak) trends and underconfidence in forecasting easy (or strong) trends. Contrary to the Lawrence and Makridakis (1989) and O'Connor et al. (1997) studies, Thomson et al. report that in general,

negative trends performed better than positive trends. These contrary results were attributed to either differing forecasting formats or differing natures of the forecasting tasks.

Other important time-series characteristics are seasonality and randomness. These characteristics have been asserted to have a significant influence on the performance of judgmental forecasts (Önkal-Atay et al., 2002; Webby and O'Connor, 1996). The research on these characteristics has found that the accuracy of judgmental forecasts declines with high seasonality (Adam and Ebert, 1976; Sanders, 1992) and high randomness (Adam and Ebert, 1976; Eggleton, 1982; Sanders, 1992; O'Connor et al., 1993). However, these observations are also valid for statistical forecasts. The performance of statistical forecasting methods also decreases with increasing seasonality and randomness in data (Webby and O'Connor, 1996).

In a related study, O'Connor and Lawrence (1992) investigated the effects of seasonality, trend and noise on the generation of judgmental interval forecasts. They developed metrics to quantify the seasonality, trend and noise so that they could regress the width of the judgmental intervals against time-series characteristics. From these analyses they concluded that the most important characteristics affecting the width of judgmental intervals were seasonality and trend. While the noise present in a series is a fundamental determinant of interval width for statistical methods, it was not as important for judgmental intervals. O'Connor and Lawrence also stated that all three factors combined could only explain 57% of the variability in judgmental interval widths; therefore, there are

other important factors influencing judgmental interval forecasts besides time-series characteristics.

2.2.3. Mental Heuristics and Biases

The heuristics and biases of the human cognitive system comprise the third factor that affects judgmental forecasts. Heuristics are mental shortcuts, simplifying tools, or rules of thumb that human beings utilize in making judgments and decisions. Being limited in capacity and processing power, human minds rely on these shortcuts or simple rules due to their ability to reduce the complex and sophisticated processes of judgment and decision-making into simpler operations. Heuristics bring speed and grant the ability to cope with the massive amount of data continuously pouring into the human cognitive system from the environment. If such simplification mechanisms were not present, human minds would be overwhelmed by the amount of data-processing and decision-making that has to be done.

Even though heuristics are fundamental for survival in the environment, like every shortcut, they carry the potential of misconceptions, severe errors and biases. Since it is the human mind and judgment that produce judgmental forecasts, these forecasts' sensitivity to the heuristics and biases of human cognition is inevitable. Many heuristics and biases identified in the psychology literature have been found to be relevant to the forecasting process (Goodwin and Wright, 1994; Önkal-Atay et al., 2002).

One of the most prominent heuristics that influences judgmental forecasters is the anchoring and adjustment heuristic (Tversky and Kahneman, 1974, 1982). This heuristic states that in their pursuit to ease the process of decision-making, people often take a number or a piece of information as a base or initial value, and then make adjustments to that number to obtain their final answer. Moreover, in general, the amount of adjustment made to the anchor is not sufficient; it is less than what it should be. For a judgmental forecasting task, this heuristic works in a similar way. A forecaster often takes a value as an anchor and makes adjustments to it to obtain his/her judgmental forecast.

As to the nature of the anchor point, there has been some debate. Lawrence and O'Connor (1992) determined that for untrended series, judgmental forecasters show an anchoring-and-adjustment behavior by taking the long-term average of the series as the anchor point. For trended series, Harvey et al. (1994) suggest that the anchor point is the last data point in the series, and they argue that forecasters add a proportion of the last difference in the data (i.e., the difference between the most recent data points) to this anchor as the adjustment. Moreover, forecasters may also add some noise to the output of this process (Bolger and Harvey, 1993).

Like the original formulation by Tversky and Kahneman, in judgmental forecasting, the amount of adjustment made on the anchor is generally insufficient. In the studies done by Harvey et al. (1994) and Harvey and Bolger (1996) it is suggested that through this insufficient adjustment process, trend-dampening

behavior can be frequently observed (Eggleton, 1982; Lawrence and Makridakis, 1989; Sanders, 1992). However, these suggestions and findings are not conclusive since there have also been divergent findings about the amount of adjustment. For example, in a judgmental point-forecasting task on the series of the M-competition, subjects were reported to have generated higher than expected, even excessive, adjustments on the anchors (Lawrence and O'Connor, 1995).

Another crucial heuristic that finds application in the judgmental forecasting process is the representativeness heuristic, also developed by Tversky and Kahneman (1974, 1982). According to Tversky and Kahneman, the representativeness heuristic is in effect when people use a representative event to evaluate the likelihood of another event. In such a case, the event is judged to be likely or unlikely based on the degree to which it resembles another event. When individuals try to make a judgment about an event, they first access the stereotypes and prototypes in their minds and then make a decision about that event by the amount of resemblance between the two. Alternatively, they may project the event mentally in order to deduce the outcome by referring to the outcomes of similar events that have been encountered previously. This heuristic, like the others, can provide quick and effortless judgments, but it may not be sensitive to certain factors that must be taken into account when making judgments. These include prior probabilities, sample size and predictability of the event. Representativeness may also cause misconceptions about the probability of occurrence of events and may create illusions of validity (Tversky and Kahneman, 1974, 1982).

These arguments are also valid for judgmental forecasting tasks. Harvey (1995) and Harvey and Bolger (1996) assert that when generating a judgmental forecast, the forecasters may try to simulate the time-series in their minds by making use of the representativeness heuristic. In this process, they try to project not only the pattern, but also the amount of noise in the series. However, when they cannot do this very efficiently they may produce highly variable and inconsistent estimates about the forecasted event. In a relevant work, Lories et al. (1997) showed that when naïve subjects were asked to produce forecasts about time-series, they made use of a sort of representativeness heuristic.

The overconfidence effect observed in interval forecasts and probabilistic forecasts constitutes another primary bias. As was stated before, individuals are generally miscalibrated in their probability assessments so that they overestimate the chance of occurrence of certain events (Lichtenstein et al., 1982; Keren, 1991, 1997; Ayton and Wright, 1994; McClelland and Bolger, 1994; Ayton and McClelland, 1997). This effect finds repercussions in judgmental interval and probability forecasting as well, since both of these formats involve subjective probability assessments. With respect to this fact, judgmental forecasters were found to be overconfident in their estimations. In the case of judgmental interval forecasts, the intervals generated were observed to be excessively narrow (O'Connor and Lawrence, 1989; Arkes, 2001; Chatfield, 2001).

In a pertinent study, Bolger and Harvey (1995) asked subjects to forecast the probability that the next point in the series would be below or above a certain

reference value. In this task, participants were found to overestimate the probabilities that were less than 0.50 and underestimate the probabilities that were more than 0.50. These results were consistent for both 'below' and 'above' cases, but the amount of over/underestimation was indicated to be higher in the 'above' case.

Aside from being overconfident, people also have a poor understanding of the randomness in a series. They have a tendency to perceive systematic patterns in a completely random set of data. This is known as the "clustering illusion" (Gilovich, 1991). In their struggle to cope with uncertainty, individuals try to form clusters or patterns out of completely random sequences of events. The clustering illusion is undoubtedly valid for time-series forecasting. O'Connor et al. (1993) found evidence for the presence of the clustering illusion in a judgmental point-forecasting task. In trying to generate forecasts for series having structural changes, subjects were observed to react excessively to random fluctuations in the series, perceiving random movements in the series as if they were signaling structural changes, where there had actually been none. O'Connor et al. also added that this effect was reinforced especially for series with high variability.

Other biases that may be relevant to the judgmental forecasting process are hindsight, illusory correlations, recency, selective perception, attribution of success and failure, unrealistic optimism, underestimating uncertainty and inconsistencies in judgment (Goodwin and Wright, 1994; Önkal-Atay et al., 2002). In a nutshell, brief descriptions of these biases are as follows: The hindsight bias represents the tendency in people to judge events to have been more likely,

retrospectively, when they have learnt the outcome. When individuals look and think back about past events, they generally give higher probabilities to them than they would have given at that time in the past. Illusory correlation is the tendency to perceive false relationships among events. The recency bias is the overreaction of forecasters to the most recent data in the series. Selective perception involves the tendency of individuals to ignore or degrade information that is contrary to their expectations or anticipations. Attribution of success and failure occurs when forecasters tend to attribute their success to their skill or expertise, but asymmetrically, attribute their failure either to external influences that are impossible to control or else completely to chance.

In a financial setting, Önkal-Atay (1998) also reported the presence of "base-rate neglect" in the generation of probabilistic forecasts. This bias occurs when forecasters tend to ignore the base-rate of an event (i.e. the frequency of occurrence of an event as observed in the past) and provide likelihood assessments that conflict with the historical frequencies.

In conclusion, because it is a product of the human cognitive system, judgmental forecasting is inevitably affected by heuristics and biases. As long as human beings are involved in forecast generation, the impact of heuristics and biases will be observed in one way or another. This is one of the reasons that judgmental forecasts perform much better than mechanical forecasts under some circumstances and much worse under others.

2.2.4. Presentation Format

The fourth important factor that has an effect on judgmental forecasts is the presentation format of time-series data. It has been reported that the presentation format has a deep impact on the accuracy and perceptions of judgmental forecasters (Goodwin and Wright, 1993, 1994; Webby and O'Connor, 1996; Önkal-Atay et al., 2002). One of the fundamental research topics on this issue is the effect of presenting a time-series in a table vs. presenting it in a graphical format. The work of Wagenaar and Sagaria (1975), Lawrence et al. (1985), Angus-Leppan and Fatseas (1986), Dickson et al. (1986) and Harvey and Bolger (1996) can be cited as the relevant work on the comparison of graphical vs. tabular presentation of data.

Wagenaar and Sagaria (1975) made this comparison in an exponentially growing time-series prediction task, and in such a setting they showed the effectiveness of tabular presentation over graphical. On the other hand, in the studies done by Angus-Leppan and Fatseas (1986), Dickson et al. (1986) and Lawrence et al. (1985), participants given graphical displays of data were shown to perform better than those receiving tabular presentations of the same data.

To clarify the apparent contradiction of the previous research, Harvey and Bolger (1996) conducted an extensive study to explore the conditions in which graphical presentation was better than tabular presentation, and vice versa. In an experimental setting, subjects received half of the series graphically and the other half in tabular format, and generated forecasts. The results obtained indicated that

for *trended* data, graphical presentation induced the subjects to generate more accurate point forecasts than tabular presentation. However, this effect was reversed in the case of *untrended* series. Participants performed better when they received the untrended time-series in a tabular display than in a graphical display. These findings were consistent for both high- and low-noise series.

Harvey and Bolger (1996) argue that there are at least two plausible explanations for different presentation formats inducing differences in forecasting performance. The first reason involves the overforecasting or trend-dampening tendency of the subjects. Both of these tendencies are related to the anchoring and adjustment heuristic (Tversky and Kahneman, 1974, 1982). Presentation format may create a difference in the selection of the anchor point and the amount of adjustment made, which, in turn, leads to distinctions in the trend-dampening and/or overforecasting behavior of the subjects. If this is the case, it is natural to expect the distinct influence of graphical vs. tabular format between trended and untrended series.

The second explanation is the susceptibility of individuals to misconceiving the presence of noise in a series. Due to the representativeness heuristic (Tversky and Kahneman, 1974, 1982), besides the pattern of the series, forecasters also have a tendency to represent the variability of the series in their forecasts. If the presentation format causes distinct perceptions of the variability in a series, these distinct perceptions will be reflected in the generated forecasts as well, and dissimilar forecasts will be observed.

Whatever the reasons might be, Harvey and Bolger have successfully shown that even though graphical presentation was often found to be better than tabular display in the literature, its success depends on the circumstances. The presence of trend seems to have critical importance in the matter. Likewise, Önkal-Atay et al. (2002) assert that some external factors, such as forecast horizon and environmental complexity, also have an influence on the effectiveness of the presentation format of the data.

Another important issue in this area is the presentation scale of the graphical data. Lawrence and O'Connor (1992) manipulated the length of the vertical axis of the time-series graph on which subjects generated point forecasts. All the time-series used in this study were artificial and involved no trend. In such a setting, the researchers did not observe any significant effect of the presentation scale on the accuracy of judgmental point forecasts. Aside from involving no trended series, another important limitation of the study was the confounding of the variability of the data with the presentation scale. A time-series may be perceived to be fluctuating slightly on a graph that has a small vertical scale; on the other hand that same time-series might be perceived to fluctuate widely if the vertical scale of the graph is large enough. In this manner, the interaction of the effect of variability and the effect of the presentation scale becomes inescapable.

Having noticed this limitation, Lawrence and O'Connor (1993) conducted another study. In this study, the authors examined the calibration of judgmental intervals while presentation scale and variability were experimentally manipulated. Subjects were asked to generate judgmental interval forecasts using data that

differed in both noise level and in presentation scale. Lawrence and O'Connor concluded that the amount of variability had an influence on the calibration of the subjects. Regarding the effects of scale, it was observed that the presentation scale of data also had a significant impact on calibration scores. Even the horizontal lines and the symbols on the graphs were shown to exert some influence on the calibration of the subjects. Overall, participants tended to provide excessively wide intervals for small-scale and low-noise series, while they tended to provide excessively narrow intervals for large-scale and high-noise series.

Up to this point, I have mentioned the basic concepts, formats and factors in judgmental forecasting in order to provide the preliminaries on this field of research. The knowledge conveyed so far is crucial for the comprehension of the concepts and ideas given in the remaining parts of this review. Having said that, it is now the time to turn our attention to user perspective in forecasting.

2.3. User Perspective in Judgmental Forecasting

A significant issue in forecasting theory and practice is the recognition that the forecasting process usually involves two perspectives. The first perspective is the *provider* perspective and the second is the *user* perspective. The first perspective is related more to the generation and supply of forecasts, and it is generally carried out by people who are formally educated and experienced in forecasting theory. For this point of approach, the nature and application of forecasting models and techniques and the accuracy of the provided forecasts gain primary

importance. On the other hand, the user perspective presumes the point of view of the managers or decision makers who actually demand, receive and use the forecasts that are generated by the suppliers. However, it must be noted that this distinction in perspective does not necessarily involve separate persons. A decision maker may be a forecast provider in a certain situation and a forecast user in another. The important point is that he cannot assume both mantles at the same time. He may be framed either as a provider or as a user for a particular context and forecasting task.

As has been rightfully acknowledged by some researchers (Wheelwright and Clarke, 1976; Gross and Peterson, 1978; Fischhoff, 1994), there are some perception differences between these two perspectives. That is, what the suppliers of forecasts think to be important in forecasts may be very different from what the users of these forecasts believe or think. The forecasts the suppliers are generating may not address the needs of the forecast users, which may lead to a communication problem between the two sides. At the same time, this distinction indicates that the acceptance of a forecast may be more important than its accuracy. A generated forecast can be as accurate and justifiable as possible; however, it will mean nothing if it is not utilized by the users, or if it becomes subject to excessive modifications.

Between the two sides, the concept and criteria of a successful forecast may be different as well. What is perceived as a successful forecast and what constitutes that success may have distinct explanations. Even basic definitions,

such as quality of forecast or what a ‘good’ forecast might be, may differ among them.

Although the differences in perceptions of providers and users have been recognized since the 1970s, relatively few studies have appeared in the literature directly addressing the issue. This branch of research is very promising for discovery. The research gap in the area is urgently in need of descriptive and prescriptive work for the investigation and understanding of the distinctions between the two sides. This is critically important for the generation of new forecasting tools, models and decision support systems that will successfully meet the expectations of and gain acceptance by users. Only in this way can a real and practical value be thoroughly attained. Accordingly, this section of the review is devoted to providing a detailed examination of the research conducted on this essential topic.

For the investigation of provider and user perspectives in forecasting, Wheelwright and Clarke (1976) conducted a survey spanning 127 major US companies. They produced different questionnaires for providers and users of forecasts and obtained responses from both sides. The results of the survey indicated distinct perceptions regarding the adequacy of forecasts and the confidence placed in each other. First of all, the providers of forecasts saw their companies as more technologically advanced than those of the users. Second, when evaluated by users, the skills of the providers were perceived to be less than when they were evaluated by the providers themselves. Third, the users’ confidence in the ability of the providers to choose the best technique was less

than the confidence the providers self-reported to have. Next, the users perceived themselves to be more knowledgeable about forecasting techniques and their applications than the providers rated them to be. Lastly, the providers rated themselves much more favorably in understanding the needs and problems of the management than the users rated the providers.

In this regard, all these different perceptions signify a clear communication gap between users and providers, discernible even within a single company. Wheelwright and Clarke commented that these perceptual differences, and hence, the communication gap, are observed because the criteria used to evaluate the forecasts are dissimilar between the two groups. Their conclusion is noteworthy: "…users and preparers must be more informed and practical about one another's roles if the potential of forecasting is to be fully realized" (1976: 40, in italics).

Gross and Peterson (1978) also reached similar conclusions. In an article written on their consultancy experience, they assert that the source of problems between the users and providers of forecasts is not technical in nature, but is related to the perceptions and beliefs of the two sides. The failure to communicate effectively was identified as the preliminary problem. Other problems included problems of distrust, confidence and commitment. It is argued that forecast users perceive providers as highly technical people with little understanding of real-world problems, and providers perceive users as intuitive people with insufficient knowledge to understand the forecast generation and development process. In the end, these differing perceptions lead to problems of distrust and confidence,

resulting in the insufficient commitment of management to the process of forecasting.

After these studies the forecasting literature remained silent for more than a decade on the issue. Only at the International Symposium on Forecasting (ISF) in 1991 was the matter reintroduced in a larger perspective. In a study based on the roundtable discussions among the researchers, practitioners, preparers and users attending the symposium, Mahmoud et al. (1992) renewed the interest in this subject. In their study, Mahmoud et al. compared the differences in perceptions among these four groups of people. It must be noted that differences among the groups were somewhat subtle, so that in many parts of the paper, researchers were grouped with preparers, while practitioners were grouped with users. Mahmoud et al. suggest that the discussions held designated three important gaps between theory and practice in forecasting.

The first gap was named the "*understanding gap*" (1992: 253), which was a restatement of the idea that the forecast providers and academicians fail to communicate effectively with the forecast practitioners/users. The forecast practitioners/users do not fully comprehend the measures, techniques and interests of the forecast preparers/researchers, and the forecast preparers/researchers fail to meet the needs of the practitioners/users.

The second gap is the "*data sharing gap*" (1992: 254). This gap is focused on the idea that researchers and preparers cannot always utilize the available data due to reasons of accessibility, quality or confidentiality. A much needed piece of

data may not be available, may be available but not accessible, or may be both available and accessible but lacking in sufficient quality. In such cases, the generated forecasts will most probably fail to satisfy the expectations of the practitioners/users.

The last gap between forecasting theory and practice is the "*political gap*" (1992: 254). This gap is argued to exist when the aims and motives of the forecast practitioners/users and academicians/ preparers are different from each other. The forecasting process is not just a statistical or judgmental process, but an organizational process that is vulnerable to various pressures, biases and hidden agendas. Mahmoud et al. exemplify that salespeople generally prefer underforecasts since they will receive a higher bonus when they outperform their forecasted sales. Likewise, manufacturers prefer overforecasts in order to prevent backlogs or delays. Managers may overforecast the success of certain projects to receive higher budgets. For reasons similar to these, deliberate biases are introduced to forecasts in organizations. Mahmoud et al. suggest that this is an important reality of the forecasting practice, and all the groups must become aware of its presence. In line with these arguments, some recent studies have found evidence on the presence of organizational pressures on forecasting (Jones et al. 1997) and have investigated the deliberate biases introduced to sales forecasts (Ehrman and Shugan, 1995; Lawrence and O'Connor, 2005), production forecasts (Goodwin, 2004) and financial and economic forecasts (Pons-Novell, 2003).

As measures toward closing the three gaps mentioned, Mahmoud et al. (1992) proposed providing incentives to the academia for more practice-oriented research and suggested more collaborative work between researchers and practitioners. They also pointed out the need for exploration of the criteria, expectations and needs of the forecast users.

Yokum and Armstrong (1995) carried out a similar comparison on a slightly different topic. The topic of their study was the criteria used in selecting forecasting methods. Although method selection is more directly related to the generation of forecasts, and thus, more closely related to the provider perspective, in their comparison Yokum and Armstrong surveyed four groups of people including decision makers, who are the most common forecast users. The other groups surveyed were forecast researchers, educators and practitioners. The inclusion and comparison of the user perspective with the other groups was one of the unique points of this study, setting it apart from previous research conducted on method selection in forecasting (e.g.,Carbone and Armstrong, 1982; Mentzer and Cox, 1984; Mahmoud et al., 1988). In this respect, the primary aim of the researchers was the examination of the similarities and differences in the criteria used for the selection of forecasting methods among these four groups of people.

The study was carried out via a questionnaire that was sent to the members or affiliated experts of the International Institute of Forecasting (IIF) and the forecast-selection criteria involved were the following: accuracy, timeliness in providing forecasts, cost savings resulting from improved decisions, ease of interpretation, flexibility, ease in using available data, ease of use, ease of

implementation, incorporating judgmental input, reliability of confidence intervals, development cost, maintenance cost and theoretical relevance (1995: 593).

The results of the survey indicated that 'accuracy' was rated higher than all other criteria, regardless of the role of the forecasters, replicating the findings of previous research conducted on the topic. However, many other criteria were rated to be almost as important as accuracy, especially the criteria related with the implementation and application of forecasting methods, such as 'ease of use', 'ease of implementation', 'flexibility' and 'ease of interpretation'. Yokum and Armstrong also reported many differences among the criteria across the groups. Although accuracy received the highest rating in all groups, the forecast researchers provided significantly higher ratings for accuracy than all the other groups. This indicates that practitioners and users of forecasts do not take accuracy to be as crucial as academicians consider it to be. Another major observation was that the implementation-related criteria were deemed to be more essential by the forecast users than by all the other groups. This finding presents direct evidence of the distinction between providers and users of forecasts, even on a topic that is more closely related to the provider perspective.

Another relevant study was conducted by Fischhoff (1994). In this study, Fischhoff analyzed a particular problem of a weather channel. The weather channel had complaints that its storm warnings were not taken seriously by the users of its forecasts, which resulted in heavy losses. Extrapolating from this case, Fischhoff identified four reasons for miscommunication between the providers and users of forecasts. The first reason is the "*ambiguity*" of forecasts (1994: 388).

Forecast providers do not clearly reveal the nature and the probability of the forecasted events to the users. The presentation of these entities is generally done in ambiguous and hard-to-comprehend language full of technical terms, which prevents the users from understanding the conveyed message. The second reason is "*irrelevance*" (1994: 391), which states that what the forecasters provide is irrelevant to what the forecast users want and seek. "*Immodesty*" (1994: 393) constitutes the third reason. Forecasters generally do not properly relay the true value of uncertainty in their forecasts to the users. In many cases, it is even impossible to find any indication of either uncertainty or confidence in the forecasts. This causes misconceptions about the certainty of those forecasts. The last reason is stated to be "*impoverishment*" (1994: 397), which asserts that forecasts and their consequences are not provided in the larger context of society, the environment and daily life, but are provided in a technical context that is isolated from real life. In this manner, impoverishment leads to the forecasts' being difficult to interpret, thus, hindering acceptance by users.

All the aforementioned studies carry a similar characteristic in that they remain only in the descriptive side of research. They define and explain the problem of the distinct nature of provider and user perspectives, and speculate on the possible differences and their causes. They do not present any hypotheses or try to prove them by empirical means. Quite expectedly then, after these descriptive studies were published, more empirically-oriented research began to appear in the literature (e.g., Yates et al., 1996; Ackert et al., 1997; Price and Stone, 2004 and Önkal and Bolger, 2004).

The first of these studies (Yates et al., 1996) was conducted in a probabilistic-forecasting setting. Yates et al. explored whether forecast users evaluate the performance of forecasts differently than the standard theoretical measures used by providers. In a series of experiments involving weather and stock-price forecasts, Yates et al. found that the users' performance criteria were indeed different from the standard analytical performance measures used by providers. The forecast users did not employ standard measures at all. Instead, they utilized more intuitive criteria to evaluate forecast performance, and correspondingly, the performance of the providers.

The most important intuitive criteria were determined to be the number of correct forecasts, the presence or absence of extreme forecasts, and the magnitude of the deviation when the provided forecasts turned out to be wrong. The first subjective criterion states that users pay a lot of attention to the number of correct forecasts in assessing the performance of forecast providers. Moreover, the keenness of accuracy is not so important; what matters is whether those probabilistic forecasts are in the correct category or not. In other words, users do not pay attention to how accurate a forecast is as long as that forecast turns out to be in the correct category. This result seems intuitive, since forecast supplying companies with a history of success are more frequently chosen by users.

For the second subjective criterion, it can be said that forecast users are sensitive to the extremeness in the provided forecasts. However, there were two kinds of responses to this extremeness. One group of subjects saw extreme

forecasts as indicators of the confidence of the forecasters, which resulted in a positive effect on ratings. On the other hand, the other group of users believed that extreme forecasts are against the uncertain nature of forecasting, and by this logic, they gave degraded ratings.

The third criterion indicates that forecast users are receptive not only to the presence of error, but also to the magnitude of that error. Unlike the case of categorically correct forecasts, when a forecast turns out to be in the wrong category, users also seem to be attending to how large the error is in the probability given for that category.

In addition to these three criteria, Yates et al. also proposed that users expect forecasters to be able to give good explanations for their forecasts. Finally, users were observed to be susceptible to the presentation and/or display of the forecasts as well.

A follow-up to the Yates et al. study was conducted by Price and Stone (2004). In this study, the researchers investigated the effects of the forecast providers' overconfidence on the users of those forecasts. They utilized a similar methodology to Yates et al. and employed stock-price probabilistic forecasting as the medium for their experiments. The results obtained indicate that forecast users do prefer an extreme forecaster who is overconfident to a moderate forecaster who is practically better-calibrated than the overconfident one. Moreover, the user perceptions of extreme forecasters were also found to be favorable over the moderate forecaster. Forecast users not only prefer the extreme/overconfident

forecasters, but also perceive them to be more knowledgeable and correct than the moderate ones. Price and Stone attribute these results to a mental heuristic where individuals assume that confident forecasters have their reasons for being confident; namely, they are more knowledgeable and more accurate than other forecasters, and they reflect this through their confidence.

On the other hand, what happens if this overconfidence turns into a systematic bias in a provided forecast? Will forecast users notice the presence of this bias and still persist in acquiring these forecasts? The answers to these questions were the main theme of the research done by Ackert et al. (1997). They investigated the behavior of forecast users when the forecasts provided to them included systematic biases. The chosen setting was stock-price forecasting, and bias was introduced in the form of over-optimism. Ackert et al. asked the subjects to buy or sell stocks to make the highest profit, and in this task, subjects had the opportunity to buy a forecast report on the stock prices. There were three experimental groups in terms of the bias introduced to these forecast reports: the unbiased group, the low-biased group and the high-biased group. The results of the experiment pointed to a significant impact of the amount of bias on both the acquisition behavior of forecast users and how those forecasts were employed by the users in their decision making. Quite expectedly, the participants showed a tendency to continue acquiring unbiased or low-biased forecasts, while their willingness to acquire the high-biased forecasts was found to be much lower.

The participants also showed adaptive behavior to the presence of bias in the provided forecasts. Over the sessions, they learnt how to use the given forecast

efficiently regardless of the presence of bias. The performance of the unbiased forecast group was significantly better than the biased forecast group for the initial sessions. However, the differences between the performances of the unbiased and low-biased groups diminished over the later sessions. These findings evidently suggest that decision makers may continue buying forecasts from companies, even if they have slight biases, and still achieve good performance; however, the same decision makers may discontinue use of the forecasts if the amount of bias becomes intolerable.

As has been argued before, an important communication tool between providers and users is the forecasting format used. Through the forecasting format, producers are able to convey their predictions and judgments of uncertainty to the users. It is probable that there are differences in the perceptions of various forecasting formats between users and providers. For instance, forecast users may perceive one particular format to be more difficult to understand than another, while for producers, expressing the forecasts in the former format may be preferable. If this is the case, then these distinct preferences will also contribute to communication problems between the two sides.

The only available research regarding this issue was conducted by Önkal and Bolger (2004). In this study, researchers explored the differences between the perceived usefulness of different forecasting formats. The compared formats included point forecasts, directional probabilistic forecasts, 95% interval forecasts and 50% interval forecasts. In a stock-price forecasting experiment, participants

assumed the roles of both providers and users, and conveyed usefulness ratings in both roles.

The experimental task required the subjects to first take the mantle of providers and asked them to generate stock-price forecasts in all four formats and to convey usefulness ratings. They then assumed the role of forecast users. In this case, they received external forecasts in all four formats and used these forecasts to generate stock portfolios. Afterwards, they were asked about their perceptions of the usefulness of the formats again.

The analysis of the results indicate that in both provider and user roles, participants rated 95% interval forecasts to be the most useful forecasting format, with probabilistic forecasts, 50% interval forecasts and point forecasts following after. Point forecasts were deemed to be the least useful format. It must be noted that there seems to be no difference in the usefulness perceptions between provider and user roles. This lack of difference may be attributable to the fact that the role of provider and user were performed by the same subject consecutively, which may have caused the subject's preference of different formats to persevere, regardless of the role.

Even though the presence of all these studies has contributed seminal knowledge to the understanding of user perspective in forecasting, they have only been successful in touching the tip of the iceberg. There have been many areas left under-attended or even unattended. The forecasting literature seems to be in urgent need of research directed primarily at exploring the needs, desires and

criteria of forecast users in order to develop effective new techniques and methodologies that will achieve high levels of acceptance. Many fundamental questions need answering, such as what the users' expectations are of the forecasts they receive, how important accuracy is to users, whether accuracy is as important as researchers wish to believe it is, and most importantly, what the notion of forecast quality is in a forecast user's mind.

Regarding these fundamental questions, to our knowledge up to now no studies have been conducted that have directly investigated the expectations of forecast users from the financial forecasts they obtain. Yates et al., in their influential study (1996), have utilized the term 'good' forecasts. As was mentioned previously, researchers found that the properties users expect from a good forecast cannot be captured by the standard theoretical measures used by forecast providers. What forecast users anticipate or want to see in financial forecasts still remains largely unknown to the forecasting field. Moreover, the results of the Yates et al. study were limited to the probabilistic forecasting format. Whether they can be generalized to point and interval forecasts has yet to be examined.

Prior to this study, Murphy (1993) also tried to analyze the meaning of a 'good' prediction in the context of weather forecasting. He argued that the concept of goodness for weather forecasts implies multiple dimensions including consistency (the match between the forecast providers' predictions and their judgment), quality (the accuracy of the predictions) and value (the benefits the forecast users will obtain by using those forecasts). Additionally, he strongly

emphasized that the concept of goodness was quite different between the forecast providers and those who actually utilize them, providing further support to the arguments made in this section.

As for the importance of forecast accuracy to users, only a few studies have been located that were focused specifically on the issue. One of these studies was based on sales forecasts used in strategic resource planning (Wacker and Lummus, 2002). The researchers argued that a high number of forecast errors do not necessarily mean a competitive disadvantage, and that companies learn to cope with a certain amount of error in the forecasts they use. They mention that forecasts only need to be accurate 'enough' and this does not require them to attain the smallest amount of error. Lawrence and O'Connor (2005) also mention a similar concept. They state that some of the companies they are associated with talked about 'tolerable error' or 'acceptable error'. These terms signify that companies have a tolerance level up to which errors in forecasts were not very important. Up to that level, however high the number of forecast errors might be, the forecasts would be perceived as accurate. Based on this notion, Wacker and Lummus introduced some forecasting paradoxes observable in resource-allocation activities.

The first paradox is that "the most important strategic decisions a company makes are based on the least accurate information" (2002: 1021), meaning that the most important strategies in a company generally involve long time-horizons, and forecasts of long time-horizons are less accurate than short-term forecasts. The second paradox states that the "forecast information that is most useful for

resource planning is the least accurate" (2002: 1022). The more detailed the product grouping is, the more valuable it becomes for resource planning, but at the same time, the more detailed the product grouping is, the more difficult and less accurate its forecasts become. "The organizations that need the most accurate forecasts have the largest forecast error" (2002: 1023) is the last paradox, which asserts that the industries that are in constant need of forecasts, like pharmaceuticals, electrical appliances and durable goods, generally face less accurate forecasts due to fluctuations in demand.

Likewise, the study of Winklhofer and Diamantopoulos (2002) brought up the notion that managers' concept of forecast effectiveness or forecast performance is not limited to accuracy, but also includes forecast bias, timeliness of forecasts and cost of forecasts. To provide evidence for these arguments, the authors have developed an empirical model and estimated it with data obtained from export-sales forecasts in UK firms. The analysis of the model output stated that short-term accuracy and absence of overestimation bias were equally influential in managers' perceptions of forecast effectiveness, while timeliness, cost and long-term accuracy carried less importance than the other two.

The findings of Yokum and Armstrong (1995) were in line with these results. As was mentioned before, they pointed out that there are many criteria deemed to be almost as important as accuracy in the field of forecasting-method selection.

All these studies provide ample evidence that forecast accuracy is not the sole criteria for forecasting effectiveness. Moreover, there is a tolerable or acceptable amount of error in forecasts, so that a moderately accurate forecast may still be good enough for the proper functioning of the forecasting process in a company. One important point to be mentioned is that the results of the aforementioned studies, unfortunately, remain limited to the sales-forecasting and forecasting-method selection fields, and the applicability of these results to other forecasts and their users has yet to be discovered.

The last fundamental question is about the quality of forecasts. For this notion, the literature has generally witnessed the term "quality of forecast" used in an equivalent meaning to the accuracy of that forecast (e.g., Granger, 1996; Murphy, 1993; Murphy and Wilks, 1998; Aukutsionek and Belianin, 2001). According to Granger (1996), a forecast was of high quality if it achieved the lowest amount of error. According to Aukutsionek and Belianin (2001), in a probabilistic-forecasting setting, a forecast was of high quality if it was well-calibrated. Although Murphy (1993) and Murphy and Wilks (1998) also use quality as accuracy, they point out the multi-dimensional nature of quality and argue that the concept of quality cannot be captured by a sole dimension or error parameter. However, their suggestion for resolving the issue is normative, and involves adding more parameters to the equations used for measuring forecast accuracy. These studies indicate the presence of a significant opportunity in that there is no currently available research which attempts to discover perceptions of forecast quality and its dimensions in the eyes of forecast users.

Before concluding this section, there is one last point to make. Provider perspective in forecasts generally deals with forecast production. This involves collection of data, selection of an appropriate model (either statistical or judgmental) and the generation of forecasts to be utilized by decision makers and managers in the companies. On the other hand, the perspective of decision makers or forecast users generally does not deal with forecast generation. Instead, they receive already generated forecasts and either use them the way they are or apply judgmental adjustments to them. In this respect, judgmental adjustment of forecasts is more relevant to the user perspective than forecast generation. In the next section, a review of the current state of the art in the judgmental adjustment of forecasts will be presented.

2.4. Judgmental Adjustment of Forecasts

There are two components of the judgmental adjustment process. The first component is the provided forecasts, which function as the 'baseline' for the adjustments. These forecasts are generally generated quantitatively (i.e., statistically), though qualitatively-originated forecasts are also frequently encountered (e.g., Lim and O'Connor, 1995). The forecast providers usually produce these baseline forecasts and convey them to the forecast users. The second component of the judgmental adjustment process is the judgment applied on these baseline forecasts by the receivers or users of those forecasts. This

application of judgment may lead the forecast users to introduce some changes, or judgmental adjustments, to these forecasts.

Judgmentally adjusted forecasts serve as a vessel for the integration of human judgment, intuition and experience into methodologically generated forecasts. In this case, forecast users need not spend time and effort producing the forecasts they need; instead they leave the job entirely to experienced forecasters. They then take the generated forecasts, inspect them and add their opinions and judgment to them by making adjustments. Due to its nature, this method of judgmental forecasting is more prevalent among forecast users than judgmental generation of forecasts. Carbone et al. (1983), state that judgmentally adjusted forecasts are "inevitable" in organizations (p.559). Quite expectedly, then, the forecasting literature provides extensive evidence of the presence and popularity of judgmental adjustments in the forecasting practice and real business settings (Mathews and Diamantopoulos, 1986, 1989, 1990; Diamantopoulos and Mathews, 1989; Sanders and Manrodt, 1994, 2003; Önkal and Gönül, 2005).

An important question on this topic is what the cognitive nature of the judgmental adjustment process is. Judgmental adjustments may owe their existence to the well-known heuristic of anchor and adjustment (Tversky and Kahneman, 1974, 1982). The baseline forecasts may serve as the anchor point, and forecast users may apply adjustments on that anchor by using the heuristic. Another perspective regarding this process is that forecast users generate an independent judgmental forecast in their minds, and when they encounter the already-generated forecast they mentally combine these two in some manner,

producing a judgmentally adjusted forecast. Although there have been some attempts to determine the underlying process (e.g., Lim and O'Connor, 1995), many of the research findings offer mixed results, and the cognitive nature of the process is still largely unknown (Goodwin and Wright, 1994).

Aside from this issue, the judgmental adjustment literature seems to concentrate on two major fields (Webby and O'Connor, 1996). The first area is judgmental adjustments conducted in the absence of contextual influence. The second area of focus is the judgmental adjustment of forecasts with the help of contextual information.

2.4.1. Judgmental Adjustments without Contextual Information

When there are no contextual cues available to forecast users, they are left with only the information contained in the time-series data and the provided forecasts. In this case, the forecast users have to make do with only these two sources of information for their adjustments. Since there are only two sources, then, only two factors are found to affect adjustment behavior, one for each source of information. The first factor is the nature of the underlying time-series (Sanders, 1992), while the second factor is the accuracy and reliability of the provided forecasts (Carbone et al., 1983; Carbone and Gorr, 1985; Willemain 1989, 1991; Lim and O'Connor, 1995).

As argued before, the time-series characteristics of the underlying series is one of the major factors that have an influence on judgmental forecasts. Likewise,

judgmental adjustments on provided forecasts are affected by these characteristics. Sanders (1992) explored this issue in a point-forecast generation and adjustment task. The task first required subjects to generate judgmental forecasts on a set of time-series data. Afterwards, they were provided with statistical forecasts on the same series and were asked to judgmentally adjust those. The time-series were artificially constructed to control the series characteristics, namely trend, seasonality and noise. Sanders discovered that judgmental adjustments on statistical forecasts led to improved accuracies when the series had low noise and when there was a definite and identifiable pattern (like seasonality) in the series. Judgmental adjustment of statistical forecasts worsened in accuracy for high-noise series. In a later study, Lim and O'Connor (1995) also showed that the seasonality of the time-series data had an influence on the performance of judgmental adjustments.

On the issue of the accuracy and reliability of the provided forecast, one of the preliminary studies that varied baseline forecasts was conducted by Carbone et al. (1983). In this study, the participants generated point forecasts on series selected from the M-competition. Another task they had to complete was the judgmental adjustment of the statistical forecasts they had generated. The accuracy measures on the generated and adjusted forecasts showed that judgmental adjustments did not improve the accuracy of the statistical forecasts. For one particular method, adjustments led to no significant changes in accuracy, while for the others, judgmental adjustments degraded the accuracy of statistical

forecasts. Thus, the researchers found evidence that the type of forecast on which adjustments are made has an effect on the accuracy of the adjustments.

In their subsequent study (Carbone and Gorr, 1985), the researchers carried out a similar experiment with a larger number of subjects. In this case, they tried to emphasize the time-series patterns via enhanced graphics. However, this attempt produced no significant differences among the groups, and overall, they achieved results conflicting with their previous work; that is, the researchers found that judgmental adjustments applied to statistically and judgmentally generated forecasts led to improved accuracy.

Like Carbone and Gorr, the highly cited studies of Willemain (1989, 1990) also utilized graphical presentation of the time-series data and the provided forecasts. For these presentations, Willemain asked the subjects to conduct their adjustments graphically. The main motivation of the author was to investigate the conditions under which judgmental adjustments may lead to improvements in the accuracy of the provided forecasts.

In the earlier study (1989), the time-series were artificially generated by an ARIMA process, and the provided forecasts were generated by applying simple methods to those time-series. Since the nature of the underlying series was known, Willemain was able to generate optimal ARIMA forecasts that specifically fitted the data and outperformed all the simple forecasts. For the adjustment task, the participants received the time-series and the simple forecasts. They were not

aware of the presence of the optimal forecasts, which were used only for accuracy comparisons.

Willemain achieved highly influential results. When the initial forecasts have a substantial opportunity for improvement (i.e., they are more erroneous than the optimal forecasts) the judgmental adjustments conducted on them attain increased accuracy. However, if the statistical forecasts are highly accurate (with respect to optimal forecasts), the judgmental adjustments have no effect on the accuracy of the initial forecasts, and may even lead to slight degradations in performance. In this respect, the performance of the graphical adjustment of point forecasts is contingent upon the accuracy of the baseline forecasts. If there is room for improvement in the statistical forecasts, the judgmental adjustments will use this opportunity to increase the accuracy.

The later study of Willemain (1991) was very similar to the previous one except that it contained real time-series data. In this study, optimal forecasts were not possible, and automatic and naïve forecasts were used for comparisons. The results obtained, even though statistically insignificant, were in the same direction as the previous study, and thus had a reinforcing effect on the findings.

All of the above studies successfully provide evidence for the effects of baseline forecast performance on judgmental adjustments. However, they lack a clear experimental manipulation of the forecast accuracy and reliability variable. The study conducted by Lim and O'Connor (1995) remedied this deficiency. In a series of experiments, they specifically manipulated the reliability of the provided

forecasts. Another major purpose of the authors was the exploration of the heuristics and biases inherent to the adjustment process. For these aims, they utilized the following setting: the participants first generated judgmental point forecasts on artificial sales data. Afterwards, they were provided with statistical forecasts and were asked to produce judgmental adjustments on their preliminary forecasts. For the statistical forecasts, two levels of reliability were selected: either high-reliability predictions or low-reliability ones.

The aggregated results indicated that the reliability of the provided forecasts had a significant impact on the accuracy of judgmentally adjusted forecasts. The high-reliability group performed better than the lower-reliability groups. Even when the provided forecasts had low reliability, the groups receiving them achieved some improvements in accuracy over their initial forecasts. However, their final accuracies were less than those of the provided statistical forecasts.

Moreover, as a cognitive bias, the subjects showed a tendency to place more weight on their initial forecasts, even if the provided statistical forecasts were highly reliable, and this behavior persisted over time.

2.4.2. Judgmental Adjustments with Contextual Information

The literature on judgmental adjustments with the presence of contextual information generally includes either experimental studies that investigate the effects of contextual cues (Lim and O'Connor, 1996a; Goodwin and Fildes, 1999;

Goodwin, 2000) or studies done on real organizational settings, which cannot be devoid of contextual influence (Mathews and Diamantopoulos, 1986, 1989, 1990; Diamantopoulos and Mathews, 1989; McNees, 1990).

To explore the effect of contextual information on the adjustment of point forecasts, Lim and O'Connor (1996a) utilized a similar setup to that of their previous study (Lim and O'Connor, 1995). As in the previous study, subjects first generated judgmental point forecasts. Afterwards, they received statistical forecasts and were asked to judgmentally adjust their initial forecasts. The forecasted event was the sales of a soft drink on a beach. The researchers artificially generated the sales data so that they were highly correlated with temperature. Quite expectedly, the contextual information was the temperature on that day.

The main manipulation of the experiment was this contextual information. The groups either received no contextual information, somewhat-reliable contextual information or highly-reliable contextual information. The reliability was manipulated by the amount of noise added to the temperature data.

The analysis of the findings suggests that, dependent upon the reliability of the contextual information, there were improvements in the judgmentally adjusted forecasts with contextual information over both the initial forecasts and the judgmentally adjusted forecasts without contextual information. The high-reliability contextual information group obtained the highest accuracies, even better than the statistical forecasts provided. The somewhat-reliable and no-

contextual-information groups did not achieve significant improvements over the initial judgmental forecasts. The authors also mentioned a side-finding which had also been observed in their previous study: participants showed strongly conservative behavior in their initial forecasts; subjects seemed to show a tendency to persistently put too much weight on their initial forecasts over time.

Another experimental study was conducted by Goodwin and Fildes (1999). The main novelty in this study was the inclusion of special events affecting the time-series data. The time-series were introduced as data of sales of a product, and the contextual event was sales-incentives. These incentives served as special events leading to extra amounts of sales during certain periods. The participants were required to generate judgmental forecasts for both normal periods (where the effects of incentives were insignificant) and special periods (where the effects of incentives were significant). Some of the experimental groups were also required to apply judgmental adjustments to their initial forecasts after receiving some statistical forecasts. These statistical forecasts provided highly accurate predictions for normal periods, while they were not so accurate for special periods owing to the fact that the effects of incentives could not be included.

The results obtained pointed out that the inclusion of statistical forecasts led to some improvement in accuracy under some conditions, especially in normal periods with series having high noise levels and complex signals. Another important observation of Goodwin and Fildes was that the discrimination of the subjects between when to use the statistical forecasts as they were and when to adjust them was flawed. It seems that they could not make efficient use of the

information provided to them via forecasts. In an ideal situation, during normal periods forecasters would have introduced no adjustment, or a relatively small one, to the forecasts that were highly accurate, while during special periods they would have taken the less-accurate statistical forecasts as a baseline for adjustment and introduced judgment to decide on the amount of adjustment. Such ideal behavior, however, was not observed. Based on this finding, the researchers decided that designing new methods that would improve the subjects' utilization of statistical forecasts was very important.

This was the main motivation of the Goodwin (2000) study. He tried to explore ways that would lead the subjects to introduce smaller adjustments to reliable statistical forecasts and proper adjustments to the less-reliable forecasts. He employed the same setup as in the previous study; namely, the forecasted series were sales of a product that were affected by incentives during certain periods. In this setting, three simple methods that would affect the judgmental adjustments were proposed. The first method was explicitly asking the subjects whether or not they wished to make adjustments after the provision of the statistical forecasts. The second method differed from the first method in the sense that if the subjects chose to adjust, they were required to convey only the amount of adjustment, not the revised forecasts. The third method was just like the second method with the addition that if the subjects chose to adjust, then they were asked to indicate their reason for adjustment. Thus, in the experiment there were four groups, one group for each of the three elicitation methods and a control group for comparison.

The results obtained from this experiment were quite successful for the aims of the researcher. During normal periods, when the statistical forecasts were highly reliable, the elicitation methods decreased the amount of adjustment conducted on these forecasts, leading to an increase in accuracy compared against the control group. There were no significant differences among the elicitation methods in terms of their effects on adjustment, but the third method seemed to be slightly more efficient than the other methods. During special periods, when the statistical forecasts were less reliable, no significant differences were reported in the amount of adjustment and the accuracy among the four groups. The elicitation methods led to adjustments in the same amounts as in the control group. In the end, the elicitation methods were shown to be successful in increasing the utilization of the statistical forecasts. They discouraged the subjects from making excessive adjustments to accurate forecasts, but they had no effect when the statistical forecasts were less accurate and needed judgmental adjustments.

Aside from these laboratory-based studies, there is also a series of studies which explored the nature of judgmental adjustments on sales forecasts for a real-organizational setting (Diamantopoulos and Mathews, 1989; Mathews and Diamantopoulos, 1986, 1989, 1990). It seems appropriate to review these studies in this category, since in a real-organizational setting, the contextual information is a natural part of the sales-forecasting process and this type of information exerts a strong influence on those forecasts.

Mathews and Diamantopoulos conducted these studies by using sales-forecast data collected from a huge manufacturing company operating in the UK health-care industry. Their major finding was that judgmental revisions (i.e., adjustments, but the term revision was preferred by the authors) conducted by the managers on quantitatively generated sales forecasts led to improvements in accuracy (Mathews and Diamantopoulos, 1986, 1989). Thus, the managers were encouraged to apply judgmental adjustments on forecasts generated by statistical methods.

The second study, published in 1989 (Diamantopoulos and Mathews, 1989), examined two factors that may potentially affect post-revision forecast accuracy. The type of the product being forecasted was the first factor, and the nature of the managers making the adjustments constituted the second factor. The analysis indicated that only the first factor had a significant influence on the outcome of forecast revisions. There were no significant differences among the managers in terms of revision accuracy. Moreover, there seemed to be a positive correlation between the amount of adjustment applied and the amount of improvement obtained.

The last study in the series (Mathews and Diamantopoulos, 1990) was an attempt to discover whether managers effectively select the forecasts on which application of revision is appropriate. The results obtained suggest that the managers were efficient discriminators of the forecasts which needed judgmental adjustments. They selected forecasts for adjustment that would benefit more from the revision process (i.e., less-reliable forecasts with high error levels). Indeed, the

error levels of the chosen forecasts were found to be higher than the error levels of the ones that were left behind.

These findings were contradictory to the laboratory-based findings of Goodwin and Fildes (1999). However, in the Goodwin and Fildes study there were special events affecting the forecasts, and the setting was a generic sales-forecast setting. The major advantage of the Mathews and Diamantopolus studies was the utilization of real sales forecasts and real managers, but these results seem to be limited to the health-care industry, and the generalizability of the results to other industries was a noteworthy concern.

Lastly, Mathews and Diamantopoulos (1990) mentioned a slight bias observable in the managers. The managers had a tendency to adjust pessimistic forecasts more than optimistic ones. That is, the managers chose to employ revisions to the forecasts that were initially on the low side more frequently than to those that were on the high side.

Another study that explores judgmental adjustments under the influence of contextual information was carried out by McNees (1990) on macroeconomic forecasting. There is extensive evidence in the literature that judgment has a strong influence on the generation of macroeconomic predictions and judgmental adjustments are quite common and popular in this area (Young, 1982; Turner, 1990; Donihue, 1993; Clements, 1995).

In the McNees (1990) study, macroeconomic point forecasts like GNP, real GNP, treasury-bill rate, unemployment rate, etc., were used. As the

macroeconomic forecasting process needs much more information than just time-series data, the presence of contextual information is unavoidable. A few well-known forecasters provided the macroeconomic forecasts used in the study. The forecasters first generated the forecasts using quantitative methods, and then judgmentally adjusted them. McNees had access to both the unadjusted and the adjusted forecasts so he could make accuracy comparisons. These comparisons showed the benefits of judgmental adjustments in general. Most often, the accuracy of initial quantitative forecasts was improved through judgmental adjustments. However, this improvement was sensitive to both the macroeconomic forecast type and the forecast horizon. The accuracy of the forecasters decreased over longer horizons. McNees also reported a bias of the macroeconomic forecasters: in some cases, the forecasters showed a tendency to overestimate their judgments by making excessive adjustments to the quantitative baseline forecasts.

The summary of the cited research so far is that the judgmental adjustment of statistical forecasts was frequently shown to increase accuracy over baseline forecasts in both laboratory settings and real settings. However, this improvement is highly dependent upon the accuracy and reliability of the baseline forecasts, as well as the presence and quality of any contextual information. Biases affecting this process were also reported from time to time, which shows that forecast users must be careful when applying their judgments to the provided forecasts.

Sanders and Ritzman (2001) made some suggestions regarding this issue. To achieve adjustments which are as bias-free as possible, they advise the

application of adjustment only on certain occasions. The first occasion is when the forecasting environment is highly uncertain but is subject to predictable changes. The second occasion is when forecast users are able to access highly-relevant contextual knowledge. However, to achieve success, the users should be experienced in properly processing and integrating that knowledge into their forecasts; namely, they should possess a good level of domain knowledge. Furthermore, systematic recording of both the initial forecasts and the adjustments was strongly suggested. In this way, feedback and performance comparisons may provide invaluable clues to forecast users.

Another suggestion for improving the process involves systematic application of judgmental adjustments. Some researchers have criticized judgmental revisions in the sense that they are applied on baseline forecasts in an informal and ad hoc manner by forecast users (Bunn and Wright, 1991; Bunn, 1996). On the other hand, the proposed systematic adjustments require the application of judgment in a structured, methodological or rule-based manner. Some of the main methods offered for this purpose are the analytic hierarchy process—AHP (Wolfe and Flores, 1990; Flores et al., 1992), a computerized system for the automatic adjustment of forecasts (Lee et al., 1990), model-consistent expectations (Bunn and Salo, 1996), neural networks (Lee and Yum, 1998), IF-THEN rules accompanied by fuzzy logic (Ghalia and Wang, 2000) and a forecasting-support system for the integration of contextual information (Webby et al., 2005).

In general, all these studies on systematic adjustments reported to have achieved improvements in the adjustment process. However, they carry problems of implementation. None of the studies reported above documents the applied use of the techniques in a real-life organizational setting. The forecast users need specialized training and have to devote much time and effort to the adjustment process if they want to apply the suggested systematic adjustments. To do this would be a real burden for most managers and forecast users. Due to this problem, it seems that it will be difficult for systematic adjustments to find widespread application. On the other hand, there is extensive evidence in the literature that the criticized intuitive and informal adjustments may lead to improvements in the accuracies of baseline statistical forecasts. Whether there is an improvement or not and the degree of this improvement seems to be contingent upon the quality of the baseline forecasts and the presence, reliability and amount of contextual information accompanying those forecasts. There is no indication that improvement is contingent on whether the adjustment is systematic or not. In a real-life organizational setting, managers and forecast users seem to be fond of intuitive judgmental adjustments, and they are able to attain increased accuracies by applying them.

2.4.3. Research Gaps and Nested Judgmental Adjustments

This review on the judgmental adjustment of forecasts clearly identifies many research gaps in the field for future research. First of all, none of the above studies has directly and systematically investigated the reasons behind judgmental

adjustments. Sanders and Manrodt (1994) and Goodwin (2000a) were the only researchers that mentioned the reasons behind adjustments. However, this issue was not their main focus, and therefore, the knowledge gained from these studies remains too limited to be of any use in the clarification of the issue.

The studies conducted by Mathews and Diamantopoulous (1990) and Goodwin and Fildes (1999) tapped into the effectiveness of the forecast-selection process for adjustment purposes, but they did not research the motivations behind these selections. The reasons and conditions in which forecast users apply adjustments on the forecasts they receive is still largely unknown. Symmetrically, the factors that lead the initial forecasts to be left unadjusted are also unexplored. Questions such as "what is the rationale behind adjustments?", "what are the conditions that make forecast users accept provided forecasts the way they are?" and "under what conditions and for what purposes do forecast users select and apply adjustments?" do not have clear answers yet. Mathews and Diamantopoulous (1990) and Goodwin and Fildes (1999) provide evidence for both the ability and the inability of forecasters in selecting the forecasts for adjustment. At this point, investigating the motivations and characteristics behind the forecast adjustment/acceptance process seems crucial so that we can have a better understanding of when forecast users are able to make an efficient selection of forecasts to be adjusted, and when their selections turn out to be deficient.

The second major research gap is that the judgmental adjustment literature seems to concentrate on adjustments applied on sales forecasts expressed in point format. Only a few studies have investigated adjustments outside the context of

sales. Especially the judgmental adjustment process on financial forecasts is left largely untouched, even though it is critical that we improve our understanding on this matter. This is because just like sales forecasts, financial forecasts are very important for organizations. Finance is a vital activity of firms, and financial forecasts are preliminary tools used to aid this activity.

Another related research opportunity is the nature of judgmental adjustments on other well-known forecasting formats. The use of point predictions has inevitably dominated the research done on the issue; however, as was mentioned before, probabilistic forecasts and especially interval forecasts are as highly-demanded by decision makers as point predictions. Given the situation, research on the judgmental adjustment of interval predictions still remains scarce and limited. As interval forecasts become more available and popular in academia (e.g.,O'Connor et al., 2001; Önkal and Bolger, 2004; Bolger and Önkal, 2004), it seems vital to enhance our knowledge and comprehension of the adjustment process for interval forecasts as well.

Although all the aforementioned research opportunities are essential, there is an issue that carries even greater importance. This issue bears critical significance for forecast users, but has never before been examined and explored in the literature. It is the judgmental adjustments conducted on already-adjusted forecasts, namely multi-tier or nested judgmental adjustments of provided forecasts. Before explaining these concepts, it is useful to talk briefly about information-flow in firms.

In an organization, a piece of information is generally utilized at a level different from the one where it originates. The information enters the firm at a certain point and is then communicated through different units before it reaches its final destination. In each unit that piece of information is processed according to the function of that unit, and then transmitted to the next one in the chain. In this way, information entering the firm experiences a series of operations before it is finally used in the decision-making process.

As important pieces of information themselves, forecasts are no different. A forecast from an external source typically undergoes a series of adjustments in different departments or at different levels before finally being utilized for a decision. For example, a company may require a financial forecast (i.e., a stock-price forecast or an exchange-rate forecast) to plan or hedge its investments. In such a case, an external forecast may be acquired from a consulting company or bank by the finance department of the company. These external forecasts will first be examined in the finance department. After this department examines and appropriately adjusts the forecasts, the adjusted forecasts will pass through to the executive level, where top-level managers further examine them to make the final set of adjustments on the already-adjusted forecasts. In this respect, in a multi-tier or nested manner, a series of adjustments may be implemented on a set of forecasts while they are being transferred from department to department or from one level to another in a company.

The characteristics of these nested adjustments are not known. It is not known whether making adjustments on an already-adjusted forecast produces

different effects than making adjustments on the initial forecasts in terms of accuracy and acceptance of those forecasts. Moreover, the factors affecting the nested adjustment of forecasts and the effects of these factors are also not identified. The exploration of this issue seems extremely important, and it is expected that it has the potential to draw great attention from both forecasting research and practice.

Having made this introduction to the nested judgmental adjustment of forecasts, which will constitute one of the unique contributions of this thesis, it is preferable to pass along to another important concept that is highly relevant to the user perspective. The next and last section of the literature review will focus on the issue of improving the users' acceptance and adjustment process of the provided forecasts.

2.5. Provision of Explanations Along with Forecasts

In their struggle with uncertainty, decision makers generally seek and acquire forecasts from external sources. As was mentioned before, this behavior leads to a distinction of perspectives between forecast providers and users. Forecast providers analyze data and generate forecasts, while forecast users receive or acquire those forecasts and then utilize them in their decision-making process. An important concept in this relationship is the acceptance of the received forecasts by the forecast users.

From the perspective of forecast users, acceptability is directly related with the perceptions of how useful and beneficial those received forecasts are. It can be said that forecast users choose to trust the provided forecasts only if they believe they are not merely wild guesses or groundless numbers, but rather, are justifiable, informative and useful (Goodwin et al., 2004). A provided forecast may be considered justifiable, functional and accurate from a provider's perspective; however, its utilization will be contingent on whether the users are convinced that this is the case. The accuracy of a forecast alone will not make it successful, but its acceptance by the users will. A provided forecast will be of no value if the users of that forecast are not willing to implement it. Additionally, acceptance is also relevant to the amount of adjustment conducted. Decision makers adjust less if they perceive the provided forecasts to be broadly acceptable; otherwise they introduce extensive adjustments or simply cease acquiring forecasts from that source. Due to this nature, this issue carries survival value for forecast providers, since a forecast source whose forecasts are not accepted and acquired cannot continue in business. This undoubtedly demonstrates the need for forecast providers to effectively communicate the quality of their forecasts to the users.

The provision of an explanation about a forecast is a crucial tool for communication between providers and users. In this regard, explanations may prove to be highly effective in improving the acceptance and adjustment process of the provided forecasts. It has been suggested that if the provided advice or forecast is accompanied by a relevant explanation, its chance for acceptance is

improved (Lawrence et al., 2001). Moreover, our own experience with business professionals has revealed that explanations for acquired forecasts are both sought and highly appreciated. Such explanations could assume many forms: they may provide information about the line of reasoning used in generating a particular prediction, they may refer to specific theoretical knowledge for justification, or they may simply describe the procedure or data used.

Research on the provision of explanations along with forecasts is relatively new to the judgmental forecasting field. To the extent of our knowledge, currently, there are very few studies that have specifically explored the effects of explanations on the acceptance of provided forecasts. Given this scarcity of research in the field of forecasting, it is preferable to start this review by first describing the research on explanations based on the knowledge gathered from other related fields such as expert systems/decision support systems, advice taking/giving and organizational justice.

2.5.1. Previous Research on Explanations

Explanations have enjoyed widespread use within the domain of expert systems (ES) and decision support systems (DSS). These systems typically manipulate a knowledge base or model to generate advice and provide this advice to their users in order to aid their decision-making process (see Turban and Aronson (2001) for detailed information on these systems). The acceptance of the advice provided by these systems carries critical importance since it is directly related with its

utilization. If the advice provided is not employed by the users of a DSS then that system can hardly be considered successful or beneficial. There is indeed evidence that decision makers have difficulties in accepting DSS advice provided to them (Kleinmutz, 1990; Davis and Kotteman, 1995), and acceptability is the primary attribute that a successful DSS should have (Fildes et al., 2005).

In order to improve the acceptability of a recommendation, effective communication between the system and the decision makers is obligatory. Serving this function, explanations have been employed extensively within the expert systems and decision support systems domain.

One of the first systems that incorporated an explanation function was an expert system for medical decision making in the 1970s (Shortliffe, 1976). After this milestone work, the interest and research in the area deepened, attempting to find answers to preliminary questions such as why explanations are provided and what the general types and uses of explanations are among expert systems and decision support systems.

The first preliminary question is why users of a decision support system may request an explanation from the system. One of the answers to this question is that decision makers may require an explanation for an output if they perceive that output to be contrary to their expectations (Mao and Benbasat, 2000). The explanation may serve to reduce the amount of discrepancy between the expectations of the user and the advice of the system, or it may simply try to show that the output may in fact be aligned with the user's expectations. In case of an

irregular or abnormal output, explanations may successfully convey the reasons behind those results so that the decision makers will not perceive them as errors and will not attribute them to the inability of the system.

Another reason to demand an explanation may be to enhance problem-solving performance. When attending to a problem, if a difficulty arises, decision makers often require supplementary and timely information to successfully reach a solution. In this sense, explanations might be invaluable in providing that extra information to the decision makers. Gregor (2001) has shown that performance in a cooperative problem-solving task was improved by the use of explanations.

Aside from such single-use problem needs, explanations may also be demanded to facilitate learning for long-term benefits. Decision makers not only utilize extra information to attend to the current problem, but also to address problems that may arise in the future. If they can understand the provided knowledge and how to apply it to the problem, then their effectiveness for future tasks will be enhanced as well. In simpler terms, decision makers always appreciate an explanation if the benefit they receive is greater than the effort they would have spent in its absence (Mao and Benbasat, 2000; Papamichail and French, 2003).

Regardless of the reasons behind them, explanations can be classified with respect to their content (Dhaliwal and Benbasat, 1996; Gregor and Benbasat, 1999; Mao and Benbasat, 2000; Irandoust, 2002). If the explanation is written as an answer for the question “how”, that explanation is often called a “*trace*” or “*line*

of reasoning" explanation. This type generally explains the processes used in a system while generating advice. Which information or rules were utilized, which steps were followed and how the system reached its recommendations are the main focus of these explanations. This type of explanation can be easily generated and presented to decision makers. They do not require any extra work since all the work they need will already have been done to generate the output of the system. Hence, they are the most common type of explanation (Mao and Benbasat, 2000). Moreover, this type of explanation produces more transparent systems. The internal outcome generation of the system can be traced step by step by examining these explanations. This transparency will help the user to better understand the system, leading to improved user perceptions (Irandoust, 2002; Papamichail and French, 2003).

The second type of explanation generally aims to provide the overall problem-solving strategy that is used in the system. These explanations often convey the goals and objectives according to which recommendations are generated. They explain the tactics and planning used to accomplish the task the system is designed for. Therefore, these explanations are often called "*strategy*" explanations. When compared to other explanation types they are more restrictive, and on many occasions this type of explanation may be very difficult to generate.

The last type of explanation is called the "*justification*" type of explanation, and they are written to answer the question "why". They often relate the outcome to the theory or model used to generate it in order to justify the validity of the recommendations. The background and reasoning behind the

advice-generation procedure is presented in this type of explantaion. Ye and Johnson (1995) have demonstrated that justification type explanations are more effective than both "trace" and "strategy" type explanations in improving the acceptance of an expert system recommendation.

Another important issue to be considered is the implications of explanations for decision makers. Based on cognitive learning theory, Dhaliwal and Benbasat (1996) propose that the presence of explanations leads to improved understanding of the provided advice and allows learning from the system. When learning takes place, the performance of the decision makers using the system will be enhanced in future tasks. This increase in accuracy and speed will also be followed by an improvement in user perceptions of the system. Dhaliwal and Benbasat argue that this improvement in perceptions will be noticed in perceptions of ease of use, usefulness, satisfaction and trust. The system will be perceived as an easy to use, useful and satisfactory system. At the same time, it will lead to accurate and quick decisions. In the end, these two aspects will result in higher intentions of future use, and thus, also in the enhanced acceptance of the expert system or the decision support system.

Closely related to the acceptance of recommendations, research on advice-taking demonstrates that the main reason for taking advice is the expectation that the decision maker's judgment will be improved by the receipt of another person's opinion (Harvey and Fischer, 1997; Yaniv, 2004a; 2004b). In organizations, advisors are often chosen due to their superiority over the decision makers in terms of knowledge and expertise in particular areas. When decision makers turn

to their advisors for opinions, there is an inherent belief that a recommendation made by a more knowledgeable and expert person will result in a better final decision than the decision maker could make by himself.

Supporting this line of reasoning, Schrah et al. (2006) argue that decision makers seek advice in order to increase their accuracy. Furthermore, the harder the decision, the greater the demand for advice will be. It has also been indicated that decision makers integrate this advice after conducting their own information search and reasoning about a task and after forming a preliminary opinion. Following this initial generation, they take the advice, integrate it with their preliminary opinion and form a finalized decision. This independent form of advice integration was found to be very effective in improving the accuracy of and confidence in these decisions (Sniezek and Buckley, 1995).

In real business settings, the advice often does not come free-of-charge. Experts are paid for the recommendations they provide to decision makers. The commitment of money for the receipt of advice is also a motivation to use it. The decision makers may feel obliged to use the advice they have put money on. Sniezek et al. (2004) demonstrated that, having paid for expert advice, the utilization of that advice and the subsequent decision accuracy was improved.

Another reason for seeking advice is the need to share the responsibility of an outcome with other individuals (Harvey and Fischer, 1997; Yaniv, 2004a; 2004b). When a manager makes a decision after receiving advice from a consultant, the final decision will be the responsibility of both the manager and

the consultant. In case the outcome of that decision turns out to be bad, it will be psychologically relaxing for the manager to know that he was not the only one behind the decision. The burden of a probable negative outcome will be alleviated.

Social pressure also has a say in the acceptance of advice (Harvey and Fischer, 1997). When an advisor submits his recommendations to a decision maker, that decision maker does not wish to seem unappreciative of the help offered him. He feels that if he refused the help of an advisor, that advisor would be reluctant to offer his recommendations when the decision maker really needed it. Therefore, he feels that he should accept and use the advice.

This willingness to employ advice is also observed in the case of recommendations which may be derived from unreliable pieces of information. Decision makers generally show a tendency to access and use any information that could be helpful in their decisions, even though this information may be based on rumors that may turn out to be false in the end and, thus, may have negative effects on accuracy (DiFonzo, 1997). Harvey and Fischer (1997) stated that a provided recommendation will never be completely rejected, and will be utilized at least to some extent.

On the other hand, advice will never receive full and complete acceptance, either; it will always be subject to some discounting. The decision makers tend to favor their own opinions over the opinions of others, and they partially ignore the recommendations presented to them. This effect is known as the “self/other” effect and it has been observed in a consistent manner (Yaniv and Kleinberger,

2000; Harvey and Harries, 2004; Yaniv, 2004a; 2004b). Regarding this effect, Yaniv (2004a) argues that decision makers allocate more weight to their own opinions than to those coming from advisors. The greater the difference between the opinions, the more discounting advisor recommendations will receive. Extreme or nonconforming advice will be severely discounted. Furthermore, if the knowledge and expertise of a decision maker increases relative to the advisor, the favoring of the decision maker's own opinion will also increase. A decision maker trusts his own judgment more when he receives advice from a relatively novice consultant rather than an expert consultant.

This effect generally results in an inefficient combination of a decision maker's self-estimates with the recommended ones (Harries and Harvey, 2000; Yaniv 2004a; Harvey and Harries 2004). Even though it is normal to observe some improvement in the accuracy of the final estimate with respect to a case where no advice was present (Yaniv, 2004a; 2004b), the final combination of opinions are generally observed to be far from optimal combinations (Yaniv 2004a; Harvey and Harries 2004). Yaniv (2000b) proposes that the aggregation of opinions will be more accurate if the opinions are produced independently from each other and if there is no significant relationship between the advisors and the decision maker.

In this sense, decision makers seem to be inefficient in combining advice with their own opinions. However, their ineffectiveness in this area is not paralleled in their assessment of the quality of the source providing the advice. Harvey et al. (2000) argue that decision makers are better at determining the

quality of advice than integrating that advice. They can effectively discriminate a good recommendation from a bad recommendation, and discounting is also applied accordingly. A bad recommendation is discounted more and is given less weight in the final decision than a good one. The quality of the recommendations coming from a source may shape its reputation rapidly, and decision makers will understand that the source is a good one. At the same time, it may be even easier for that source to lose its good reputation. Yaniv and Kleinberger (2000) state that there is asymmetry in the formation of reputation, so that losing a good reputation is easier than forming one.

Swol and Sniezek (2005) investigated some factors that may affect the acceptance of advice. They focused on five factors: advisor confidence, advisor accuracy, the decision maker's trust in the advisor, the decision maker's prior relationship with the advisor and lastly, the decision maker's power to pay for the advisor's recommendations. Out of these five factors, advisor confidence was found to have the most significant impact on the decision maker in accepting and utilizing the provided advice. If an advisor has confidence in the recommendations he is making, there is a greater chance that those recommendations will be accepted and used. In line with these findings, in a previous study, the researchers demonstrated that the confidence of an advisor was highly related to the decision maker's trust in that advisor as well (Sniezek and Swol, 2001). Confidence is generally perceived as a sign of expertise, eventually leading to the formation of a decision maker's trust in his advisor.

2.5.2. Explanations and Forecasting

Within the forecasting domain, Goodwin and Fildes (1999) were the first to mention and briefly utilize explanations accompanying forecasts. As was mentioned before, in this study the authors tried to investigate the effect of statistical forecast provision on the adjustment of sales forecasts affected by special events (i.e., sales incentives). One of the many factors present in this study was short trace-type explanations. These explanations were provided to one of the experimental groups in addition to statistical forecasts. However, the impact of those explanations was minimal, and therefore, they were only briefly mentioned and commented upon. Moreover, explanations were not the main focus of the study and lacked appropriate manipulations, so that no beneficial knowledge about explanations was gained from this study.

The only proper study that specifically examines the effects of presenting explanations on forecast acceptance was conducted by Lawrence et al. (2001). In this study, along with the forecasts, the decision makers were provided with two kinds of explanations that differed with respect to content. The first type of explanation was technical in nature, and explained the method used in generating the forecast and the reasons behind the selection of that method. The second type was a managerial explanation conveying the meaning of the forecast to the decision maker in the context of the time-series given. The results of the study indicate that the presence of either kind of explanation had a highly positive

impact on the acceptance of the provided forecasts and the confidence of the users in those forecasts.

With a slightly different focus, a recent study involved explanations within a forecast-generation task (Mulligan and Hastie, 2005). In this study the effects of the order of sentences in the provided explanations were investigated. The primary question of the researchers was whether the influence of explanations would differ if the sentences forming them were presented in story order (i.e., context, problem, goal, plan and outcome) or a mixed order of these five semantic pieces. After the explanations were provided, participants were asked to provide directional stock-price forecasts and confidence ratings on the forecasts they had generated. In this context, explanations were not used to increase the acceptance of provided forecasts, which is the main focus of this thesis and the Lawrence et al. (2001) study. Instead, explanations served as general information to aid in the generation of forecasts on the stocks of the mentioned firms.

The main findings of Mulligan and Hastie (2005) indicate that the only explanations which had a direct effect on generated forecasts were those presented in story order. Moreover, the semantic piece which had the highest impact on the participants was the outcome sentence. Based on these results, it is plausible to suggest that forecast users can better understand and utilize explanations when they are presented in a semantically meaningful order.

Given the scarcity of studies, it can be confidently argued that research in explanation provision for the forecasting field is in its infancy stage. The

characteristics of explanations, the factors affecting them and their effects on the acceptance/adjustment of the accompanying forecasts are largely unknown.

As previous research in other areas suggest, the most important characteristics of explanations can be considered to be related to their content and structure (Shapiro et al., 1991; 1994). Content characteristics of explanations deal with *what* the explanations are all about, while structural characteristics of explanations are more related to *how* this content is presented to the users. For instance, the technical or managerial nature of explanations is more closely related with content characteristics, while the length of explanations and style of language used are more closely related with structural characteristics. Lawrence et al. (2001) was successful in the investigation of content characteristics; however, there is no research available for the effects of structural characteristics on the acceptance and/or adjustment of provided forecasts. Thus, there is urgent need for research about explanations and their structural characteristics, since they carry paramount importance for proper communication between forecast providers and forecast users. Without the proper examination of explanations, the investigation of user and provider perspectives cannot be considered complete.

As such, the literature review part of this thesis is complete. At this point, it seems appropriate to turn to the research questions and details of the research methodology that is targeted to fill the aforementioned gaps in judgmental forecasting research.

CHAPTER 3

RESEARCH QUESTIONS

One of the crucial points of the contemporary forecasting literature is the distinction between user and provider perspectives. The majority of research up to the last decade has explored the issues related with provider perspective, involving forecasting methods, selection of those methods and generation of forecasts.

Only recently has the user perspective begun to take the focus, and a few studies on the subject have begun to appear. It is the users who ultimately hold, inspect, and adjust the forecasts, and it is they who will decide whether the forecasts will be employed or not. In this respect, the last word in the forecasting practice will always belong to the users. This fact by itself indicates the significance of this perspective. Since many areas of this perspective remain largely unexplored, the forecasting literature needs research directed primarily toward exploring the needs, desires and criteria of forecast users, in order to develop effective new techniques and methodologies that will benefit both the providers and users of forecasts. Many fundamental questions need answering,

such as what the expectations of users are from the forecasts they receive, how important accuracy is to users, whether accuracy is as important as researchers wish to believe, and most importantly, what the notion of forecast quality is in a forecast user's mind.

Forecast users either accept the provided forecasts the way they are or introduce judgmental adjustments to them, which are proportional to the level of acceptance of those forecasts. Owing to their frequent use in business practice, judgmental adjustments have drawn much attention from academia; nevertheless, they still offer many opportunities for research. For instance, the motivations of forecast users in adjusting forecasts and the situations in which they introduce adjustments have not been investigated properly. Accompanying this issue, the situations in which users choose not to adjust forecasts are also unknown. The exploration of these motivations and situations will undoubtedly bestow invaluable insight to our knowledge regarding user perspective.

Judgmental adjustment of forecasts mostly concentrates on point forecasts in the context of sales. The properties and effects of the adjustment process on other forecasting formats (especially interval forecasts) and in other settings (especially the financial context) remain to be explored.

Regarding judgmental adjustment of forecasts, a highly exciting research opportunity exists through multiple or nested judgmental adjustments, namely, the adjustment of already-adjusted forecasts. This topic seems to be highly relevant to the user perspective. To clarify, in an organization, information does not stay in a

single unit or department. It is passed from one unit to another while being inspected and processed. Being information themselves, forecasts may also change hands and may be subject to consecutive adjustments before reaching their final destination. Therefore, nested adjustments are highly important in organizational settings.

However, no research done particularly on this issue has been located. The exploration of this process will form one of the unique contributions of this thesis to the current academic field. It is not known whether making adjustments to an already-adjusted forecast produces different effects than making adjustments to the initial forecasts in terms of accuracy and acceptance of those forecasts. Moreover, the factors affecting nested adjustment of forecasts and the effects of these factors have also not been identified.

Another relevant issue on this topic is whether the presence of the initial (i.e., unadjusted) forecast along with the already-adjusted ones will be beneficial for decision makers in their attempt to adjust the already-adjusted forecasts. In other words, if the forecast user receives both the initial forecast and the adjusted forecast at the same time instead of receiving only the adjusted forecast, will he/she benefit from this presentation while further adjusting those forecasts?

For the improvement of the adjustment and acceptance process of provided forecasts, explanations seem to be critical tools. They form a highly important communication vessel between the two perspectives. Despite their significance, the research on the provision of explanations in the forecasting field is in its

infancy stage and has acquired attention only recently. The characteristics of explanations, the factors affecting them and their effects on the acceptance/adjustment of accompanying forecasts are largely unknown.

Basically, not all types of explanations are equally effective. The persuasiveness of explanations could be directly related to their characteristics; in particular, structural and content characteristics of explanations may play a critical role in their acceptability. Lawrence et al. (2001) have investigated the content characteristics (i.e., whether technical or managerial) of explanations. Therefore, investigation of the structural characteristics (i.e., the length of explanations and the confidence conveyed via language-style) is urgently required.

Additionally, our own experience with managers has shown that forecast explanations are demanded and appreciated by business professionals. Because of this, it is justifiable to expect them to influence not only the initial forecast adjustments, but also the nested adjustment of these forecasts. In the process of nested adjustments, how the forecast users will react to the explanations regarding the adjustment of initial forecasts and what their effects will be constitute important research endeavors.

Aiming to capture all these research opportunities, the proposed research questions are the following:

1) What are the expectations and quality perceptions of forecast users regarding financial forecasts?
    a) What do users expect from financial forecasts?

b) How important is the accuracy of financial forecasts? Does "accurate" correspond to the smallest amount of error, or to a tolerable margin of error?

c) What does "quality of forecast" mean in the eyes of forecast users? Is it equivalent to "accuracy"?

d) Is a forecast that meets the expectations of users considered a high quality forecast?

2) What are the reasons and situations behind judgmental adjustment and acceptance of provided forecasts by forecast users?

a) What are the reasons/situations in which forecast users choose to adjust provided financial forecasts?

b) How frequently are these reasons/situations encountered?

c) What are the reasons/situations in which forecast users choose to accept provided financial forecasts the way they are?

d) How frequently are these reasons/situations encountered?

3) What are the effects of the structural characteristics of explanations on the adjustment of financial forecasts?

a) What is the effect of the length of explanations (either long or short) on the adjustment of point forecasts?

b) What is the effect of the length of explanations (either long or short) on the adjustment of interval forecasts?

c) What is the effect of conveyed confidence (either confident or weak) on forecast users in the adjustment of point forecasts?

d) What is the effect of conveyed confidence (either confident or weak) on forecast users in the adjustment of interval forecasts?

e) What is the effect of the structural characteristics of explanations on the accuracy of point forecasts?

f) What is the effect of the structural characteristics of explanations on the accuracy of interval forecasts?

g) What is the effect of length on the perceived information value of explanations?

h) What is the effect of conveyed confidence on the perceived information value of explanations?

4) What are the basic properties and effects of nested adjustments on financial forecasts?

a) How will forecast users adjust point forecasts when they think they have received already-adjusted ones?

b) How will forecast users adjust interval forecasts when they think they have received already-adjusted ones?

c) Does making adjustments to already-adjusted forecasts produce different effects than making adjustments to the initial point forecasts?

d) Does making adjustments to already-adjusted forecasts produce different effects than making adjustments to the initial interval forecasts?

e) How is the accuracy of point forecasts affected if the forecast users think they have received already-adjusted forecasts?

f) How is the accuracy of interval forecasts affected if the forecast users think they have received already-adjusted forecasts?

g) What is the effect of the receipt of the initial point forecasts in addition to the adjusted ones on the task of further adjustment? Will it benefit the accuracy?

h) What is the effect of the receipt of the initial interval forecasts in addition to the adjusted ones on the task of further adjustment? Will it benefit the accuracy?

5) How does the provision of explanations affect the nested adjustment of financial forecasts?

a) How will forecast users adjust point forecasts when they receive already-adjusted ones accompanied by explanations about the rationale of the initial set of adjustments? What will be the effect of explanation provision on accuracy?

b) How will forecast users adjust interval forecasts when they receive already-adjusted ones accompanied by explanations about the rationale of the initial set of adjustments? What will be the effect of explanation provision on accuracy?

c) What is the effect of the receipt of both initial (unadjusted) forecasts and explanations on the nested adjustment of point forecasts? What is its effect on accuracy?

d) What is the effect of the receipt of both initial (unadjusted) forecasts and explanations on the nested adjustment of interval forecasts? What is its effect on accuracy?

The research that targeted these questions was composed of two parts. The first part of the research sought answers to the first two questions. These questions are specifically related with user perspective, and thus, this part of the research included a study carried out with business professionals who make use of forecasts in their routine business practice. Qualitatively oriented methodology seemed to be the best for this investigation. Accordingly, the first part of the research included a field survey conducted with practitioners and professionals.

The second part of the research sought answers to the remaining questions, namely the third, fourth and fifth questions. These questions are related with the effects of nested adjustments and explanations contingent on the aforementioned factors. The best way to explore these questions was laboratory experimentation for the controlled manipulation of the factors and identification of their effects. Therefore, the second part of the research included three laboratory-based studies to be employed for each of the third, fourth and fifth questions.

CHAPTER 4

SURVEY ON FORECAST USING PRACTICE

4.1. Methodology and Design

Survey studies are popular qualitative data collection tools that allow statistical analysis and inferences based on data (Frankfort-Nachmias and Nachmias, 1996). Owing to this property, the primary instrument of the first part of the research was a survey study conducted via a questionnaire prepared specifically according to our needs.

To prepare this particular questionnaire, a preliminary study was required to gain some insight and perspective on the forecast utilization and adjustment habits of business professionals and practitioners. For this purpose, structured interviews were conducted. These interviews formed a basis for the survey. In other words, they served as an exploratory study that provided us with ideas, directions and points of origin for the preparation of the questionnaire used in the survey.

The interviews were conducted with five business professionals who had stated their familiarity with financial forecasting. The interviewed professionals were as follows:

a. The CFO of a large alcohol & spirit factory
b. The branch manager of a large commercial bank
c. The CFO of a construction company
d. The MIS director of a government bank
e. The CFO of another large construction company

These interviews were conducted in a structured manner (please consult Frankfort-Nachmias and Nachmias (1996) for more details on structured interviews). The interview questions were asked in the same order and in the same way as far as the time allocated to us and the conditions allowed. All of the interviews were done face-to-face in the professionals' office except for a single case. In that case, the professional (the CFO of a large alcohol & spirit factory) had to be interviewed briefly on the telephone, since he had only allowed us that possibility. Otherwise, the interviews lasted around an hour on average. The interviews were audio taped if permission was acquired. Field notes were written based on these recordings. When recording was not permitted, field notes were written down immediately after the interviews to minimize cognitive distortion and forgetting.

Highlighted in these interviews, the following points about the financial forecasting practice are worth mentioning:

- Firms do not rely on a single source for the provision of forecasts. They try to obtain forecasts from different sources and in different formats if possible.
- Judgmental adjustments are frequently applied to received forecasts.
- The accuracy of forecasts is not the sole criteria for consideration.
- Interval forecasts are known and utilized in addition to point forecasts. Alternative forecasts and scenarios are also employed.
- Explanations regarding the rationale behind forecasts are sought and appreciated.
- Forecasts prepared by banks and specialized firms are not always fully trusted or accepted. Some financial managers find them leading and speculative.

Interpretations based on these interviews, coupled with suggestions from the literature, enabled us to develop the seven-page questionnaire used in the survey (The questionnaire form is provided in Appendix A). It was organized around four main concepts: the expectations of forecast users, the reasons/situations in which forecast users choose not to adjust, the reasons/situations in which forecast users choose to make an adjustment, and finally, the users' perceptions of quality in the forecasts.

At the start of the survey, a one-paragraph cover letter was provided conveying our primary aim and purpose along with the scientific value of the study. The standard discourse regarding anonymity and the generalization of results was also disclosed in the letter.

After the cover letter, there were general questions asking for details about the participant. These questions inquired about the profession, position and experience of the practitioner and asked about their familiarity with one or more kinds of financial forecasts. There was also a screening question to eliminate professionals who had no experience with forecast use in their business practice.

These general questions were followed by the first main concept in the survey, namely, the expectations of users from the forecasts they are using. In order not to lead the participants, we first asked about their expectations using an open-ended question. After this, six possible expectations derived from the initial interviews and literature were given. The participants were requested to provide their agreement with these expectations via a 7-point Likert scale. These expectations were:

a) The forecast should have the smallest amount of error possible.
b) A forecast should have a tolerable amount of error. It does not need to be the smallest.
c) A forecast should have a plausible and justifiable basis and assumptions.
d) A forecast should include scenarios and alternative forecasts against a variety of circumstances.
e) A forecast should have a reasonable cost.
f) A forecast should be timely.

The second and third major themes of the survey came after the questionnaire items on user expectations. Following a few questions about acquiring and using forecasts from multiple sources, an open-ended question was asked to learn about the reasons or motivations for which forecast users choose not to adjust a provided forecast. Next, another open-ended question inquired about the reasons or motivations behind making adjustments. These two open-ended questions were given together to allow the participants to consider these two sets of reasons in comparison to one another, and to think about them freely based on their past experience.

Afterward, the participants were given some predetermined reasons and situations, each of which they were asked to provide frequency-of-occurrence and ratings-of-importance for on 7-point Likert scales. Like the predetermined expectations, these reasons and situations were also generated from the preliminary interviews and existing research in the literature. There seemed to be eight reasons and situations in which business professionals might accept a forecast the way it was. These were as follows:

a) The knowledge and experience of the user is not adequate to make an adjustment.
b) The source providing the forecasts is well-known and famous.
c) The source of the forecast is believed to be unbiased and objective.
d) The methods and analysis in the provided forecasts are highly persuasive.

e) The presentation and the style of language of the provided forecasts are highly persuasive.
f) The explanations provided with the acquired forecasts are highly persuasive.
g) The forecast user does not want to be responsible for the consequences of the adjustment.
h) The forecast user is not authorized to make any adjustment.

There seem to be nine reasons and situations in which business professionals might adjust the provided forecasts. These are as follows:

a) The forecast user wishes to integrate his knowledge, experience and initiative into the provided forecast.
b) To integrate unexpected events and new information into the provided forecasts.
c) The source providing the forecasts is small and barely-known.
d) The source of the forecast is believed to be biased and leading.
e) There are extreme forecasts present.
f) To gain control of and take responsibility for the acquired forecasts.
g) The methods and analysis used in the provided forecasts are not persuasive enough.
h) The presentation and the style of language used in the provided forecasts are not persuasive enough.
i) The explanations provided with the forecasts are not persuasive enough.

These survey items were followed by some detailed questions regarding the adjustment process as experienced in the participant's routine business practice. Some questions about the presence and usefulness of a forecast-feedback mechanism accompanied these. Then came the final major theme of the survey study, which was about perceptions of the quality of an acquired forecast.

What the professionals understood from the concept of a high-quality forecast was first asked by using an open-ended question to allow free thinking. Since expectations from a forecast can be directly related with perceptions of quality, the question following inquired whether this was indeed the case. The survey ended by requesting the participants to give ratings of agreement on six previously-used expectations in order to elaborate on the relationship between expectations and quality.

Having provided some details about the structure and items of the questionnaire, we shall now discuss the issues regarding the population of the survey study and the sample plan used.

The population of the survey study was all the organizations in Turkey that utilize financial forecasts. It would have been excellent if there had been a database, society or NGO keeping track of the forecasting professionals and practitioners in Turkey; however, this was not the case, so there was no available way to determine which firms utilized financial forecasts. On top of this complexity, the administration of a survey study to real-life professionals involves many inherent difficulties. Professionals generally have very limited time and

they lack interest in anything beside their routine business activities. Our own experience and evidence in the literature have shown that professionals are generally very reluctant to participate in such academic studies and their overall response rate is generally very low.

Given these conditions, the best available option seemed to include firms that were known through formal or informal contacts to use financial forecasts, or firms that had been referenced as using financial forecasts. It was decided that when participation in the survey was requested through personal acquaintance and reference, the interest and concentration of the professionals would increase dramatically. This indeed occurred. The initial surveys started with professionals that were known thorough personal contacts. After these surveys, the surveyed professionals informed and asked some of their contacts in other firms to participate in the survey, and these professionals were surveyed next. In this way, the sample size grew larger and extended to a large variety of industrial sectors.

In this respect, the sampling plan was a mixture of a snowball sample and a convenience sample. Even though both of these sampling methods provide non-probability samples, they are known to achieve representative samples, and thus, they have been used extensively in marketing research (Frankfort-Nachmias and Nachmias, 1996; Burns and Bush, 2000).

For the administration of the survey, a supervisor and two surveyors were hired. The supervisor selected for the job had ample experience in survey research and had previously held various positions (including branch manager) in a

respectable marketing research and consultancy company in Turkey. She was mainly responsible for the coordination of the surveyors and the quality of the questionnaire forms received. She also had access to a large base of business professionals with whom she had had contact in her previous projects.

The two surveyors were university students who had past experience as field surveyors. They had good references and had been involved in numerous projects with the supervisor before. The primary duty of the surveyors was to contact the professionals and make appointments with them for the administration of the survey. Afterwards, they were required to attend the appointment and conduct the survey. At the same time, they acquired new contact information from the surveyed professionals in order to include it in the next set of surveys.

In the end, this data-collection process took around three months, and a total of 124 questionnaires were collected from business professionals who used financial forecasts in their routine business practice. Confidential information about the identities of these participants suggested that the surveyed sample was representative in terms of company size and type. A good industry-wide coverage including a wide range of sectors was also achieved. The industries involved were basic materials (i.e., chemicals and mining); construction, electrical & electronics and other general industrials; consumer goods & services; health care; telecommunications & technology; and financial services.

4.2. Results

The surveyed practitioners in the participating companies were mainly composed of managers, CEOs, executive board members and partners/owners. Without exception, all of the 124 respondents indicated that they were using financial forecasts in their routine business conduct. Table 1 summarizes the distribution of the participants with respect to their position in their companies and their experience in using financial forecasts.

Table 1. The position and forecast-using experience of the participants. The numbers in parenthesis indicate the actual number of responses in the category.

Practitioner position	
Partner/owner	46.8% (58)
CEO/ executive board member	16.9% (21)
Manager	28.2% (35)
Staff member	8.1% (10)

Financial forecast-using experience	
Less than 3 years	26.2% (32)
4-9 years	40.2% (49)
More than 9 years	33.6% (41)

A high percentage of the participants (73.8%) claimed to have more than four years of forecast-using experience, with 33.6% of them constituting a highly experienced group having over nine years of familiarity with financial forecasts. The distribution of the practitioners strongly indicates that the data collected is suitable for gaining good insight into actual forecast-using practices in companies.

These practices were reported to involve a variety of different forecast types on various financial instruments. Foreign-exchange rate and parity forecasts

were utilized by the majority of the respondents (68.5%); 29.0% of the practitioners employed interest-rate/bond-price forecasts; 21.0% worked with forecasts on investment funds; 16.1% utilized stock-price forecasts; and 7.3% of the practitioners employed macroeconomic predictions. Some respondents, constituting 13.7% of the whole group, claimed to receive more specialized forecasts, such as predictions on sector financials and forecasts on new economic/financial instruments and investments. As may have already been observed from the percentages, 38.7% of the practitioners asserted that they use more than one type of financial forecast.

Aside from forecast type, the number of forecast sources also seem to vary among the users. Only 25% of the practitioners asserted that their externally acquired predictions come from a single source. The rest, 75% of the respondents, stated that they acquire forecasts from multiple sources. When there is more than one source, there also exist quite a few alternatives to proceed with these various forecasts. The first and simplest alternative may be selecting a forecast among multiple ones and sticking solely to that single prediction. Thirteen percent of the respondents who stated that they receive multiple predictions said that they adhere to this alternative. They select advice from only one source and disregard the rest.

Other alternatives involve using combinations of multiple recommendations. Sixty-nine percent of the practitioners reported that they combine various forecasts using their own judgment and experience, and eventually use this resulting combination. Likewise, 11% of the practitioners also claimed to use a combination of predictions; however, their methods of blending

involved simple statistical methods. The remaining 7% declared that the forecasts they receive are distinct, and thus, not suitable for combining.

Regardless of the type and sources of external forecasts, judgmentally adjusting these predictions seems to be a common practice for forecast users. Among all the respondents to the survey, only 23.4% of them indicated that they rarely employ judgmental adjustments to the forecasts they receive. The remaining 76.6% of the practitioners claimed to apply adjustments on an occasional-to-frequent basis. On the perceptions of benefits gained from making judgmental adjustments, 41.1% of the practitioners believed that the application of adjustments increases the accuracy of the forecasts quite frequently, while 46.0% of them believed this improvement in accuracy occurs only occasionally. The remaining 12.9% believed that only on rare occasions are the adjustments beneficial in terms of accuracy improvement. Similarly, 85.5% of forecast users believed that judgmental adjustments applied to external forecasts increase their persuasive power on an occasional-to-frequent basis. This apparent direct association between judgmental adjustments and the benefits expected from their application is statistically important. There was a significant positive correlation between the frequency of judgmental adjustments and the frequency of accuracy improvements ($r=0.541$, $p=0.002$). A similar relationship can be observed between the frequency of judgmental adjustments and the frequency of improvement in the persuasiveness of predictions ($r=0.526$, $p=0.002$).

Another general observation from the survey concerns the presence of a feedback mechanism as part of the forecasting practice. The participants who

stated that they utilize feedback on either a systematic or personal level made up only 35.5%. The remaining 64.5% do not employ any sort of feedback in their routine forecasting activity.

It may be recalled from the methodology part that there were four main themes for the survey: expectations of users from the forecasts, their perceptions of forecast quality, the reasons and motivations when the practitioners choose to make an adjustment and the rationale when they refrain from making one. In the next two sections, a preliminary analysis on these four themes will be presented. Afterwards, a more detailed and comprehensive analysis will follow. One thing to emphasize is that in all sections, every rating is done on a seven-item scale; therefore, every score mentioned is out of seven unless otherwise noted.

4.2.1. User Expectations from Forecasts and Perceptions of Quality

Table 2 represents the participants' ratings of agreement provided to the dimensions of forecast expectations. As is evident from the table, the highest ratings for expectations were given for the dimensions of timeliness and forecasts having a sound and justifiable basis and assumptions. Forecast users seem to demand and value most the forecasts that are at their disposal when they need them (i.e., timely), as well as having them based on reliable and justifiable grounds and suppositions. The accuracy of the acquired forecasts also seems to be important, and the practitioners expect the predictions they receive to have low levels of error. They even anticipate that these forecast errors will be as small as

possible rather than at a tolerable level, and this difference in expectations is statistically significant ($t_{123} = 3.99$, $p < 0.0001$). Among all the aspects, forecasts having reasonable costs appear to have received the lowest ratings. The practitioners seem to be willing to pay well for solid and accurate forecasts.

Table 2. Expectations of forecasts users

	Mean rating
A forecast should be timely.	6.44
A forecast should have sound and justifiable basis & assumptions.	6.43
A forecast should have the smallest amount of error possible.	6.38
A forecast should include scenarios and alternative forecasts covering a variety of circumstances.	6.02
A forecast should have a tolerable (not necessarily minimal) amount of error.	5.83
A forecast should have a reasonable cost.	5.42

Table 3 displays the ratings for the dimensions representing perceptions of a high-quality forecast. Similar to expectations, perceptions of quality seem to have received the highest ratings on timeliness, followed by the accuracy aspect. Again, the smaller the amount of error, the better from the quality perspective. A high-quality forecast with the smallest amount of error received significantly higher agreement ratings than a prediction having a tolerable error level ($t_{122} = 5.15$, $p < 0.0001$). Predictions based on a reliable basis and assumptions are also deemed important for quality, receiving very similar ratings. The least favorable dimension was the cost dimension. Similar to the case in expectations, practitioners seem to disagree with the idea that forecasts having reasonable costs is an essential attribute of quality.

Table 3. Perceptions of a high-quality forecast.

	Mean rating
A high-quality forecast is timely.	6.31
A high-quality forecast has the smallest amount of error possible.	6.29
A high-quality forecast has sound and justifiable basis & assumptions.	6.24
A high-quality forecast includes scenarios and alternative forecasts covering a variety of circumstances.	6.14
A high-quality forecast has a tolerable (not necessarily minimal) amount of error.	5.63
A high-quality forecast has a reasonable cost.	5.18

The ratings provided for both the expectation and quality perceptions appear to be closely related to each other. On this issue, another question asked the participants directly whether a forecast satisfying their expectations could be considered to be a high-quality forecast. The mean rating to this question was 6.17, indicating a strong agreement. Both of these factors have led us to further investigate their relationship via canonical correlation analysis. The results of this analysis indicate that the first canonical correlation linking the two sets was highly significant (Chi-sq_{36} = 172.54, $p<0.001$). The variability that the expectation dimensions can accommodate in the quality aspects (i.e., redundancy) was 32.13%, while the redundancy of the quality aspects on expectation aspects was 28.94%. These redundancy levels and the significant first canonical variates convincingly indicate that there is indeed a high positive association between the two sets of dimensions.

4.2.2. Reasons and Motivations for Adjustment

The dimensions that were investigated for adjusting/not adjusting behaviour are provided in Table 4 as matching pairs. Two sets of questions explored these dimensions. The first set asked for frequency ratings for the occurrence of these events, and the second set asked for their importance ratings. Figure 1 provides these mean ratings for the whole group.

Table 4. Reasons and motivations for adjusting/not adjusting behaviour

not adjusting	adjusting
• My knowledge and experience on the subject is not adequate to make an adjustment.	• To integrate my knowledge, experience and initiative to the forecasts.
	• To reflect the unexpected events and new information in the forecasts.
• The source providing the forecasts is well-known and famous.	• The source providing the forecasts is small and barely-known.
• I believe that the source providing the forecasts is unbiased and objective.	• I believe that the source providing the forecasts is biased and leading.
	• To intervene in case of extreme forecasts.
• I do not want to be held responsibility for adjusting the forecasts.	• To gain control of and be responsible for externally acquired forecasts.
• The methods and analysis used in the acquired forecasts are persuasive.	• The methods and analysis used in the acquired forecasts are not persuasive enough.
• The presentation and style of language used in the acquired forecasts are persuasive.	• The presentation and style of language used in the acquired forecasts are not persuasive enough
• The explanations provided with the acquired forecasts are persuasive.	• The explanations provided with the acquired forecasts are not persuasive enough.
• I'm not permitted to make an adjustment.	

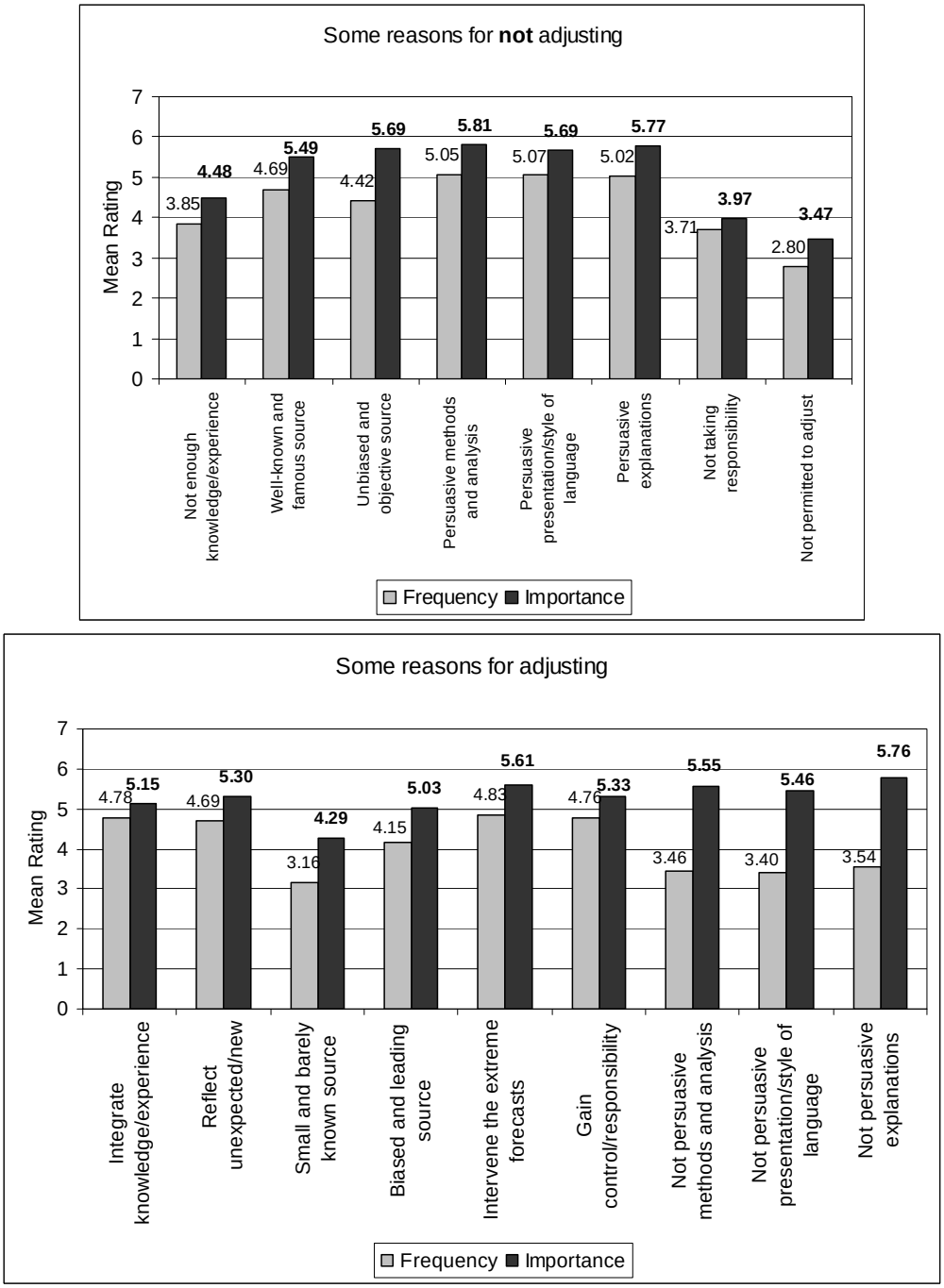


Figure 1. The frequency and importance ratings

Overall, the most frequent reasons for refraining from making any adjustments were related with the persuasive power of the provided forecasts. When methods, analysis, presentation, style of language and explanations

provided with the predictions were perceived to be influential, the practitioners more often accepted them without changes. At the same time, they also provided high importance ratings for them. The source-related reasons were also frequent and important. If the source of the forecast was well-known and reliable, this also constituted a focal motivation in order not to adjust. The least frequent reason for not making an adjustment was having no authority to employ revisions. Not only was this dimension rare, it was also perceived to be less important than the others.

For the cases where judgmental adjustments were applied, the most frequently occurring motivation was related with making some contribution to the forecasts. When the practitioners felt the need to integrate their own knowledge and experience, to incorporate unexpected events, to intervene in case of extreme forecasts or simply to gain responsibility for external forecasts, they frequently chose to apply judgmental adjustments. They also thought that these motivations carried high importance. The absence of persuasive cues seemed to be perceived as a crucial factor as well. Even though it occurred less frequently, a lack of convincing methods, analysis, presentation, style of language or explanations was an essential reason for introducing an adjustment. The least frequent reason for adjusting was the case where the predictions were acquired from a small and barely known source. This reason received the lowest importance rating as well.

These points strongly suggest that the importance ratings provided to the dimensions for both adjusting and not adjusting occasions seem to be highly related with each other. This relationship was further investigated via canonical correlation analysis. The results of this analysis have shown that the first

canonical correlation linking the two sets was highly significant (Chi-sq_{72} = 116.13, $p<0.001$). The variability the reasons for making adjustments can accommodate in the reasons for refraining from introducing adjustments (i.e., redundancy) was 17.95%, while the redundancy of the reverse case was 15.66%. These redundancy levels and the significant first canonical variates are strong indicators for the presence of a high positive association between the two sets of dimensions.

A comparative look at the common reasons and motivations for introducing an adjustment or refraining from making one revealed the presence of some similarities and differences in their strengths to induce adjusting behaviour. Figure 2 provides these common reasons with their respective importance ratings.

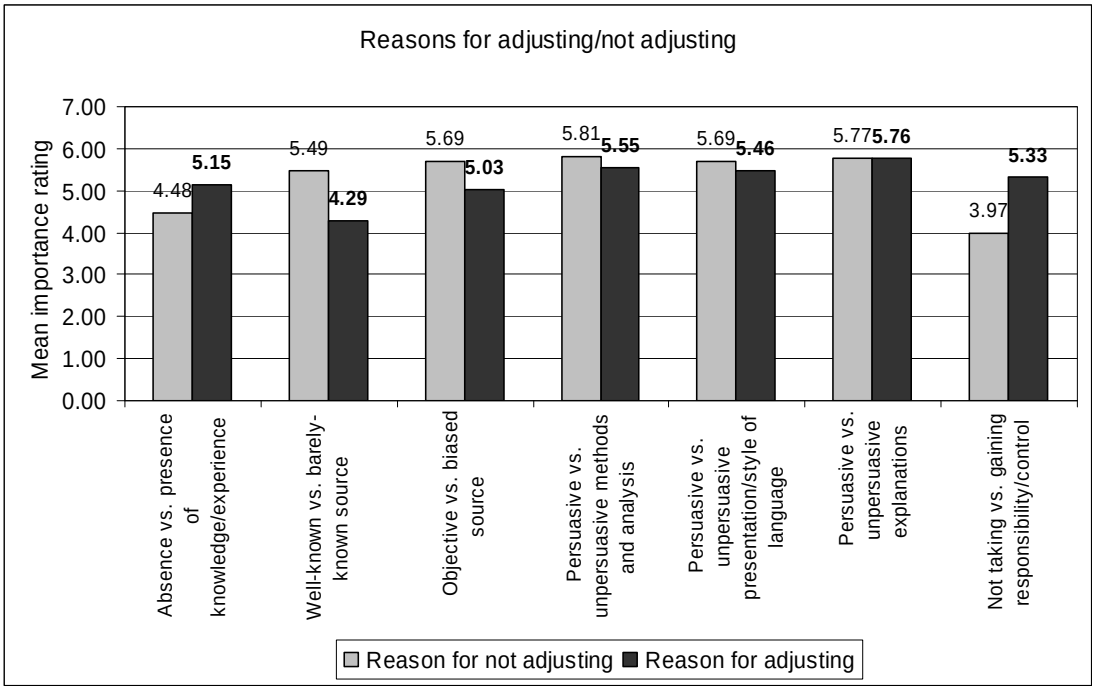


Figure 2. The relative importance of reasons for making an adjustment/not making an adjustment

The presence of and desire to integrate one's knowledge and experience was significantly more important as a reason for making an adjustment when compared against their absence as a reason for acceptance (t_{122}=-2.40, p=0.018). Likewise, making an adjustment in order to take responsibility for a forecast was rated to be a more essential motive than not making one in order not to be held responsible (t_{122}=-5.03, p<0.0001). The practitioners valued a reliable and objective source as a discouraging factor against introducing an adjustment more than they valued an unreliable and biased source as encouragement for adjustment (t_{122}=6.57, p<0.0001 for a well-known vs. barely-known source, t_{122}=3.75, p<0.0001 for an objective vs. biased source). However, these differences in ratings were not paralleled for the persuasiveness of methods, analysis, presentation, style of language and accompanying explanations. All these dimensions were rated to have statistically similar importance for encouraging adjustments when they were lacking, and for discouraging adjustments when they were adequately present.

Even though these analyses provided an overall picture for the whole group of practitioners, they did not provide any detailed information in terms of differences between groups that chose to apply adjustment in diverse frequencies. To further explore whether there are distinctions regarding reasons and motivations between the practitioners who rarely applied adjustments and those who were frequent adjusters, the data were split with respect to the question on frequency of adjustment. The participants who provided ratings of 1 and 2 (out of 7) formed the group of rare adjusters, while the participants who provided ratings

of 6 and 7 (out of 7) formed the group of frequent adjusters. The frequency and importance ratings on the dimensions of adjusting/not adjusting behaviour for these groups are provided in Figure 3.

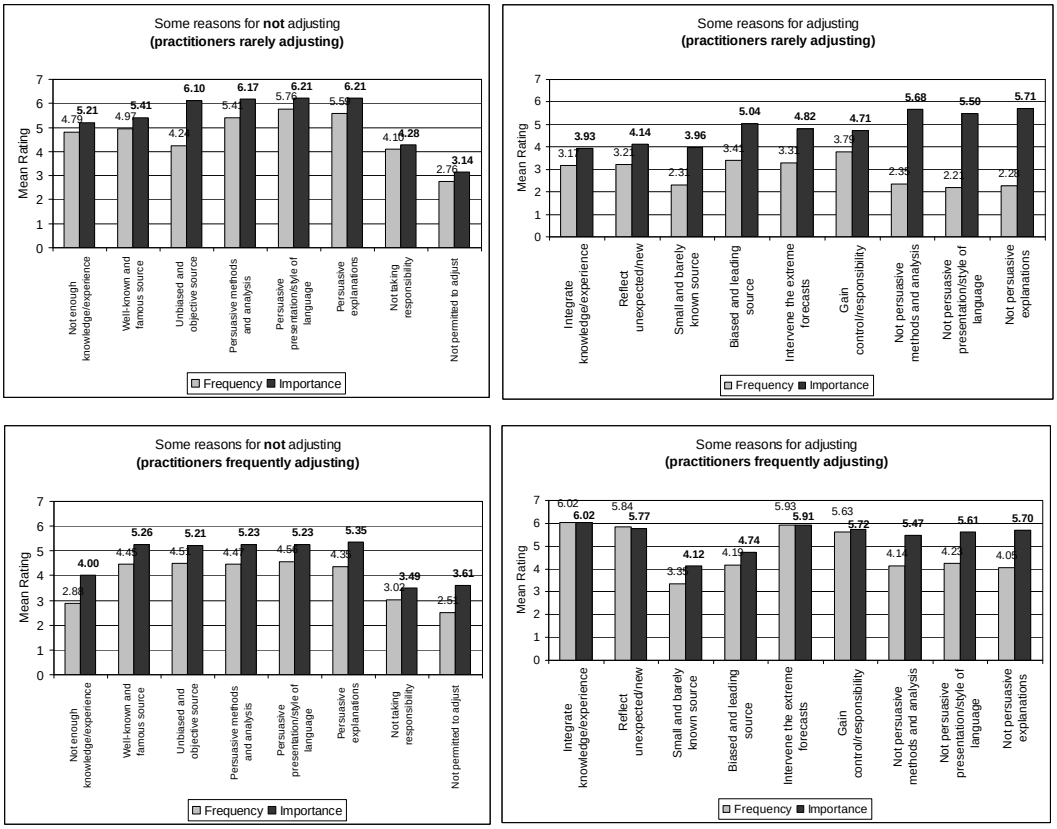


Figure 3. The frequency and importance ratings (grouped with respect to the frequency of making adjustments)

As may be observed from the figure, practitioners who identified themselves as rare adjusters indeed provided relatively high-frequency ratings for the reasons for accepting forecasts without adjustments when compared with the frequent adjusters. Only source-related dimensions (i.e., a well-known and famous source, or an unbiased and objective source) and lack of authority to make an adjustment received statistically similar frequency ratings between the two groups

in terms of not making a revision. The reverse situation occurred for frequent adjusters, where the dimensions of adjusting behaviour were reported to occur relatively more often than those given by the rare adjusters. On only one dimension (the source being biased and leading), was there no statistical difference between the frequency ratings provided by the two groups on making a revision.

There were some parallel patterns between the importance ratings provided by rare adjusters and frequent adjusters as well. Regarding the reasons for accepting a forecast the way it is, frequently adjusting practitioners attributed statistically similar importance to some of the reasons when compared with rare adjusters (i.e., there were no statistical differences between the importance ratings given to the corresponding dimensions by the two groups). These reasons were having a well-known and famous forecasting source, having no desire to take responsibility for the forecast, and lacking permission to apply adjustments. When the acquired forecasts had suitable venues for these conditions to take place, statistically speaking, rare and frequent adjusters had equally important reasons for not making adjustments.

Similar to this case, concerning the reasons for making judgmental revisions, rare and frequent adjusters gave statistically similar importance to some of the reasons. These reasons were related with the source characteristics and the absence of persuasive cues. When the acquired forecasts had an unreliable, untrusted source and the methods, analysis, presentation, style of language and explanations were not persuasive enough, that would constitute an equally

essential motive to introduce an adjustment among both rare and frequent adjusters.

One final point to emphasize in this section concerns the presence of further adjustments conducted on already-adjusted forecasts. Of all the practitioners participating in the survey, 30.7% of them reported that there is another person checking the forecasts after they adjust them. Out of this group, 80.0% responded that the colleagues inspecting their predictions introduce further adjustments on an occasional-to-regular basis. Only 20.0% of them mentioned that the introduction of further adjustments was a rare event. Table 5 provides the details.

Table 5. The presence of another person inspecting the adjusted forecasts

	Rarely	Sometimes	Frequently
Does this person introduce further adjustments?	20.0% (8)	62.5% (25)	17.5% (7)
demand explanations for your adjustments?	17.5% (7)	30.0% (12)	52.5% (21)
have access to the original/unadjusted forecasts?	10.0% (4)	40.0% (16)	50.0% (20)

As can be observed from the table, the persons working on the already-adjusted forecasts frequently expect explanations from the practitioners about the adjustments they have conducted, and they usually have access to the original/unadjusted forecasts. Regarding the issue of benefits gained from making further adjustments, practitioners generally agreed that this process leads to more accurate and persuasive predictions. This improvement in accuracy was believed to occur more than occasionally by 95% of the practitioners.

4.2.3. Effects of Receiving Forecasts from Multiple/Single Sources and Presence of Feedback

One of the preliminary questions that can be asked about the similarities and differences between practitioners acquiring forecasts from a single source or multiple sources is on the issue of how often they apply revisions. That is, whether practitioners receiving forecasts from multiple sources employ as frequent adjustments as those receiving them from a single source. The mean rating of adjustment frequency for the multiple-source group was 4.18, and for the single-source group it was 4.68. This difference was statistically non-significant, indicating that there were no distinctions in adjustment frequencies between the two groups. The practitioners' employment of revisions seems to be independent from whether they receive forecasts from multiple sources or a single one.

Another difference between these two groups of practitioners was related with the content and comprehensiveness of externally acquired forecasts. Concerning this issue, the presence of alternative forecasts, scenarios and explanations seemed to differ between the two groups. As may be observed from Table 6, the practitioners acquiring forecasts from multiple sources seemed to receive scenarios, alternative forecasts and explanations more frequently. It may be the case that these forecast users might be adopting a comparative approach to the multiple predictions they receive. They may demand more information and knowledge so that they can have a better understanding of what these predictions were all about, as well as better comprehension of the similarities and differences. The multiple-forecast-receiving users also engage in a combination task with

these predictions. By having better understanding and insight, the forecast users might believe that they have a better chance of making more efficient and accurate combinations.

Table 6. Frequency ratings of participants on their acquisition of scenarios/alternative forecasts and explanations with their forecasts.

	Frequency of scenarios/alternative forecasts	Frequency of explanations
Forecasts acquired from multiple sources.	4.73	5.29
Forecasts acquired from a single source.	3.84	4.10
Difference significant?	t_{41}= 2.24 p=0.03	t_{37}= 2.87 p=0.007

In terms of differences in forecast expectations and quality perceptions, there are some supporting findings. It seems that practitioners receiving forecasts from multiple sources have higher expectations about them arriving on time and being based on solid foundations and assumptions. The multiple-forecast group gave a rating of 6.56 for the timeliness aspect and 6.58 for the justifiable basis aspect, while the single-forecast group provided ratings of 6.10 and 6.00 to the corresponding dimensions. The differences between both ratings were statistically significant (t_{43}= 2.06 p=0.046 and t_{37}= 2.09 p=0.043, respectively). As was also discussed regarding the comprehensiveness issue, it may be the case that practitioners receiving multiple forecasts may be more involved with the forecasting process, taking the task more seriously, thus resulting in them being more critical and stricter in regard to the forecasts they acquire. A similar finding was also observed for quality perceptions. The multiple-forecast group provided 6.52 for timeliness and 6.42 for justifiable basis, while the single-forecast group

gave a rating 5.68 for both of the aspects, with the differences being statistically significant (t_{44}= 3.44, p=0.001 and t_{38}= 2.54, p=0.015, respectively).

Distinctions between the two groups were also noticeable in the reasons for refraining from making adjustments. One major difference existed in the ratings given to the importance of having no authority to employ adjustments. Practitioners receiving forecasts from a single source gave a rating of 4.45, which was significantly higher than 3.14, the score provided by practitioners acquiring their forecasts from multiple sources (t_{47}= -2.72, p=0.009). It is possible that for the single-source case, the practitioners have to be content with what they have since there is a dearth of information gathered in the form of other forecasts and no way to make comparisons. Thus, they may have less freedom to engage with an acquired forecast and are forced to accept it more than the other group. Another difference was observable in the persuasiveness of accompanying explanations. The multiple-source group attributed more importance to the presence of persuasive explanations as a reason for not making an adjustment. They provided a rating of 5.92, while the single-source group provided a rating of 5.32, and this difference was marginally significant (t_{52}= 1.99, p=0.052).

The persuasiveness of explanations was also an issue for the reasons of introducing judgmental adjustments. Paralleling the previous case, the absence of persuasive explanations was seen to be a more essential reason for adjusting a prediction by practitioners acquiring forecasts from multiple sources. They provided a rating of 5.97, which was significantly higher than 5.13, the rating given by practitioners receiving forecasts from a single source (t_{42}= 2.19,

p=0.034). At the same time, forecast users in the multiple-source group rated the lack of persuasive methods and analysis to be more imperative than those in the single-source group. Their rating was 5.75 against 4.93, and the difference was statistically significant (t_{45}= 2.26, p=0.028).

An additional point of inquiry about the two groups might be whether there exists similar feedback using behaviour. Practitioners receiving forecasts from multiple sources have to deal with and evaluate many forecasts at the same time. Without any sort of feedback mechanism, this task can prove to be quite unmanageable and burdensome given the accumulation of these forecasts over a period of time. In this respect, the use of systematic or personal feedback systems can be expected more often for practitioners acquiring forecasts from multiple sources than for those receiving them from a single source. Data obtained from the survey were indeed supportive of this argument. Significant proportion tests have shown that feedback applications were more common for practitioners obtaining forecasts from multiple sources. Approximately 41% of the first group claimed to utilize feedback mechanisms, while this percentage was around half that for the second group. Table 7 shows these results.

Table 7. The presence of feedback and the source of acquired forecasts. The numbers in parenthesis indicate the actual number of responses in the category.

	Systematic/personal feedback present	No feedback mechanism
Forecasts acquired from multiple sources.	40.9% (38)	59.1% (55)
Forecasts acquired from a single source.	19.4% (6)	80.6% (25)
Difference significant?	z = 2.46, p = 0.014	z = -2.46, p = 0.014

As shown by the previous analysis, the presence of systematic or personal feedback seems to vary among forecast users with different approaches to financial forecasting. It may be the case that the existence of feedback also fluctuates with the frequency of adjustments introduced. Practitioners who adjust their forecasts more frequently are much more likely to be the ones employing feedback systems more often. This can be expected, since making more adjustments should also be accompanied by a better means of controlling and checking the resulting accuracies in order to achieve a real improvement in the process. In order to shed more light on this issue, the distribution of data on adjustment frequency was further investigated. Figure 4 provides this distribution with respect to the presence of a feedback mechanism.

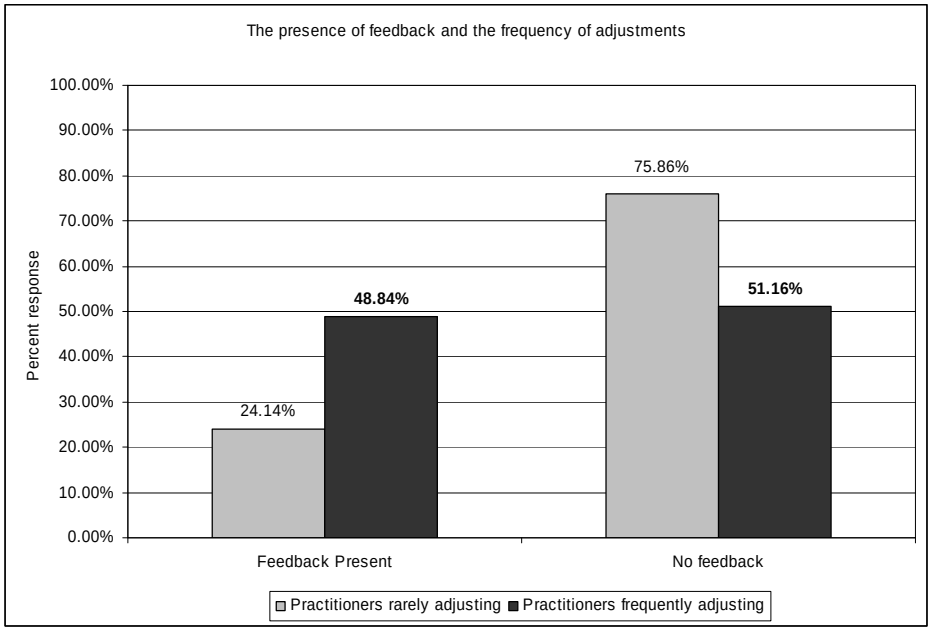


Figure 4. The presence of feedback and the frequency of adjustments

Proportion tests have shown that the presence of feedback differed significantly when the practitioners were frequent rather than rare adjusters (z = -

2.24, p = 0.025). The ratio of participants employing systematic/personal feedback rose from 24.1% to 48.8% among the group making adjustments frequently when compared against the practitioners who rarely applied adjustments.

The distinction in the adjustment frequencies owing to the presence or absence of feedback also found some reflections in the reasons and motivations behind adjusting/not adjusting behaviour. Table 8 provides these reasons, which vary significantly between the two groups.

Table 8. The presence of feedback and the importance ratings provided for the reasons behind adjusting/not adjusting behaviour.

Importance of Reasons for <u>not</u> Adjusting

	Well-known and famous source	Unbiased and objective source
Systematic/personal feedback present	4.95	5.11
No feedback mechanism	5.79	6.01
Diff. significant?	t_{63}=-2.63, p=0.011	t_{65}=-2.83, p=0.006

Importance of Reasons for Adjusting

	Unpersuasive methods and analysis	Unpersuasive presentation/style of language	Unpersuasive explanations
Systematic/personal feedback present	5.09	4.72	5.23
No feedback mechanism	5.80	5.86	6.05
Diff. significant?	t_{74}=-2.15, p=0.035	t_{75}=-3.23, p=0.002	t_{68}=-2.42, p=0.018

For practitioners not utilizing any sort of feedback, the importance attributed to the source seems to be greater than for practitioners who utilize feedback at either a systematic or personal level. This might be expected, since the no-feedback group generally has no way of checking the accuracy and reliability of their adjusted forecasts, given that there is no tool available for the purpose. For

this reason, a trustworthy and high-quality source would mean a lot more to them and would restrict their adjustments, due to their belief in the accuracy of forecasts acquired from that source. On the other hand, the group utilizing feedback can compensate to some extent for the existence of errors in the forecasts, possibly decreasing their dependence on a good and reliable supplier.

In terms of reasons for making an adjustment, the primary distinctions between the two groups concerned the absence of persuasive cues in the acquired forecasts. Any lack in the persuasiveness of methods, analysis, presentation, style of language and accompanying explanations was a bigger concern for practitioners having no feedback function. They rated these reasons as more crucial for employing their adjustments than the other group. Being more in control of the forecasting process, the feedback group seems to be less affected by the lack of some of those persuasive cues.

4.2.4. Effects of Forecast-Using Experience

The question asking about the period of forecast-using experience revealed that there were distinct groups with respect to the levels of forecasting familiarity: novice practitioners (those with less than three years of experience), familiar practitioners (those with four to nine years of experience) and experienced practitioners (those with more than nine years of experience). It can be expected that there are some differences and similarities among these groups in terms of the various aspects asked in the questionnaire.

The first aspect is whether there exists a distinction in frequency of adjustments when the practitioners' experience in forecasting practice varies. The mean frequency of making an adjustment was reported to be 4.38 among novice practitioners, 4.14 among familiar practitioners and experienced practitioners provided a score of 4.42. Furthermore, the expectations of the groups from the adjustments they conducted displayed a similar portrait. The novice group provided 4.88, the familiar group provided 4.90 and the experienced group provided 4.93 to the inquiry asking how often they believed making an adjustment improved accuracy. On the question exploring expected improvement in persuasion, novice practitioners provided 4.88, familiar practitioners provided 4.86, and finally, the experienced group gave a rating of 4.98. One-way ANOVAs showed that these ratings were not significantly different across the three experience groups. Based on this information, it can be deduced that the introduction of judgmental adjustments and the benefits expected from applying them were not dependent upon the years of familiarity and experience with forecasts.

The second aspect is the presence of another person checking the forecasts after they were adjusted by the practitioners. It may be argued that the presence of such a person might be directly related with the experience and position of the particular practitioner conducting the adjustments. On the former dimension, the task of forecast revisions is delegated to practitioners only if they are deemed worthy to be entrusted with the task. For those practitioners, this will result in the accumulation of forecasting experience over the years. Being experienced in such

a way will also lead to a situation where there are fewer people above them checking their work. This is because there is confidence in what they have been doing and/or they have been promoted to a position where they only report to a select few. Figure 5 provides the survey data gathered on this issue.

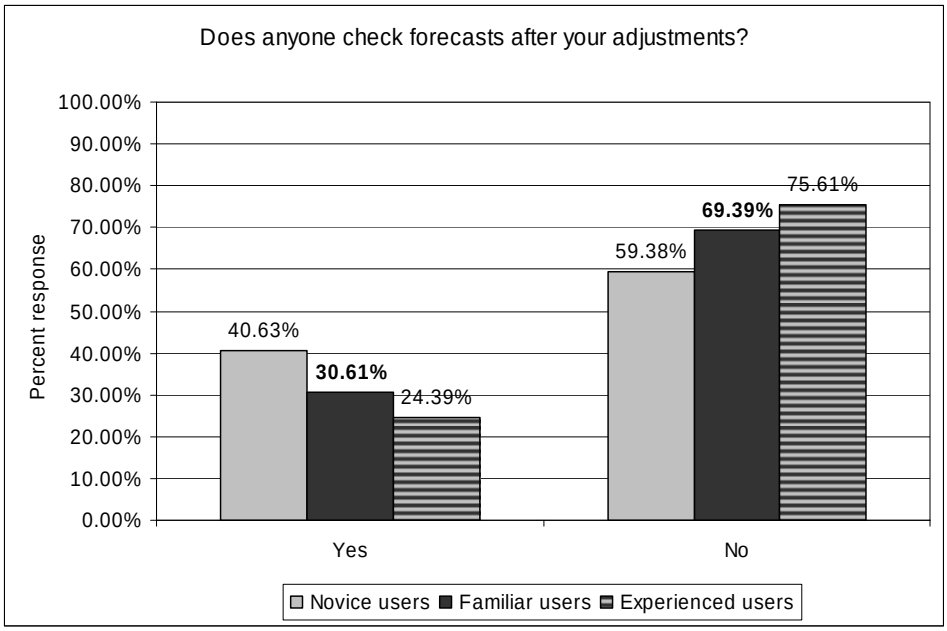


Figure 5. The presence of a person checking the adjusted forecasts and the forecast using experience

As is evident from the figure, the presence of a person checking the adjusted forecasts drops from 40.6% to 24.4% when the adjustments are conducted by experienced rather than novice forecast users. However, this effect does not seem to be strong, given that the difference between the proportions cannot reach statistical significance at the 5% level.

The third aspect inquired into differences among the three groups in terms of expectations from forecasts and perceptions of quality. The only statistical difference that originated from the varying experience levels was found to be in

the expectations of tolerable error from acquired forecasts. The novice group rated the question 5.19, the familiar group 5.86 and the experienced group provided 6.24 to the question. One-way ANOVA indicated that this difference was significant ($F_{2,119}$ = 4.96, p=0.008). At the same time, there was no significant difference in expectations regarding the smallest amount of error. This finding suggests that even though all the practitioners expect to have as accurate forecasts as possible, through experience and familiarity they may gain a more realistic perspective into the presence and nature of forecast errors. This may lead to a rise in the anticipation that there will always be error present in the predictions and, as long as they are within tolerable limits, they are not an obstacle for the forecasting function to proceed properly. Therefore, the participants might have provided higher ratings on this aspect.

This result was not paralleled for perceptions about a high-quality forecast. There were no significant differences among the groups for the existence of either a tolerable amount of error or the smallest amount of error. It may be the case that more experienced practitioners have a higher expectation of receiving forecasts with a tolerable amount of error when compared against less experienced practitioners, but this was not something they would anticipate from a high-quality forecast, just like the other groups. Indeed, experienced forecast users significantly lowered their rating of tolerable error as an aspect of forecast expectation from 6.24 to 5.68, where the second number is the rating it received as an aspect of a high-quality forecast (t_{40}= 2.34, p=0.024).

The final aspect concerns the existence of differences regarding various reasons and motivations behind the employment of adjustments and acceptance without adjusting. For all these reasons and motivations, there seem to be no significant distinctions originating from the various experience levels. Practitioners across all experience groups seemed to provide statistically similar importance ratings for the whole set of dimensions. This suggests that making judgmental adjustments may indeed be independent of the level of forecasting experience and familiarity, supporting the findings on adjustment frequencies.

4.2.5. Effects of Practitioner's Position

Among the participants in the survey, there existed four distinct groups with respect to their position in the company; the partners/owners of the company, CEOs/executive board members, managers and staff members. In terms of the various dimensions asked in the questionnaire, the presence of some differences and similarities can be expected among these groups of practitioners.

As in the previous section, the first dimension is whether there exists a distinction in the frequency of adjustments. The mean frequency of making an adjustment was reported to be 4.71 among partners/owners, 4.57 among CEOs/executive board members, 3.69 among managers and 3.60 among the practitioners working as part of the staff. Even though practitioners higher in power and hierarchy (i.e., partners/owners and CEOs/executive board members) seem to apply judgmental adjustments more frequently, this difference was only

able to reach statistical significance at the 7% level, indicating a weak relationship ($F_{3,120}$ = 2.41, p=0.07). Furthermore, the groups' expectations from the adjustments they have conducted appear to be similar to those seen in the frequency case. The partners/owners group provided 5.03, the CEOs/executive board members group provided 5.19, the managers asserted 4.77 and the staff members provided 4.20 to the question that inquired how often they believed making an adjustment improved accuracy. On the issue of expected improvement in persuasion, the partner/owner practitioners provided 5.16, the CEO/executive board member practitioners provided 5.05, the practitioners employed as managers gave 4.71 and the practitioners working as staff members gave a rating of 4.00. One-way ANOVAs showed that these ratings were not significantly different from one another among the four groups. These findings led to the deduction that introduction of judgmental adjustments and the benefits expected from applying them were not very dependent on the practitioner's position in the company. There was only a weak direction, that the practitioners with higher status applied more frequent adjustments with better expectations.

The second dimension is whether there existed another person checking the forecasts after adjustment by the practitioners. As was argued before, the presence of such a person might be directly related with the experience and position of the particular practitioner conducting the adjustments. On the latter dimension, as the position of a practitioner rises in hierarchy, there will be fewer people left above them to check their adjusted forecasts. These individuals are generally engaged in coordination and administration activity so that they can

access various sets of information coming from different persons or departments. Thus, the adjustments conducted by these practitioners can be regarded to be more valuable than those conducted by lower-hierarchy people who only have access to limited views and information. This fact is reflected by the presence of fewer people checking their forecasts after they have adjusted. Figure 6 summarizes the data gathered related with this issue.

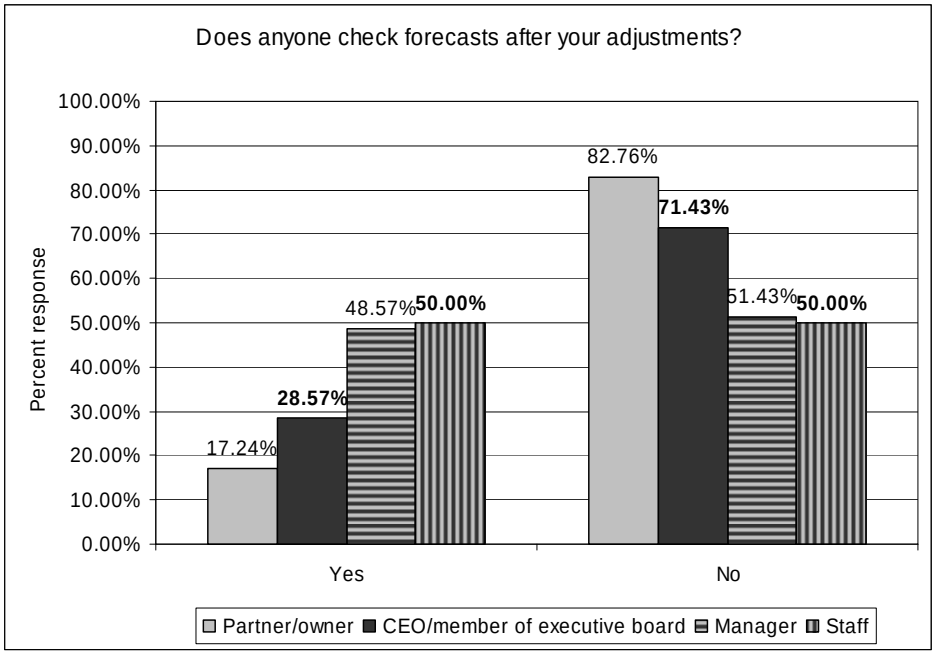


Figure 6. The presence of a person checking the adjusted forecasts and the position of the practitioner.

As is observable from the figure, the number of practitioners reporting the presence of another person checking the forecasts after their adjustments rises dramatically when the power and hierarchy of the person diminishes. The partners/owners responded "yes" with a mean frequency of 17.2%, which is not significantly different than the response of CEOs/executive board members. However, this proportion is significantly lower than the proportion of managers (z

= -3.20, p = 0.001) and that of staff members (z = -1.98, p = 0.048). These findings are completely in line with the arguments in the previous paragraph.

The third dimension concerns the existence of differences in various aspects of forecast expectations and quality perceptions. For this issue, the only significant distinction present was related to cost anticipations. Partners/owners provided a mean rating of 5.85 to the expectation that external forecasts should be obtained with a reasonable cost. CEOs/executive board members gave 4.71, managers rated 5.14 and practitioners working as part of staff chose 5.50. The difference among these ratings is statistically significant ($F_{3,120}$ = 2.79, p=0.043). This difference might stem from a possible discrepancy between the person actually authorizing the payment for the acquired forecasts and the person using them. On the other hand, such a difference was not observed in terms of quality perceptions. Practitioners in the four groups of positions rated the importance of high-quality forecasts having reasonable costs very similarly. On this issue, the greatest change was seen in the views of the partners/owners. They provided a rating of 5.35 for reasonable cost being an important attribute of forecast quality. This was significantly lower than the rating they provided for their expectation (t_{57}= 2.31, p=0.024). Practitioners in all other positions gave statistically similar ratings for both cases. It seems that from the perspective of partners/owners, the forecasts were expected to be priced reasonably; however, for forecasts to be considered high-quality this was not so important.

The final aspect inquired into the distinctions among the four groups in terms of reasons and motivations behind employment of adjustments and

acceptance without adjusting. For the reasons of refraining from making an adjustment, no significant differences were noticeable among the four groups. Practitioners in all positions provided very similar ratings to the occasions where they accepted the forecasts without revisions. On the other hand, there were some dissimilarities present for the cases where they introduced their adjustments. These dissimilarities are depicted in Table 9.

Table 9. The practitioner's position and the importance ratings provided for the reasons for making an adjustment.

	Integrate knowledge/experience	Reflect unexpected/new	Gain control/responsibility
Partner/owner	5.29	5.38	5.38
CEO/ executive board member	5.81	6.05	6.10
Manager	4.82	5.03	5.18
Staff member	4.00	4.20	3.90
Diff. significant?	$F_{3,119}$=2.93, p=0.037	$F_{3,119}$=3.09, p=0.030	$F_{3,119}$=4.59, p=0.004

As is evident from the table, the differences in the ratings were mainly related with practitioners adding something from themselves via experience, knowledge or some extra information they possessed. This also involved forming a sense of ownership of the forecasts through contribution after adjustments were made. It is not difficult to anticipate that practitioners with higher power and hierarchy are more likely to engage in such adjusting behaviour. Indeed, partners/owners and CEOs/executive board members all provided higher importance ratings across all three dimensions when compared with practitioners in other positions, thus providing supporting evidence.

CHAPTER 5

EXPERIMENTAL STUDIES

As was briefly mentioned before, the second part of the research involved three experimental studies for the third, fourth and fifth research questions. The first study in this part of the research manipulated the structural characteristics of the explanations (i.e., their length and conveyed confidence) in order to determine their effect on the judgmental adjustment of forecasts. The second study investigated the basic properties and effects of nested adjustments. The last study involved the combination of nested adjustments with explanations, and tried to explore the interaction of the presence of unadjusted/initial forecasts with the presence of explanations.

The subjects participating in the studies were all undergraduate students in the Faculty of Business Administration at Bilkent University. They were chosen from students who had received basic training in statistics, probability and forecasting theory. Small incentives (especially in the form of performance-based bonus points that would be included in their final grades) were provided to the subjects to encourage their participation, interest and commitment.

All of the studies involved stock-price forecasting tasks. Stock-price forecasting is one of the most well-known financial forecast types. Due to its coverage in newspapers, the media and lectures, every business student is knowledgeable to some extent about the ISE-100 (Istanbul Stock Exchange) index and stock prices. Owing to this property, they formed an ideal medium for the studies.

5.1. Performance Measures Used

Performance measures employed in all of the studies may be summarized under three categories: the amount of adjustments, the size of adjustments, and the accuracy of the adjusted forecasts. Table 10 depicts the various measures used under each of these categories for both the point and interval forecasts.

Table 10. Performance measures

	Amount of Adjustment	Size of Adjustment	Accuracy
Point Forecasts	% of point forecasts adjusted	APAP (absolute % adjustment in point forecasts)	MAPE
Interval Forecasts	% of interval forecasts adjusted	APAI (absolute % adjustment in interval forecasts)	Hit rate

Amount of Adjustment: The percentages of point and interval forecasts that were modified were used as indexes of the amount of adjustment for the two respective formats. Any adjustment on any of the bounds for a particular interval

forecast was counted as a single adjustment for that interval. That is, a prediction interval was considered adjusted if at least one of its bounds were modified.

Size of Adjustment: To capture the size of adjustments, absolute percentage adjustment in point forecasts (APAP) was used for point forecasts, while the absolute percentage adjustment in interval width (APAI) score was utilized for the interval forecasts. These performance measures were calculated as follows:

$$APAP = \frac{\left| adjusted \quad point \quad forecast \; - \; provided \quad point \quad forecast \right|}{provided \quad point \quad forecast} \times 100$$

$$APAI = \frac{\left| adjusted \quad width \; - \; provided \quad width \right|}{provided \quad width} \times 100$$

If no adjustments were made to the provided forecasts, these ratios automatically received the lowest possible score (0%). The further away the adjusted predictions were from the provided values, the higher were the APAP and APAI scores.

Accuracy: Accuracy of the adjusted forecasts was calculated using the mean absolute percentage error (MAPE) for point forecasts, while the "hit rate" (*i.e.*, percentage of occasions when the actual value is contained in the interval) was used for interval forecasting performance.

5.2. Study 1: Effects of Structural Characteristics of Explanations

5.2.1. Methodology and Design

Overview

This study explored the structural characteristics of explanations (i.e., their length and conveyed confidence) in order to determine their effects on the judgmental adjustment of forecasts.

Explanation length is one of the important factors that may influence the effectiveness of explanations. Explanations having the same content or carrying the same contextual information may be either short or long. In other words, the same information can be expressed in short and brief sentences that convey only the intended meaning, or in a lengthy manner with redundant sentences and obvious details. The short sentences may be a preferable way to express the planned information, since it is quicker to read and process. At the same time, reading short explanations will not be boring or time-consuming, leading decision makers to show a lower tendency to skim or skip them. The long explanation style may be disadvantageous in this sense; however, long explanations also carry the advantage of clarity and better comprehensibility. The conveyed meaning can be better understood if explanations are provided in a long and detailed way. The two sides of this argument evidently indicate that the exact effect of explanation length has yet to be determined.

Regardless of its length, the style of the language used in explanations is also very important. The language style is directly related to the amount of confidence conveyed to the forecast users. The sentences in an explanation may be formed from words that express strength, assurance and precision; this type of language may convey to the decision makers that the forecast providers are confident in the forecasts they have generated, resulting in their explanations' being more persuasive. At the same time, this style may be perceived as a sign of the expertise of the providers, which will also lead to an improvement in the acceptance of the provided forecasts (Sniezek and Swol, 2001).

Alternatively, explanations may be worded to convey uncertainty and doubt. The words chosen may reflect vagueness and ambiguity in the predictions. This style of language is more appropriate to the nature of forecasting, which is, in fact, all about uncertainty. An explanation given in weak confidence may be recognized as natural, normal and expected. If this is the case, strength and confidence communicated in explanations may be perceived as unnatural, excessively assertive and unjustifiable. In this sense, a weakly confident style may have more persuasive power in facilitating the acceptance of provided forecasts.

However, counter-arguments may also be made. Explanations that convey a weak sense of confidence may lead decision makers to perceive inability and inexperience in the providers. In such a situation, weak confidence may have a detrimental effect on the acceptance of a given forecast by the forecast users. As in the case of explanation length, then, the exact effects of language-style differences are as yet undetermined.

Participants

This study was conducted in Spring 2002. A total of 116 business students enrolled at Bilkent University participated. The students were all third-year students and they were taking courses in Business Forecasting. For their participation, they were offered an incentive in the form of extra course credits, the amount of which was dependent on their performance.

Procedure

The study was conducted in a paper-and-pencil style. Each participant was presented with a booklet including the instructions, the forms, and a wrap-up questionnaire. The form consisted of 30 time-series followed by one-period-ahead point and 95% interval forecasts for each series. There were also accompanying explanations written specifically for a particular series and its corresponding forecasts.

The participants were first requested to examine the provided time-series and its point/interval forecasts and to carefully read and consider the corresponding explanations. They were then asked whether they were satisfied with the point and interval forecasts. If they were not satisfied, they were asked to adjust the given forecasts based on their judgment (A sample form is provided in Appendix B). All the participants received the same set of time-series and point/interval forecasts; however, the explanations they received were different due to experimental manipulation.

To investigate the structural characteristics of explanations, there were two independent variables. The first independent variable was the length of the explanations, with the levels "short" and "long". The level of conveyed confidence was the second independent variable, with "strong" and "weak" confidence as its two levels. Thus, a 2x2 factorial design (please see Cochran and Cox (1957) and Montgomery (2001) for a discussion of this design) led to four groups, with the following number of participants in each group:

1st Group: *Short explanation, weak confidence (n=29)*. For this group, the explanations were 1 or 1.5 lines long, and contained the same information as the long explanations. To manipulate the conveyed confidence, words such as "probably, possibly", etc. were utilized in the explanation sentences to express weak confidence. Sample explanations for the four groups are provided in Appendix B.

2nd Group: *Long explanation, weak confidence (n=30).* For this group, the explanations were 4 or 5 lines long, and contained the same information as the short explanations. To manipulate the conveyed confidence, words such as "probably, possibly", etc. were utilized in the explanation sentences to express weak confidence.

3rd Group: *Short explanation, strong confidence (n=29.)* The explanations were 1 or 1.5 lines long, and contained the same information as the long explanations. To manipulate the conveyed confidence, the weakly confident words in the explanation sentences were replaced with words such as "definitely, surely", etc. to express strong confidence.

***4th Group**: Long explanation, strong confidence (n=28)* For this group, the explanations were 4 or 5 lines long, and contained the same information as the short explanations. To manipulate the conveyed confidence, the weakly confident words in the explanation sentences were replaced with words such as "definitely, surely", etc. to express strong confidence.

The final part of the booklet included a short wrap-up questionnaire to gain a better understanding and interpretation of the participants' perceptions. The questionnaires were prepared separately for each of the four groups.

All participants were required to do the same task except one group. This group was composed of 47 participants who were requested to complete an additional task by providing ratings of information value for every explanation. The forms they received contained an extra 7-point scale to enable them to rate the information value (i.e., the perceived usefulness of that specific explanation). The 7-point scale ranged from "very misleading" to "very helpful", with the mid-scale degree being "no value".

Generation of the Time-series and Forecasts

30 time-series were utilized in the study. Even though these series were framed as being real stock prices, they were in fact artificially constructed in order to have control over the levels of trend and variability. There were three levels of trend and two levels of variability introduced to the series, generating six different kinds of time-series. With five samples of each type, 30 series were constructed. These

series were distributed to each participant in randomized sequence to prevent any ordering effects.

The trends involved were 2% increasing, 2% decreasing and flat trend. The trend level of 2% was calculated by fitting an ordinary least-squares regression line onto a previous period's ISE-100 index. To produce the variability levels, normally distributed noise with zero mean was used. For the low variable case, the noise had a standard deviation of 5%, and for the high variable case the standard deviation was 15%. These variability levels are plausible for the ISE-100, and they have been utilized in previous work (O'Connor et al., 1993). The trend and variability components were additively aggregated to achieve the final series. One additional point was also generated to be used as the actual value for accuracy measures.

The one-period-ahead point forecasts were constructed with Holt's exponential smoothing method, using error-minimizing parameters. Based on the error variance of the point estimates, 95% interval forecasts were obtained. The exact nature of the forecasts was not revealed to the participants.

5.2.2. Results

The findings are outlined under three categories involving performance measures on point forecasts, performance measures on interval forecasts, and the perceived information value of the explanations.

5.2.2.1. Point Forecasts

Table 11 summarizes the results of the four groups across the relevant performance measures for point predictions.

Table 11. Overall results for point forecasts (numbers in parentheses are the numbers of data points in that category)

	% of point forecasts adjusted		APAP		MAPE	
	Weak confidence	Strong confidence	Weak confidence	Strong confidence	Weak confidence	Strong confidence
Short	41.95% (365/870)	42.26% (355/840)	4.72%	4.11%	19.16%	18.75%
Long	44.71% (389/870)	41.33% (372/900)	4.25%	4.14%	18.78%	18.41%
	n.s.		n.s.		n.s.	

Across all the three performance measures used, no significant differences were found among the four groups, with no significant main or interaction effects for explanation length and the level of confidence conveyed in the explanations. This suggests that the number and size of adjustments made on the point forecasts and the resultant accuracy is not affected from the manipulations made through the explanations.

5.2.2.2. Interval Forecasts

Table 12 presents the findings of the four groups across the pertinent performance measures for interval forecasts.

Table 12. Overall results for interval forecasts (numbers in parentheses are the numbers of data points in that category)

	% of interval forecasts adjusted		APAI		Hit Rate	
	Weak confidence	Strong confidence	Weak confidence	Strong confidence	Weak confidence	Strong confidence
Short	50.69% (441/870)	57.14% (480/840)	15.85%	15.98%	79.20%	79.76%
Long	56.32% (490/870)	54.22% (488/900)	16.89%	13.68%	79.54%	82.89%
	$F_{3,3476}$=2.89; p=0.03		$F_{3,3476}$=3.14; p=0.02		n.s.	

For the number of adjustments made to the interval forecasts, ANOVA results show that main effects for the length and style (conveyed confidence) variables were not statistically important, however the interaction between them was statistically significant ($F_{1,3476}$=6.42, p=0.011). Their interaction graph is provided in Figure 7. This interaction leads to a significant difference among the four groups ($F_{3,3476}$=2.89, p=0.03).

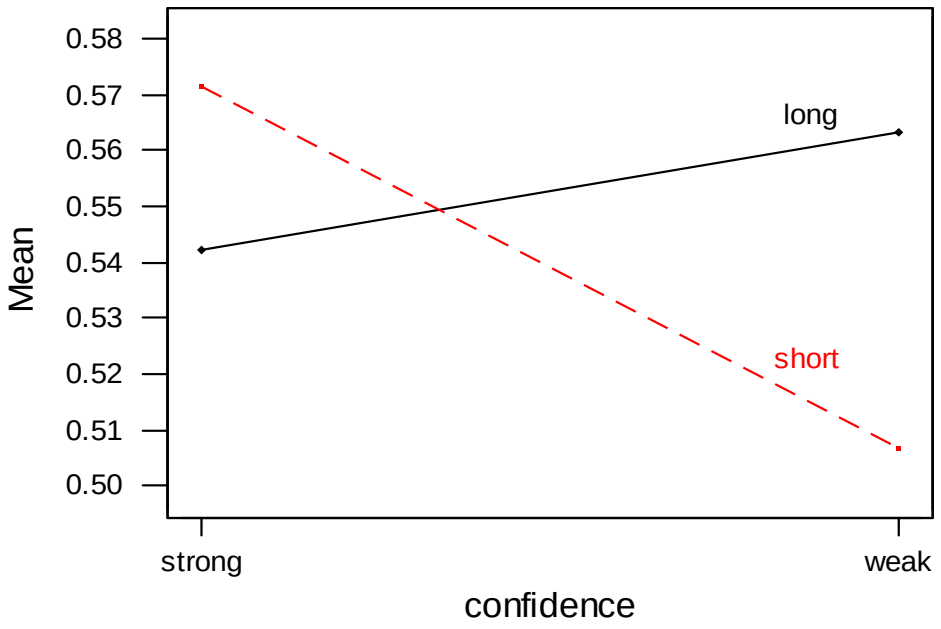


Figure 7. Interaction effect for percentage of interval forecasts adjusted

Manipulating length or style (conveyed confidence) alone seems not to have a significant influence on the number of adjustments made to the interval forecasts. However, pairwise comparisons show that long explanations combined with weak confidence appear to lead to higher percentage of forecasts adjusted then short explanations with weak confidence ($p=0.018$). Also, there are no significant differences between having a strong vs. weak confident style in case these explanations are expressed in long wording. Additionally, when an explanation is short, then the level of conveyed confidence seems to be more influential in leading to adjustments of presented forecasts. The difference between short/weakly confident and short/strongly confident groups is significant ($p=0.007$).

Resembling the case of percentage of forecasts adjusted, ANOVA results show that there was a significant main effect of conveyed confidence/style ($F_{1,3476}=3.92$, $p=0.048$) and a significant interaction effect between explanation length and conveyed confidence ($F_{1,3476}=4.64$, $p=0.031$) in terms of the size of adjustments as measured by APAI. Owing to these effects, there was a significant difference among the four groups ($F_{3,3476}=3.14$, $p=0.02$) with the least adjustment made in the long and strongly confident explanation group.

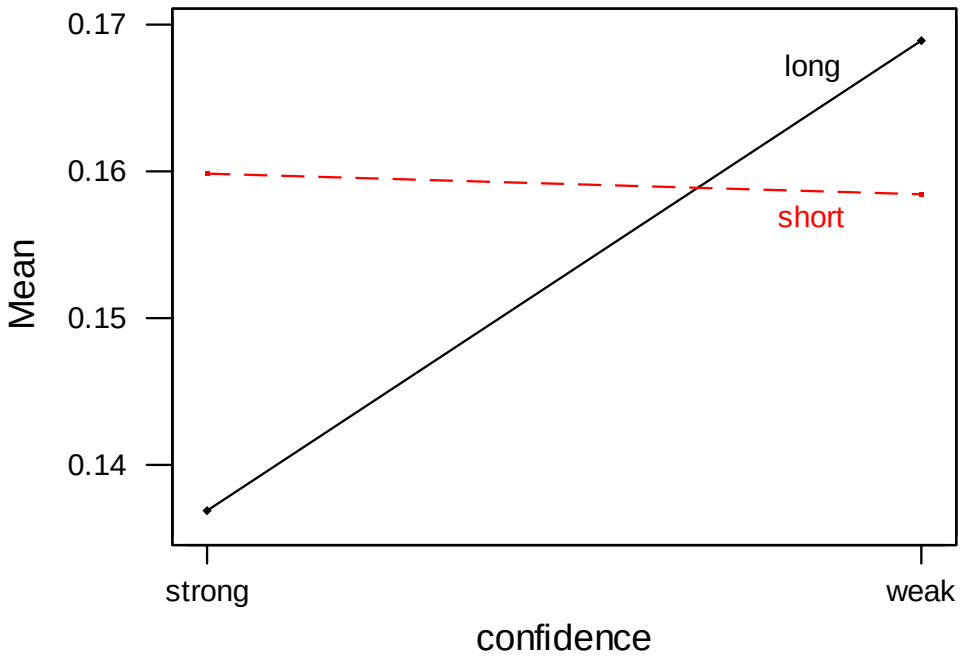


Figure 8. Interaction effect for the size of adjustments made in interval forecasts

The interaction effect depicted in Figure 8 suggests that the size of adjustments made to interval forecasts shows differences depending on the particular combination of detail and conveyed confidence in explanations. When a long explanation is accompanied by strong confidence, adjustments are much lower than all the other possible combinations. Post-hoc tests designate that all the pairwise comparisons between long/strongly confident group and the other groups are significant ($p=0.003$ with long/weakly confident group, $p=0.046$ with short/weakly confident group, $p=0.036$ with short/strongly confident group). There doesn't seem to be any significant differences due to strong vs. weak confidence when a given explanation is short.

Further insight about the adjustments made to the interval forecasts may be gained from Table 13. The participants receiving short explanations with weak wording and long explanations with strong wording were decreasing the interval widths less often than the other groups. At the same time, for these groups, percentage of intervals in which the widths remained the same was significantly higher. This may be interpreted as further support that explanations carrying this particular set of structural characteristics lead to fewer modifications applied to the provided interval forecasts.

Table 13. Percentage of intervals widened/narrowed/not changed in interval width (total number of intervals in parentheses)

	PERCENTAGE OF INTERVALS			
	widened	narrowed	no change in interval width	TOTAL
Short, weak confidence	14.14% (123)	35.52% (309)	50.34% (438)	100% (870)
Long, weak confidence	17.01% (148)	38.28% (333)	44.71% (389)	100% (870)
Short, strong confidence	14.05% (118)	41.31% (347)	44.64% (375)	100% (840)
Long, strong confidence	16.67% (150)	35.33% (318)	48.00% (432)	100% (900)
	n.s.	Chi-Sq_3=8.676, p=0.034	Chi-Sq_3=7.968, p=0.047	

It can also be observed from Table 11 and Table 12 that participants made adjustments to interval forecasts more than they made to point forecasts regardless of the group (t_{3479}=12.97, p<0.001). This may be due to the fact that there is only a single point to adjust in the point forecasts, while there are two bounds that can be adjusted for the interval forecasts. For a person not to adjust a provided point forecast, that person has to be satisfied with a single number. However, for the same time series, that person has to be satisfied with both of the provided bounds for the interval predictions in order not to introduce an adjustment. Additionally,

there is a general preference for asymmetric interval forecasts over symmetric ones, which may generate a further tendency to adjust a given symmetric interval (Lawrence and O'Connor, 2005). Hence, interval forecasts may be more likely candidates for adjustment than point forecasts. In a similar manner, the size of adjustments made to the point predictions is much less than the magnitude of adjustments made to the interval width, with the difference being statistically significant (t_{3479}=28.64, p<0.001).

In terms of the hit rates of interval forecasts, there appears to be no significant differences among the four groups. Similar to the results obtained from point forecasts, the changes in the characteristics of accompanying explanations did not create a change in the accuracy achieved. Even though the manipulations in explanations seemed to alter both the amount and size of adjustments, the induced effect was not strong enough to create a relative change in the final accuracy.

However, it is worth noting that if none of the participants made any adjustments to the provided forecasts, the resulting MAPE for the point predictions would be 16.51% and the hit rate for the interval predictions would be 86.67%. In other words, MAPE and hit rate scores for the provided forecasts are 16.51% and 86.67%, respectively. An examination of Table 11 and Table 12 shows that across all the groups, accuracy of adjusted forecasts (both point and interval) is worse than the accuracy of provided forecasts (all p<.01 for MAPE; all p<.001 for hit rate). Clearly, when the participants made adjustments to the external forecasts, they reduced the accuracy of the provided ones.

5.2.2.3. Perceived Information Value of Explanations

As an additional task, a subgroup of participants corresponding to 47 students in a given section of the forecasting course were asked to evaluate the information value of each explanation by using a 7-point scale (with 1= "very misleading"; 4= "no value"; 7="very helpful"). It was thought that such ratings may provide insights about how the participants perceived the provided explanations and in turn how these perceptions might have affected the adjustments they made.

For the analysis, the data were regrouped into a 3-point scale with 1 = "misleading" (initial ratings of 1&2 combined), 2 = "no real value" (initial ratings of 3, 4&5 combined) and 3 = "helpful" (initial ratings of 6&7 combined). The average information value attributed to the provided explanations was 2.36 across all groups, which was significantly greater than 2 = "no real value" (t_{1401}=20.91, p<0.001). Hence, supporting the participants' verbal comments at the end of the study, it may be concluded that the participants on average found the provided explanations helpful to some degree. Table 14 provides the average information value within each group.

Table 14. Average information value

	Average information value	
	Weak confidence	Strong confidence
Short	2.43	2.33
Long	2.29	2.38
	$F_{3,1398}$=3.07; p=0.03	

The ANOVA results reveal no significant main effects but a significant interaction effect in the perceived information value of explanations ($F_{1,1398}$=7.44, p=0.006), as depicted in Figure 9. The effects of this interaction can also be observed in the significant rating difference among the groups ($F_{3,1398}$=3.07, p=0.03).

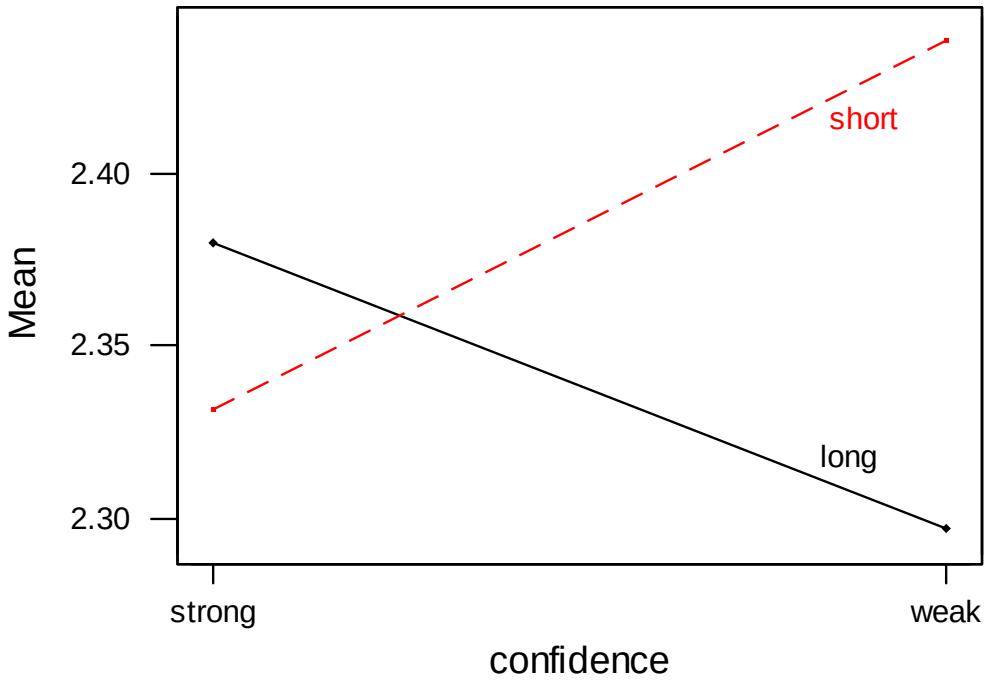


Figure 9. Interaction effect for information value of explanations

When a short explanation is given, it seems that weakly confident wording leads to much higher ratings of information value (as compared to strongly confident wording) (p=0.032). Pairwise comparisons indicate that the difference between short/weakly confident group and long/weakly confident group is also significant (p=0.005). The only group that achieved statistically similar information value rating with short/weakly confident group is long/strongly

confident group. In sum, provided explanations either have to be long/strongly confident or short/weakly confident to result in more positive perceptions.

Since it is reasonable to assume that the adjustments made on a provided forecast may be directly influenced by the relevant explanation's perceived information value, it may be expected that the measures of adjustment (i.e., percentage of point forecasts adjusted, percentage of interval forecasts adjusted, APAP, APAI) and accuracy (i.e., MAPE, hit rate) may show differences with respect to information value perceptions. Results of this analysis can be found in Table 15.

Table 15. Differences in adjustment and accuracy measures with respect to information value

Measures	Misleading	No real value	Helpful	$F_{2,1399}$	p
% of point forecasts adjusted	54.8%	41.3%	26.9%	26.49	<0.001
% of interval forecasts adjusted	63.0%	50.8%	41.1%	13.38	<0.001
APAP	8.0%	4.4%	3.1%	21.80	<0.001
APAI	18.5%	13.5%	11.7%	6.02	0.002
MAPE	26.6%	19.8%	16.0%	8.50	<0.001
Hit rate	71.8%	80.4%	84.8%	6.84	0.001

Firstly, all the adjustment scores (i.e., scores on percentage of point forecasts adjusted, percentage of interval forecasts adjusted, APAP, APAI) persistently decrease as the attributed information value of an explanation increases. Second, the accuracy of the adjusted forecasts steadily improves as the attributed information value increases (i.e., MAPE decreases and hit rate increases

at the same time). These findings confirm the previous results in that if a participant finds an explanation to be helpful in understanding the provided forecasts and time-series, he/she tends to trust and accept the provided forecasts more. In this sense, an explanation with a high perceived information value can be considered as a persuasive explanation leading to less adjustments and therefore more accuracy in the adjusted final forecasts. On the other hand, if an explanation is deemed "misleading", the provided forecasts lose some of their credibility, leading to a higher rate of adjustment and the resultant reduced accuracy.

5.3. Study 2: Effects of Adjustment Framing on Further Adjustments

5.3.1. Methodology and Design

Overview

The second study aimed to explore whether there were any adjustment differences in situations where the provided forecasts were thought to be original/unadjusted predictions as opposed to cases where the given forecasts were believed to represent already-adjusted predictions. Will there be a difference in the adjustment/acceptance behavior when a forecast user is led to believe that a given forecast has been previously adjusted? It could be that users may be influenced by the belief that the forecasts have already been worked on, had time and thought spent on them, and hence, been refined to some degree. If so, they may apply

fewer adjustments to these forecasts, indicating a higher acceptance rate. On the other hand, users may value their own judgment and opinions more than that of others (Lim and O'Connor, 1995, 1996), so it may not make any difference whether the presented predictions are thought to be original or already adjusted, leading to a lower acceptance of given forecasts.

Participants

The second study was conducted in April 2005. A total of 86 business students participated. The students were mostly second- and third-year students taking courses in Management Science and Forecasting. For their participation they were offered an incentive in the form of extra course credits, the amount of which was dependent on their performance.

Procedure

The primary focus of this study was the investigation of nested-adjustment behavior; in other words, to explore how forecast users adjust when they believe that they have received adjusted forecasts rather than unadjusted/original forecasts.

Eighteen time-series selected from the first study were utilized. It may be remembered that these series were artificially constructed to control the levels of trend and variability. Furthermore, point forecasts and 95% interval forecasts generated for the 18 time-series were already available from the first study.

As in the previous case, these series were framed as being real stock prices, with the names of stocks and time periods concealed in order to prevent potential

bias and extraneous information effects. Both the time-series and the forecasts were coded into a Visual Basic program for administration via computer. This was programmed by the researcher and named "Experimental Forecasting Support System", or EFSS for short.

To explore the framing effect, two groups were needed. These groups received identical sets of time-series and forecasts, but one of the groups was led to believe that the given forecasts were already adjusted, while the other group was told that the forecasts received were the original/unadjusted predictions, and thus served as the control group. The procedure for these groups was as follows:

1st group: This group served as the control group. Firstly, time-series plots of the stocks were provided. The participants were asked to click on a button to reveal the unadjusted/original forecasts, with a textbox saying "Click below to view the unadjusted/original forecasts". They had to click on this button to proceed. When the button was clicked on, the forecasts were shown on a graph. (Figure 10 presents a screenshot of the program for this group.) Afterward, the participants were requested to provide their modifications if there were any. They could easily perform modifications by selecting either the point forecast, the upper 95% bound or the lower 95% bound and then clicking on the graph to show the modified values. If they did not click on any of these, the forecasts remained unmodified. They could move on to the next set by clicking on a button after this stage. The program terminated when the entire 18 series had been worked on. The program automatically recorded all the

adjustments onto a diskette that had been provided to each participant before the study. A total of 43 subjects participated in this group.

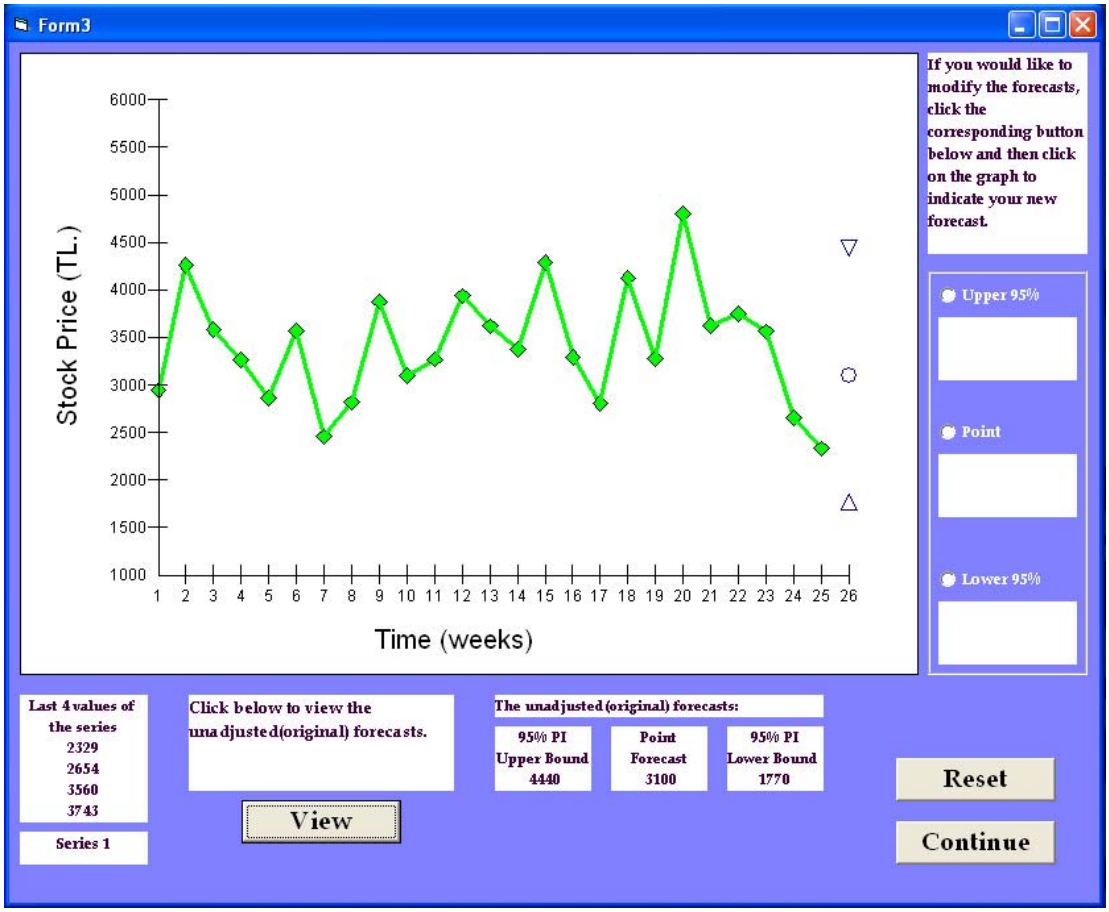


Figure 10. A screenshot from EFSS. – 1st group

2nd group: For this group, as in the previous group, time-series plots of the stocks were provided initially. At this time, participants were asked to click on a button to reveal the adjusted forecasts, with a textbox saying "Adjustments have been made to the forecasts. Click below to view the adjusted forecasts". They had to click on this button to proceed. When the button was clicked on, the forecasts (the same as in the previous group) were shown on a graph. (Figure 11 presents a screenshot of the program for this group.) Afterward, the participants were requested to

provide their modifications if there were any. They could easily perform modifications by selecting either the point forecast, the upper 95% bound or the lower 95% bound, and then clicking on the graph to show the modified values. If they did not click on any of these, the forecasts remained unmodified. They could move on to the next set by clicking on a button after this stage. The program terminated when the entire 18 series had been worked on. The program automatically recorded all the adjustments to a diskette that had been provided to each participant before the study. An equal number of subjects as in the previous group participated in this group.

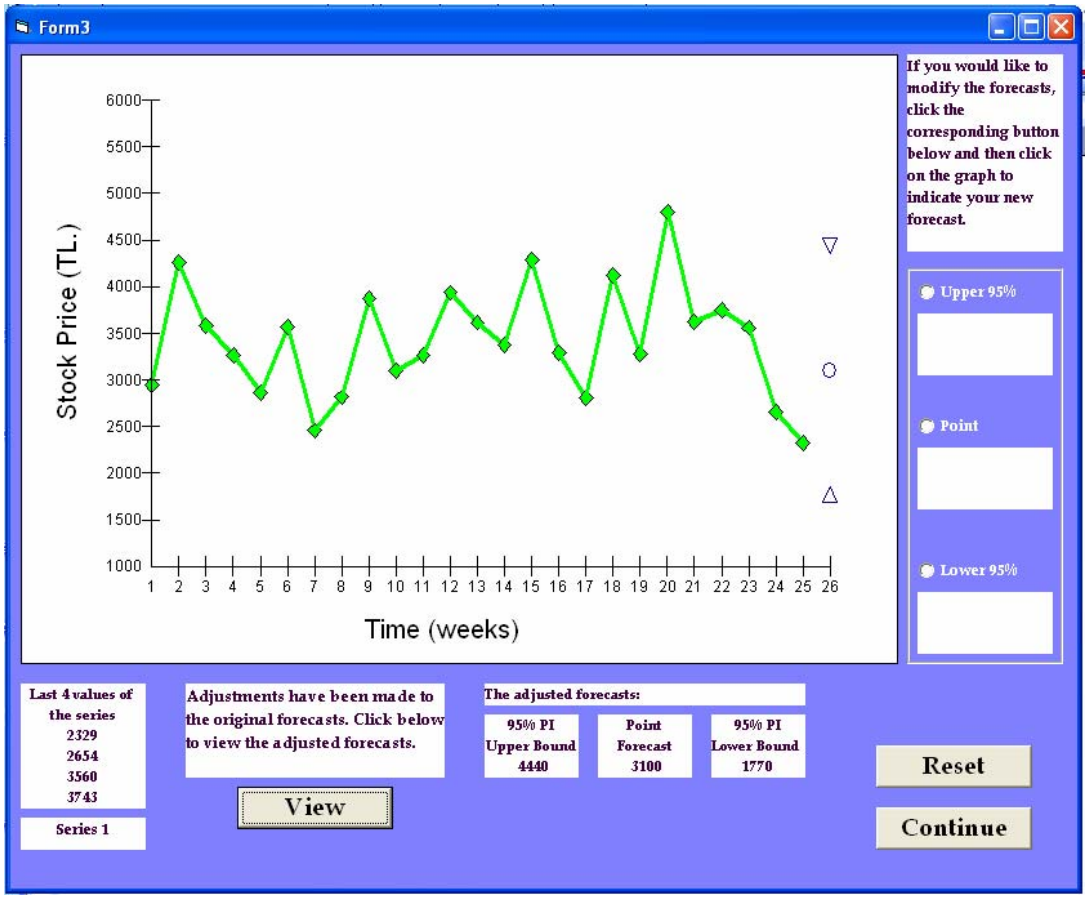


Figure 11. A screenshot from EFSS. – 2nd group

Additionally, short questionnaires were distributed to the participants after the study to gain better understanding and interpretation of their perceptions. These questionnaires were prepared separately for each of the two groups.

5.3.2. Results

The findings are outlined under three categories involving performance measures on point forecasts, performance measures on interval forecasts, and the relative accuracies of adjusted & provided forecasts.

5.3.2.1. Point Forecasts

Table 16 presents the results of the two groups (i.e., the group who believed they were receiving original/unadjusted forecasts and the group who believed they were receiving already adjusted predictions) across the relevant performance measures for point predictions. No significant differences could be found between the two groups in the amount and size of adjustments, as well as in the accuracy of the adjusted point forecasts. That is, thinking that the given point forecasts are or are not already adjusted seems to have no differential effects on further modification of point forecasts.

Table 16. Overall results for point forecasts (numbers in parentheses are the numbers of data points in that category)

	% of point forecasts adjusted	APAP	MAPE
'Original/unadjusted' framing	73.26% (774)	4.88% (774)	20.83% (774)
'Adjusted' framing	71.96% (774)	5.44% (774)	20.46% (774)
	n.s.	n.s.	n.s.

5.3.2.2. Interval Forecasts

Table 17 summarizes the results of the two groups across the pertinent performance measures for interval forecasts. There appears a significant difference on the adjustment behavior of the interval forecasts as a result of adjustment framing. Participants who believed they were receiving already adjusted forecasts have made fewer modifications; these modifications were also smaller in size (as measured via APAI). Furthermore, this adjustment difference did not affect the accuracy of the interval forecasts. The hit rates between the two groups were statistically similar.

Table 17. Overall results for interval forecasts (numbers in parentheses are the numbers of data points in that category)

	% of interval forecasts adjusted	APAI	Hit rate
'Original/unadjusted' framing	83.72% (774)	20.25% (774)	82.43% (774)
'Adjusted' framing	79.33% (774)	17.28% (774)	81.65% (774)
	$t_{(1532)}$= 2.23 p=0.026	$t_{(1409)}$=2.30 p = 0.022	n.s.

Further insight about the adjustments made to the interval forecasts may be gained from Table 18. The participants who believed they were acquiring already

adjusted forecasts were increasing the interval widths less often than the other group. At the same time, for this group, percentage of intervals in which the widths remained the same was significantly higher. This may be interpreted as further support that adjustment framing leads to fewer modifications applied to the provided interval forecasts.

Table 18. Percentage of intervals widened/narrowed/not changed in interval width (total number of intervals in parentheses)

	PERCENTAGE OF INTERVALS			
	widened	narrowed	no change in interval width	TOTAL
'Original/unadjusted' framing	35.92% (278)	47.54% (368)	16.54% (128)	100% (774)
'Adjusted' framing	28.68% (222)	50.00% (387)	21.32% (165)	100% (774)
	z=3.05, p=0.002	n.s.	z=-2.41, p=0.016	

5.3.2.3. Relative Accuracies of Adjusted & Provided Forecasts

Table 19 provides a comparison of the accuracy of adjusted point and interval forecasts relative to the accuracy of the point and interval predictions given to the participants. It can be observed that the errors increase as adjustments are introduced to the point forecasts, and this increase is significant and equivalent for both groups. Regardless of adjustment framing, both groups worsened the accuracy by adjusting the provided forecasts, displaying similar point forecasting performance.

On the other hand, hit rates of adjusted interval forecasts were not significantly different from those of the provided forecasts. Unlike point predictions, adjustments done on interval forecasts did not lead to reduced performance for either group.

Table 19. Accuracy of adjusted point/interval forecasts relative to the accuracy of provided forecasts

POINT FORECASTS	provided forecast MAPE	adjusted forecast MAPE	t-score$_{(773)}$
'Original/unadjusted' framing	17.71%	20.83%	-2.369 (p=0.018)
'Adjusted' framing	17.71%	20.46%	-2.133 (p=0.033)

INTERVAL FORECASTS	provided forecast hit rate	adjusted forecast hit rate	t-score$_{(773)}$
'Original/unadjusted' framing	83.33%	82.43%	n.s.
'Adjusted' framing	83.33%	81.65%	n.s.

5.4. Study 3: Effects of Providing Explanations along with Original and/or Adjusted Forecasts

5.4.1. Methodology and Design

Overview

The third study aimed to examine the effects of two factors that may significantly influence forecast users' adjustments of given predictions: (1) providing explanations for adjustments, and (2) presenting the original/unadjusted forecast along with the adjusted prediction. The first factor, i.e., providing explanations for

prior adjustments to forecasts, may be effective in communicating the reasons and rationale behind the previous modifications, thus persuading users that these adjustments were done because of need and not whim. If so, presenting explanations for prior adjustments may lead to an improvement in the acceptance of already-adjusted forecasts, as signalled via smaller and fewer further modifications. Alternatively, the forecast users may want to 'make their mark' or demonstrate their informational edge, regardless of whether an explanation has been provided about a prior adjustment, leading to modifications independent of whether the external predictions they are given have already been worked on or not.

The second factor concerns the availability of original/unadjusted forecasts along with their adjusted versions. When the users have the opportunity to observe both the adjusted and original predictions simultaneously, they will gain additional information on the direction and magnitude of prior adjustments. This may lead to an appreciation of the previous time and effort already invested in the given forecasts, thus enhancing the credibility of adjustments and leading to fewer and smaller further modifications. On the other hand, the presence of more information may also backfire, leading to cognitive overload, confusing the forecast users and interfering with their adjustments. As a result, users may simply disregard this extra information and carry on with their revisions, so that no differences would be observable between the groups receiving only the adjusted forecasts and the groups receiving both the adjusted and the original forecasts.

Finally, there may be a strong interaction between the two factors. In particular, the combined effects of simultaneously presenting the original forecast, its adjusted version, and explanations regarding the performed adjustments may turn out to be a powerful tool in influencing the adjustment/acceptance process. The third study was designed to explore such potential effects.

Participants

This study was conducted in May 2005. A total of 128 business students participated. The students were mostly second- and third-year students taking courses in Management Science, Production Planning and Forecasting. For their participation they were offered an incentive in the form of extra course credits, the amount of which was dependent on their performance.

Procedure

As in the second study, the current one was also conducted via a computer program written by the researcher, also called "Experimental Forecasting Support System" or EFSS.

As was proposed, the primary purpose of this study was to explore the effects of the presence of explanations and initial forecasts on the nested judgmental adjustment of forecasts. The explanations were about the reasons behind the initial adjustment of forecasts. In this respect, there were two independent variables. The first variable was the presence of explanations (present or absent) and the second variable was the presence of the unadjusted/original

forecasts (present or absent). Since there were two independent variables with two levels each, the best design for this study was a 2x2 factorial design (please see Cochran and Cox (1957) and Montgomery (2001) for a discussion on this design). In the end, there were 4 groups.

To generate these groups, two sets of forecasts were used. These forecasts were based on the same 18 time-series selected for the second study. The first set of point and 95% interval forecasts was the already-available one used in the previous study. This set of forecasts served as the unadjusted/original forecasts.

To generate the second set of forecasts and construct explanations, this initial set was given to an expert on stock-price forecasting. The expert received a similar computer program to generate her explanations and adjustments. Specifically, she was given the time-series data followed by initial forecasts and was asked to make adjustments to the forecasts while writing down the reasons for these adjustments. In this manner, the expert's modified forecasts produced the second set of forecasts.

These data and explanations were then coded into the Experimental Forecasting Support System. Based on this setting, the four groups of the factorial design were as follows:

1st group: This group served as the control group, since both types of decision-support materials (i.e., unadjusted/original forecasts and explanations) were absent. Firstly, time-series plots of the stocks were provided. The participants were asked to click on a button to reveal

adjusted forecasts, with a textbox saying "Adjustments have been made to the forecasts. Click below to view the adjusted forecasts". They had to click on this button to proceed. When the button was clicked on, the second set of forecasts (i.e., those that had been adjusted by the expert) was shown on the graph. (Figure 12 presents a screenshot of the program for the 4th group.) Afterward, the participants were requested to provide their modifications if there were any. They could easily perform the modifications by selecting either the point forecast, the upper 95% bound or the lower 95% bound, and then clicking on the graph to show the modified values. If they did not click on any of these, the forecasts remained unmodified. They could move on to the next set by clicking on a button after this stage. The program terminated when the entire 18 series had been worked on. The program automatically recorded all the adjustments onto a diskette that had been given to each participant before the study. A total of 31 subjects participated in this group.

2nd group: This group received only the unadjusted/original forecasts as decision-support material. First, time-series plots of the stocks were provided. The participants were asked to click on a button to reveal the unadjusted/original forecasts, with a textbox saying "Click below to view the unadjusted/original forecasts". They had to click on this button to proceed. Clicking on this button showed the unadjusted/original forecasts on a graph. Next, a new button appeared with a textbox, saying "Adjustments have been made to the forecasts. Click below to view the

adjusted forecasts". They also had to click on this button to proceed. When the button was clicked on, the second set of forecasts (i.e., those that had been adjusted by the expert) also appeared on the graph. Afterward, the participants were requested to provide their modifications if there were any. They could easily perform the modifications by selecting either the point forecast, the upper 95% bound or the lower 95% bound and then clicking on the graph to see their modified values. If they did not click on any of these, the forecasts remained unmodified. They could move on to the next set by clicking on a button after this stage. The program terminated when the entire 18 series had been worked on. The program automatically recorded all the adjustments onto a diskette that had been given to every participant before the study. A total of 30 subjects participated in this group.

3rd group: This group received only explanations as decision-support material. Initially, time-series plots of the stocks were provided. The participants were then asked to click on a button to reveal the adjusted forecasts, with a textbox saying "Adjustments have been made to the forecasts. Click below to view the adjusted forecasts". They had to click on this button to proceed. When the button was clicked on, the second set of forecasts (i.e., those that had been adjusted by the expert) was shown on the graph. For this group, on the left side of the screen, the explanations for the adjustments emerged when this button was clicked. Afterward, the participants were requested to give their modifications if

there were any. They could easily perform the modifications by selecting either the point forecast, the upper 95% bound or the lower 95% bound and then clicking on the graph to see the modified values. If they did not click on any of these, the forecasts remained unmodified. This group was also required to provide a rating of the information value for each explanation by selecting one of the three options: "misleading", "no real value" and "helpful". They could move on to the next set by clicking on a button after this stage. The program terminated when the entire 18 series had been worked on. The program automatically recorded all the adjustments and information value ratings onto a diskette that had been given to each participant before the study. A total of 32 subjects participated in this group.

4th group: This group received both types of decision-support material, namely the unadjusted/original forecasts and the explanations. As in the previous groups, the time-series plots of the stocks were given initially. The participants were asked to click on a button to reveal the unadjusted/original forecasts, with a textbox saying "Click below to view the unadjusted/original forecasts". They had to click on this button to proceed. Clicking on this button showed the unadjusted/original forecasts on a graph. Then, a new button appeared with a textbox saying "Adjustments have been made to the forecasts. Click below to view the adjusted forecasts". They also had to click on this button. When the button was clicked on, the second set of forecasts (i.e., those that had

been adjusted by the expert) appeared on the graph in addition to the unadjusted/original ones. As in the third group, on the left side of the screen, the explanations for the adjustments emerged when this button was clicked. (Figure 12 presents a screenshot of the program at this stage.) Afterward, the participants were requested to give their modifications if there were any. They could easily perform modifications by selecting either the point forecast, the upper 95% bound or the lower 95% bound and then clicking on the graph for their modified values. If they did not click on any of these, the forecasts remained unmodified. This group was also required to provide a rating of the information value for each explanation by selecting one of three options: “misleading”, “no real value” and “helpful”. They could move on to the next set by clicking on a button after this stage. The program terminated when the entire 18 series had been worked on. The program automatically recorded all the adjustments and information value ratings onto a diskette that had been given to each participant before the study. The remaining 35 subjects participated in this group.

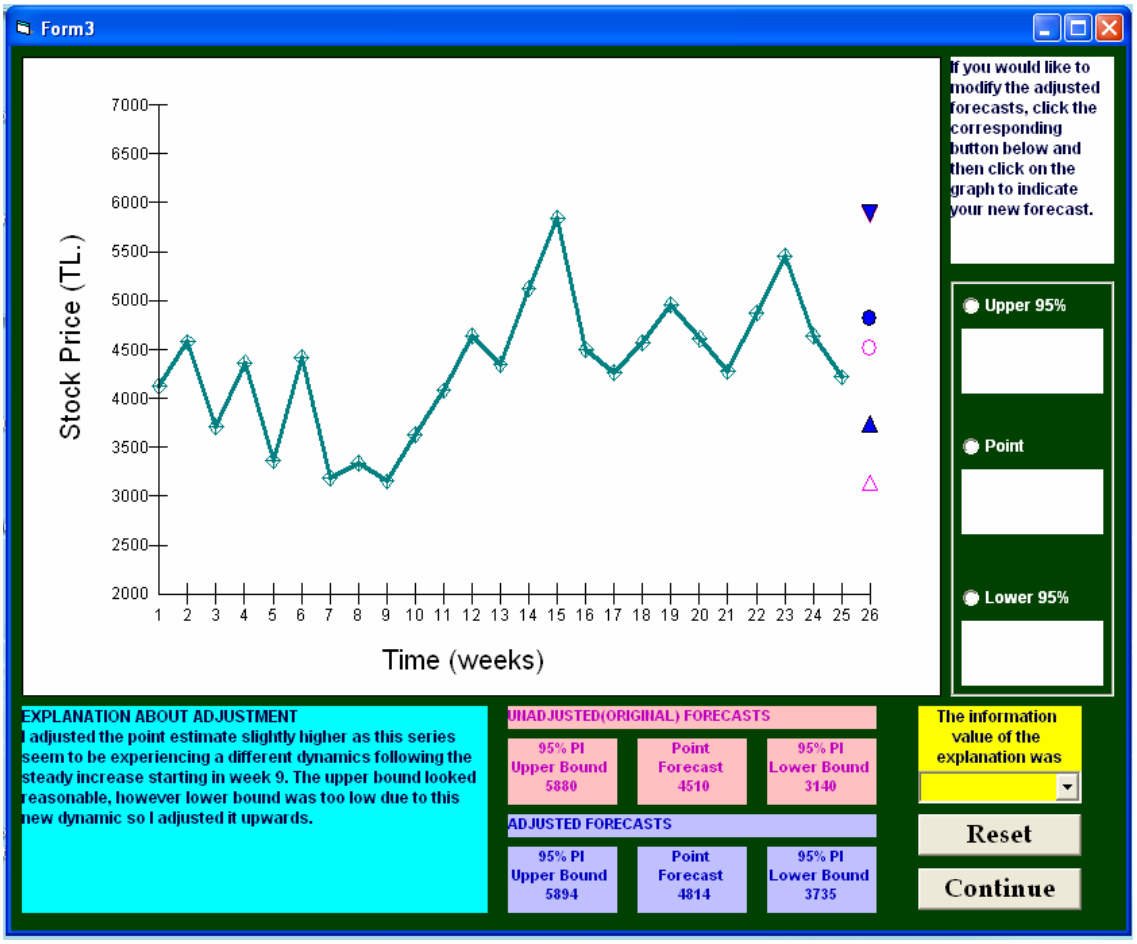


Figure 12. A screenshot from the second EFSS. – 4th group

Additionally, short questionnaires were distributed to the participants after the study to gain a better understanding and interpretation of their perceptions. These questionnaires were prepared separately for each of the four groups.

5.4.2. Results

The findings are outlined under four categories involving performance measures on point forecasts, performance measures on interval forecasts, relative accuracies

of adjusted and provided forecasts and the perceived information value of the explanations.

5.4.2.1. Point Forecasts

Table 20 presents the results of the four groups across the relevant performance measures for point forecasts. No significant differences could be found between the four groups in the accuracy of the adjusted point forecasts (i.e., MAPE). However, significant differences exist in the amount and size of adjustments (i.e., percentage of point forecasts adjusted and APAP scores).

It seems that the groups receiving the explanations (Groups 3 and 4) adjusted a smaller percentage of point forecasts presented to them. Analysis of percentage of point forecasts adjusted reveals a significant main effect for the presence of explanations ($F_{1,2300} = 135.27$, $p<0.001$), with no significant main effect for the presence of original forecasts and no significant interaction effect for the presence of explanations and presence of original predictions. Therefore the significant difference observed among the groups in the frequency of point forecast adjustments ($F_{3,2300} = 45.82$, $p<0.001$) can be attributed to the presence of explanations. Pairwise comparisons provide support to these findings. They indicate that the differences between groups receiving no explanations (Group 1 and Group 2) with those receiving explanations (Group 3 and 4) are highly significant (all p-values are less than 0.0001).

Table 20. Overall results for point forecasts (numbers in parentheses are the numbers of data points in that category)

	% of point forecasts adjusted	APAP	MAPE
Adjusted forecast only	77.06% (558)	5.71% (558)	23.17% (558)
Original + adjusted forecast	77.22% (540)	4.74% (540)	21.36% (540)
Explanation + adjusted forecast	52.26% (576)	3.75% (576)	22.33% (576)
Explanation + original + adjusted forecast	57.14% (630)	4.06% (630)	21.65% (630)
	$F_{3,2300} = 45.82$ $p<0.001$	$F_{3,2300} = 10.52$ $p<0.001$	n.s.

In terms of the size of adjustments made, there exists a significant main effect for the presence of explanations ($F_{1,2300} = 24.39$, $p<0.001$), and a significant interaction effect between the presence of explanations and the presence of original forecasts ($F_{1,2300} = 5.83$, $p=0.016$). This interaction also leads to the significant difference on APAP scores observed among the four groups ($F_{3,2300} = 10.52$, $p<0.001$). The plot given in Figure 13 suggests when no explanations are provided, the presence of the original forecasts in addition to the adjusted predictions has a reducing effect on the size of the further adjustments conducted on the point forecasts. Post-hoc tests signify the difference between Group 1 and 2 ($p=0.012$). On the other hand, when there exists an explanation about the prior adjustments, the presence of the original forecast does not result in a significant difference (between Group 3 and 4) on the size of adjustments made.

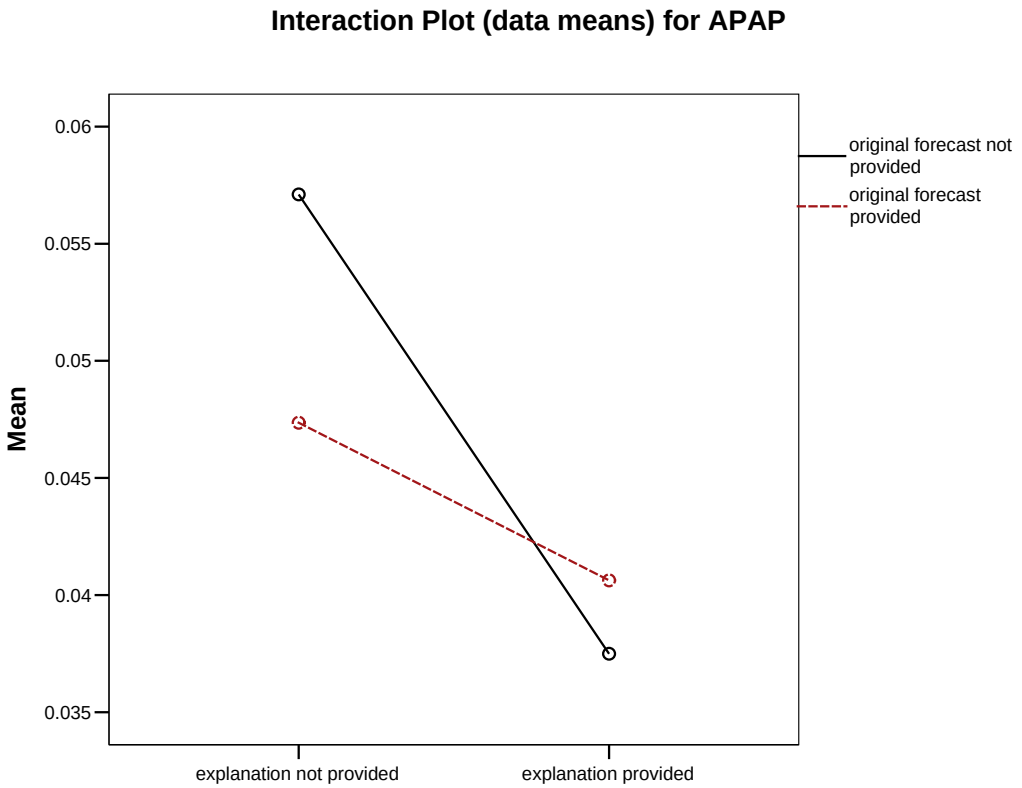


Figure 13. Interaction effect for the *presence of explanations* and the *presence of original forecasts* on APAP scores

It seems that when the users receive already adjusted point forecasts, the factor that influences them the most is the presence of explanations. When explanations about the reasons behind prior adjustments are included, the users appear to show a greater tendency to accept those forecasts. Any further adjustments they might make are not only smaller but also less frequent. In the case of no explanations, if the users have access to the original forecasts in addition to the adjusted predictions, this improves the acceptance of provided forecasts as well (as reflected in smaller further modifications). Interestingly, even though there exist differences between the adjustments of users receiving

explanations and/or original forecasts, this is not reflected in the overall point forecast accuracy of their modified forecasts.

5.4.2.2. Interval Forecasts

Table 21 provides an overview of the findings for interval forecasts. The percentage of interval forecasts adjusted and the size of adjustments are significantly different among the four groups, with no corresponding differences in hit rates.

Table 21. Overall results for interval forecasts (numbers in parentheses are the numbers of data points in that category)

	% of interval forecasts adjusted	APAI	Hit rate
Adjusted forecast only	88.35% (558)	24.09% (558)	77.96% (558)
Original + adjusted forecast	85.19% (540)	14.79% (540)	81.48% (540)
Explanation + adjusted forecast	69.10% (576)	15.17% (576)	80.03% (576)
Explanation + original + adjusted forecast	64.60% (630)	11.01% (630)	83.17% (630)
	$F_{3,2300} = 46.81$ $p<0.001$	$F_{3,2300} = 36.83$ $p<0.001$	n.s.

Analysis on percentage of intervals adjusted suggests that the differences may arise from the statistically significant main effects of the presence of explanations ($F_{1,2300} = 133.42$, $p<0.001$), as well as the presence of original forecasts ($F_{1,2300} = 4.93$, $p=0.026$). Due to these effects, a significant difference among the groups also emerges ($F_{3,2300} = 46.81$, $p<0.001$). The pairwise comparisons indicate that the differences between participants receiving no

explanations (Group 1 and Group 2) and those receiving explanations (Group 3 and 4) are highly significant (all p-values are less than 0.0001). Moreover, accessing original forecasts in addition to explanations leads to marginally fewer adjustments when compared against those receiving only explanations (p=0.059).

Analysis on APAI scores suggest that the differences in size of adjustments ($F_{3,2300} = 36.83$, $p<0.001$) may be due to the significant main effects of providing explanations ($F_{1,2300} = 47.93$, $p<0.001$) and providing original forecasts ($F_{1,2300} = 53.86$, $p<0.001$). In addition, a significant interaction effect between these two factors appears to exist ($F_{1,2300} = 7.81$, $p=0.005$). As also shown in Figure 14, when the original/unadjusted forecasts are present, further modifications are smaller in size independent from the condition that participants are receiving explanations or not. Pairwise comparisons designate that the differences between Group 1 and 2 – explanations absent, as well as between Group 3 and 4 – explanations present, are significant ($p<0.0001$, $p=0.001$ respectively). Similarly, when explanations are present, the participants show a tendency to reduce the size of their further adjustments independent of whether original/unadjusted predictions exist or not. The differences between Group 1 and 3 – original/unadjusted forecasts absent, as well as between Group 2 and 4 – original/unadjusted forecasts present, are significant ($p<0.0001$, $p=0.003$ respectively).

The interaction between the factors reveals a particular synergy that is created when both the explanations and the original forecasts are given simultaneously. In this case, the reduction in the size of adjustments is still

evident, although not as pronounced as the case when either the explanations or the original forecasts are given in isolation. Nevertheless, the participants receiving both explanations and original/unadjusted predictions (Group 4) have introduced the smallest adjustments across all the other groups (all p-values are smaller than 0.001).

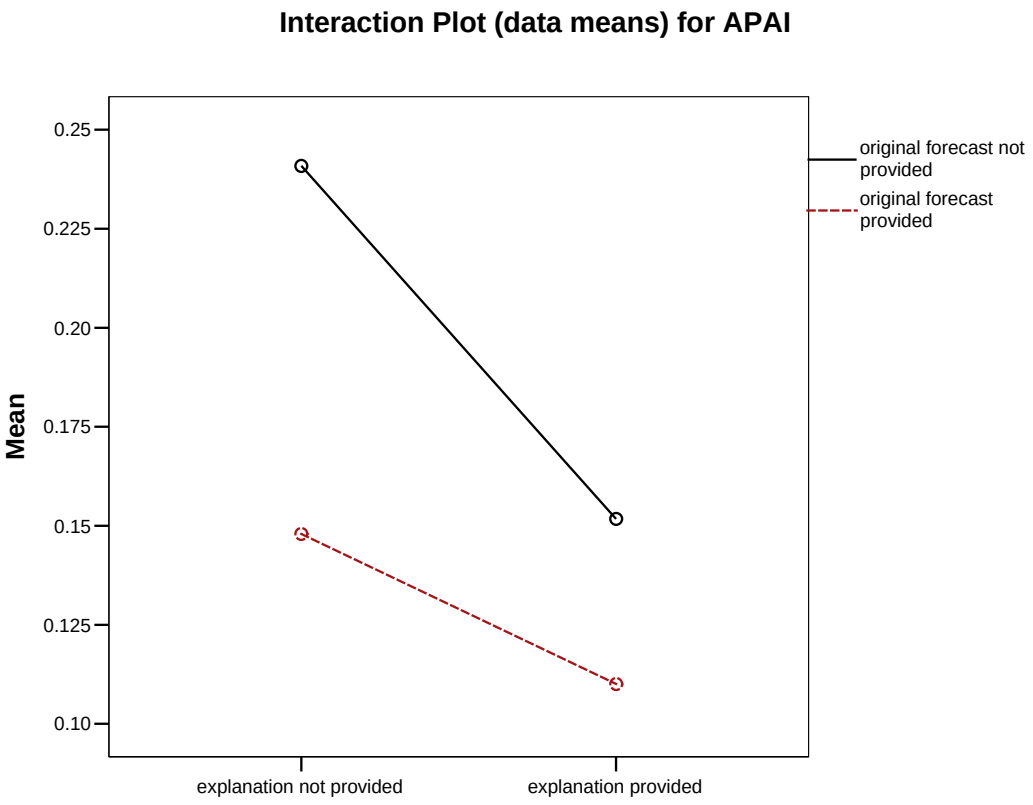


Figure 14. Interaction effect for the *presence of explanations* and the *presence of original forecasts* on APAI scores

In short, both factors appear to have significant effects on adjustments of interval forecasts. When initial adjustments are accompanied by explanations and/or when the users receiving these forecasts can also access the original/unadjusted forecasts, the tendency to accept the given adjusted forecasts

seems to be stronger. Consecutively, further adjustments tend to be less frequent and in smaller magnitudes. However, as in the case of point forecasts, these effects are not reflected in the overall interval forecasting accuracy of the modified predictions.

5.4.2.3. Relative Accuracies of Adjusted & Provided Forecasts

Table 22 provides comparisons of forecasting accuracy at two levels: (i) accuracy performance of the original/unadjusted predictions (*i.e.*, forecasts produced by the smoothing method, as also used and discussed in the second study) versus the performance of expert-adjusted forecasts (*i.e.*, the predictions labeled as adjusted forecasts when presented to the participants) and (ii) accuracy comparison of the participant-adjusted forecasts versus expert-adjusted predictions.

Table 22. Point/interval forecast accuracy of (i) original/unadjusted *vs* expert adjusted forecasts, and (ii) participant-adjusted *vs* expert-adjusted forecasts

ORIGINAL *vs* EXPERT-ADJUSTED	Original/unadjusted forecasts	Expert-adjusted forecasts	
MAPE	17.71%	21.85%	n.s.
Hit Rate	83.33%	83.33%	n.s.

PARTICIPANT-ADJUSTED *vs* EXPERT-ADJUSTED: POINT FORECASTS	Participant - adjusted forecast MAPE	Expert - adj. forecast MAPE	
Adjusted forecast only	23.17%	21.85%	n.s.
Original + adjusted forecast	21.36%	21.85%	n.s.
Explanation + adjusted forecast	22.33%	21.85%	n.s.
Explanation + original + adjusted forecast	21.65%	21.85%	n.s.

PARTICIPANT-ADJUSTED *vs* EXPERT-ADJUSTED: INTERVAL FORECASTS	Participant - adjusted forecast Hit Rate	Expert-adj. forecast Hit Rate	
Adjusted forecast only	77.96%	83.33%	$t_{557} = -3.06$, p=0.002
Original + adjusted forecast	81.48%	83.33%	n.s.
Explanation + adjusted forecast	80.03%	83.33%	$t_{575} = -1.98$ p=0.048
Explanation + original + adjusted forecast	83.17%	83.33%	n.s.

Comparisons of the original (model-driven) and expert-adjusted forecasts reveal no significant differences in either the point or the interval forecasting accuracies. However, particular changes occur when participants make further adjustments on already-adjusted forecasts. The second part of Table 22 shows that, even though further modifications do not weaken the point forecasting performance over the expert-adjusted forecasts, they lead to significantly lower hit rates (in comparison to both the expert-adjusted forecasts as well as the original forecasts) in groups not receiving the original forecasts. Further insight into these findings may be gleaned from Table 23. Results show that the groups not receiving the original forecasts changed the given intervals such that 48% of the modified intervals were made narrower. In contrast, only 39% of the modified intervals were reduced in interval width for groups that were given the original forecasts. These findings may be viewed as suggesting that when only adjusted forecasts are given to the users, their lack of information regarding the original forecasts translates to an unwarranted confidence, leading to tighter intervals and lower accuracy in their modified prediction intervals. On the other hand, when the users are supplied with original forecasts in addition to the adjusted forecasts,

their further modifications do not deteriorate the accuracy of interval forecasts, with the changes in interval widths reflecting a more realistic awareness of uncertainty.

Table 23. Percentage of intervals widened/narrowed/not changed in interval width for groups presented *vs* not presented the original forecasts (total number of intervals in parentheses)

	PERCENTAGE OF INTERVALS			
	widened	narrowed	no change in interval width	TOTAL
No original forecasts presented (GROUPS 1 & 3)	30.33% (344)	48.15% (546)	21.52% (244)	100% (1134)
Original forecasts presented (GROUPS 2 & 4)	34.53% (404)	39.05% (457)	26.41% (309)	100% (1170)
	z=-2.15, p=0.031	z=4.42, p<0.001	z=-2.76, p=0.006	

5.4.2.4. Perceived Information Value of Explanations

Participants in groups supplied with explanations about previous adjustments (Groups 3 and 4) were also asked to rate the perceived information value for each of the provided explanations by selecting one of three categories ("1 = misleading", "2 = no real value", or "3 = helpful"). It was found that the mean information value rating for the group receiving only the explanations (but no original forecasts) was 2.44, while the mean rating for the group receiving both the explanations and the original forecasts was 2.47. Both of these mean ratings are significantly greater than 2 which corresponds to "no real value" (t_{575}=14.02, p<0.001, and t_{629}=15.75, p<0.001, respectively), indicating that the participants

found the explanations somewhat helpful. Table 24 presents the adjustment and accuracy measures that are grouped with respect to the reported perceived information value ratings.

Table 24. Differences in adjustment and accuracy measures with respect to perceived information value

Group receiving adjusted forecasts and explanations for adjustment (no original forecasts)

	Misleading	No real value	Helpful	$F_{2,573}$	p
% of point forecasts adjusted	70.21%	71.43%	40.11%	28.65	<0.001
% of interval forecasts adjusted	86.17%	85.71%	58.17%	26.97	<0.001
APAP	6.88%	6.26%	1.95%	56.90	<0.001
APAI	23.29%	19.32%	11.41%	17.00	<0.001
MAPE	24.33%	28.77%	19.33%	n.s.	
Hit rate	77.66%	75.19%	82.52%	n.s.	

Group receiving original forecasts, adjusted forecasts, and explanations for adjustment

	Misleading	No real value	Helpful	$F_{2,627}$	p
% of point forecasts adjusted	81.82%	71.11%	46.21%	29.80	<0.001
% of interval forecasts adjusted	89.90%	77.88%	53.79%	31.91	<0.001
APAP	8.32%	5.25%	2.59%	53.30	<0.001
APAI	18.59%	14.36%	7.96%	32.34	<0.001
MAPE	35.89%	22.22%	17.89%	8.59	<0.001
Hit rate	72.73%	83.70%	85.61%	4.76	0.009

As shown in Table 24, perceived information values of explanations are highly related with the subsequent adjustments performed on them. For both groups, when the explanations are perceived to be helpful, the amount and size of adjustments in both point and interval forecasts seem to decrease significantly (as shown via the percentage of point and interval forecasts adjusted, APAP, and APAI). Furthermore, when original forecasts are provided in addition to the

adjusted forecasts and associated explanations, a significant improvement in accuracy (as indexed via MAPE and hit rate) is evident with increasing perceived informativeness. Even though these results are not statistically significant for the group not receiving the original forecasts, the effects are clearly in the same direction. It may be argued that as the users find the provided explanations more helpful, their trust in the already-adjusted forecasts seem to improve, leading to a higher rate of acceptance. This acceptance brings smaller and less frequent adjustments to both the point and interval forecasts, also translating to improved forecast performance when the users are equipped with the original forecasts, so that they can more clearly see the direction and the size of adjustments.

CHAPTER 6

GENERAL DISCUSSION AND CONCLUSIONS

In this thesis, research composed of two parts has been reported. The first part consisted of a field survey conducted with real business professionals to extend our knowledge regarding actual forecast-using practice in organizations. The second part involved three laboratory studies to provide a controlled environment for the investigation of various concepts closely related with forecast users.

The survey part of the research provided invaluable information about practitioners who use financial forecasts in their routine conduct of business. First of all, the collected information shed some light on what users anticipate from external forecasts, what constitutes a high-quality prediction and, consequently, the nature of the relationship between expectations and quality perceptions. The second focal point in the study concerned the acceptance and adjustment process. The reasons and motivations leading practitioners to introduce an adjustment or accept forecasts without making modifications gained a central position in the discussion. These two points basically formed the main theme of the survey. However, the information gathered also allowed us to gain notable insight and

perspective on many other relevant factors affecting the forecast-using practice. Differences and similarities among the practitioners choosing to acquire forecasts from a single or multiple sources; the effects of the presence of a systematic or personal feedback function; and how professionals resembled or differed from one another in terms of their approach to acquired forecasts when their experience and position varied were among those factors.

The laboratory studies composing the second part of the research were based on two main concepts that are highly important for forecast users: conducting multiple judgmental adjustments on acquired forecasts and the effects of explanation provision along with these predictions. The first laboratory study focused on the latter concept and investigated the impact of structural characteristics of explanations on the adjustment of provided forecasts in terms of long vs. short explanations and strong vs. weak confidence conveyed via the style of language used in the explanation. In turn, the second and third studies concentrated on the former theme; namely, further adjusting already-adjusted forecasts. Through the second study, an attempt was made to discover the basic characteristics of this process – that is, whether there existed distinct behaviour when forecast users believed they were adjusting an original/unadjusted forecast or adjusting an already-adjusted one. The third study elaborated on the second study by investigating the effects of providing explanations and the original/unadjusted forecasts along with the already-adjusted predictions. Through these studies, invaluable insight and information were gained on this scarcely-

researched judgmental process which carries fundamental significance for forecast users with essential practical ramifications.

This last section of the thesis is devoted to general discussions and conclusions on these findings in terms of their contribution to the existing literature, their limitations, their potential future research directions and their practical implications for the process of organizational forecasting.

6.1. Forecast-Using Practice: Expectations and Adjustments

Generally speaking, the fact that a forecast holds up to expectations does not necessarily mean it will be perceived as a high-quality forecast. What a practitioner anticipates from a forecast may not be the same as what he/she values. The aspects of the former concept may not match the aspects of the latter concept. In such a case, a forecast perceived as a high-quality prediction may not fully satisfy the expectations of the user acquiring it. This will create mental contradictions and user dissatisfaction about the received forecasts, and may turn out to be problematic for the forecasting process in the company. However, if these discrepancies remain at a minimal level and there is a close relationship between the major aspects of expectations and quality perceptions, this will generate a more suitable environment for the smooth processing of the forecasting practice. The data gathered from the survey mainly suggests the presence of such a situation. The dimensions of forecast expectations and quality perceptions were closely related with each other except for some minor differences between the two.

Before disclosing these distinctions in dimensions, the discussion will focus on the similarities.

The common aspects between forecast expectations and perceptions of quality reveal that what forecasting practitioners essentially seek out is an accurate, timely and reliable forecast based on justifiable assumptions. They value the forecasts' being prompt; being there when they need it. They also expect them to stand on solid ground and offer predictions as error-free as possible. The open-ended questions in the survey also provided supportive evidence for these dimensions. The most frequent answers pointed out that the most valuable forecasts were those guiding users to utilize their money in the best way possible and make profitable investments. These forecasts should also provide insights about the future and direct practitioners to plan for upcoming uncertainties as best as possible. The practitioners clearly emphasized that what they really expect from forecasts and what really makes them high-quality are not solely their attributes, but also the end results achieved in their business by utilizing these predictions. They also underlined that, for the achievement of satisfactory end-results, forecasts should have some necessary attributes like being credible, consistent, justifiable, timely and undeniably, highly accurate.

Being accurate was indeed rated to be one of the most important properties sought from forecasts, as well as being recognized as a fundamental attribute of forecast quality. Not only were forecast users in search of accurate predictions, but they also held very high expectations for the level of accuracy. The smaller the amount of error, the better it was from their perspective. The corresponding

dimensions were rated significantly more favourably than the dimensions associated with the presence of tolerable amounts of error. In this sense, the literature's insistence on forecast accuracy is well-justified. However, it was also clear that accuracy was not the only aspect that users wished to see in forecasts. The data gathered from the survey strongly suggest the multi-dimensional nature of forecast expectations and quality perceptions. If the forecasts were not reliable and justifiable, and if they were not received on time, the accuracy would definitely lose much of its meaning. Being accurate alone was not adequate for a forecast to be reckoned as a high-quality forecast that might satisfy expectations. An acquired forecast may be as accurate as possible; however, if it is not utilized or does not end up with good final results due to its lacking in other dimensions, it will definitely be less worthy than a less accurate forecast faring better in other dimensions.

For one particular group of practitioners, the concept of minimum vs. tolerable error levels constituted one of the distinctions between what forecast users expect from forecasts and what they perceive as an attribute of quality. Experienced forecast users (practitioners with more than nine years of familiarity with financial forecasts) were well aware that the existence of very small amounts of error might be unrealistic on many occasions. Thus, expecting to have errors within an acceptable range rather than at a minimum was much more natural for the forecasting process. This realization seemed to occur as experience and familiarity with forecasts advanced, and was not so strong for less-experienced groups. On the other hand, where quality was concerned, there was an agreement

about accuracy across all the experience groups: a high-quality forecast should have the smallest amount of error possible.

Overall, having a reasonable cost for forecasts was perceived to be the least important criteria when compared with the other dimensions. The price paid did not hold as much significance as the other aspects. If a prediction was faring adequately on other dimensions of expectations and quality perceptions, then the price paid for this prediction would carry less importance. This can lead directly to the deduction that practitioners are used to anticipating higher costs for better-quality forecasts, which can also meet up better with their expectations.

The concurrence of this dimension for both expectations and quality perceptions was not paralleled for one group of practitioners in a particular position within their firms. Partner/owners seemed to give more importance to the acquisition of reasonably-priced forecasts than the practitioners in other groups; however, when it came to the attributes of a high-quality forecast, similar to the other groups, the price paid did not carry much importance for this group. This situation has the potential to create problems for the acquisition and purchase of external forecasts if the users are mainly partners and owners in the company. In these situations, it seems better to emphasize the quality aspects of predictions, which may in turn receive better reactions from partners/owners, and will likely relax their price concerns.

The extent to which acquired forecasts satisfy the anticipations of users is directly related with the acceptance and adjustment of these predictions by

practitioners. Our survey data have shown that application of judgmental adjustments is a quite common and popular practice among forecast users, reinforcing the findings of previous surveys in the field (Sanders and Manrodt, 1994; 2003). More than 76% of practitioners reported that they adjust acquired predictions on an occasional-to-regular basis. There were many underlying reasons and motivations influencing their adjustments; however, the presence of adjustments itself was a pretty robust effect. The analysis revealed that the frequency of making adjustments was not affected by the forecast-using experience or position of the practitioner. It was also independent of whether forecasts were acquired from multiple sources or received from a single source. Regardless of the state of these factors, throughout this process, practitioners regularly introduced their judgments onto forecasts they received. Furthermore, application of adjustments was not necessarily confined to a single step. If inside the company there existed other peers checking the adjusted forecasts, the presence of further adjustments was reported by more than 80% of forecast users. Thus, multi-tier or nested adjustments conducted on already-adjusted forecasts seem to be an important component of the whole forecasting process as well.

Some interesting patterns emerged regarding reasons and motivations driving the acceptance and adjustment process. For occasions where practitioners chose to accept forecasts without revisions, there was a strong impact of perceived source characteristics. If practitioners believed that the source of the forecasts was reliable, well-known and objective, this generally induced them not to perform any adjustments. These reasons not only occurred quite frequently, but were also

given high importance ratings. The significance relayed to the source dimensions is emphasized by the finding that the influence of source is not affected by the frequency of adjustments employed. Regardless of how often a practitioner conducted revisions, the forecasting source was conceived to be equally influential at all times. Persuasive cues present in forecasts were also important for their acceptance. If the methods and analysis used in the forecasts were convincing and presented in a powerful way with good language style, this also constituted an essential motive for refraining from making adjustments. The presence of persuasive explanations fashioned reactions in a very similar manner.

Answers to the open-ended questions further confirmed these findings. Practitioners reported that if forecasts were acquired from a credible source, this generally comprised an adequate reason for accepting them without modifications. They also asserted that the predictions they left unadjusted were the ones perceived as logical, realistic, consistent, trustworthy, justifiable and highly persuasive. Based on this information, it wouldn't be wrong to propose that it is very important for forecast suppliers to persuade their users about the value of their forecasts if they want to facilitate the acceptance of their predictions and the eventual continuation of business. For this purpose, some suggested methods include providing ample evidence that their methods and analysis are sound and reliable, presenting findings with a good style or presentation and accompanying their forecasts with appropriate explanations. The increase in persuasive power is also expected to contribute positively to perceptions about the source, and thus have an intensifying impact.

Paralleling the importance given to the presence of persuasive components for acceptance, their absence was a strong motivation to induce the application of judgmental adjustments. Intervening in forecasts through revisions was a natural response if the employed methods and analysis, accompanying explanations, style of language and presentation were not sufficient to persuade forecast users of the worth and adequacy of the predictions. Importance given to these components was also independent of how often practitioners made their adjustments. The lack of these components was an equally essential reason among both rarely-adjusting and frequently-adjusting forecast users.

Perhaps more important than reasons based on persuasiveness were reasons related with making a contribution to forecasts. When practitioners wished to integrate their intuition, experience and knowledge into the forecasts or simply wanted to gain a sense of ownership and responsibility for the predictions, making adjustments on them was rated to be a crucial and frequent procedure. Similarly, adjusting forecasts to accommodate the effects of novel and unexpected events or excluded information was claimed to be a highly important motivation. The significance given to these dimensions increased as the position of the practitioner rose. Partners/owners and CEOs/executive board members provided higher importance ratings than managers and staff members.

Answers to the open-ended questions reinforced and expanded these results. The most frequently mentioned reasons for making an adjustment converged on practitioners' desire for the incorporation of their judgment and

experience. Additionally, making revisions was a common reaction if predictions were perceived to be inadequate, illogical, unrealistic and inconsistent. Finally, adjustments were claimed to be done when the acquired forecasts needed to be attuned to the aims, targets and future expectations of the company in the case of an existing discrepancy. This reporting strongly designates the pressures and constraints imposed on the forecasting process through organizational politics and strategies. Given the scarcity of studies done particularly on this issue, it is imperative that future research addresses this venue in order to gain better comprehension of the practical functioning of organizational forecasting.

As is evident in the influence of persuasive cues and components, there is a close association between the reasons and motivations behind making adjustments or refraining from making them. Practitioners gave similar importance to the presence of persuasive elements to induce acceptance, with their absence being a facilitator for adjusting. For some of the other dimensions, there existed some differences. Making a revision to take responsibility for a forecast was perceived to be motivationally more important than its counterpart. When the impact of the forecast source was concerned, the influence and significance of a reliable, well-known, objective source was greater than that of an unreliable source. The former's facilitation of forecast acceptance was more powerful than the latter's facilitation of forecast adjustment. All these points strongly suggest that source characteristics and trust felt in the source are critical factors that shape the forecasting process in organizations. However, our knowledge on the topic is very limited since this area has been left largely unexplored by the forecasting

literature. There seems to be a vast opportunity for research in this area, with important theoretical and practical implications.

From the survey study, we were able to gather additional interesting and important information about forecast users and the organizational forecasting practice. It is now time to direct our discussion to these issues, starting with practitioners' acquisition of forecasts from single or multiple sources.

The major difference between practitioners receiving predictions from multiple sources and those receiving them from a single source seems to depend on the importance given to the forecasting function in the practitioner's organization. Receiving forecasts from multiple sources corresponds to having more information to handle, making comparisons, having more processes (either in terms of processing forecasts to choose one among the many or preparing a combination of them) and spending more time. This kind of effort can only be meaningful if the forecasting function is allocated an essential part in decision-making activities of such firms when compared against companies receiving forecasts from a single source. This is also associated with practitioners having higher standards and expectations from acquired forecasts in the case of multiple sources, leading to more critical and selective behaviour. Analysis of the survey data provided findings in line with these deductions. Practitioners acquiring forecasts from multiple sources reported holding greater expectations of receiving timely predictions based on justifiable foundations than those in the single-source group. Their quality perceptions showed a similar pattern in these dimensions. At the same time, these practitioners demanded scenarios, alternative forecasts and

explanations more often than the latter group. Having persuasive explanations was a greater motive for the multiple-source group to accept forecasts, while a lack of persuasive explanations and methods/analysis were more important for the introduction of adjustments. All these results are logically convergent, providing support to the arguments. However, future research should further investigate these issues, to confirm, refine and elaborate the findings given.

One last point to make is about the presence of organizational or personal feedback systems. As further proof for the relatively greater importance given to the forecasting function, practitioners receiving forecasts from multiple sources implemented feedback mechanisms approximately twice as often as the other group. In this setting, implementation of feedback helps the former group in organizing, managing and keeping track of the multitude of forecasts they receive at one particular time. The incidence of such a system was also higher among practitioners making frequent adjustments with respect to rare adjusters. This is to be expected, since frequent adjusting also brings the need for more control and more organized accuracy checks in the former group.

Independent of these varying characteristics among the practitioners, all their open-ended answers univocally converged on the benefits gained from such systems. They reported that a feedback function was essential for checking the accuracy of acquired/adjusted forecasts and also the quality of past decisions. Feedback applications were also claimed to prevent practitioners from repeating past mistakes and provide an opportunity to learn. Finally, they provide better control of the process and serve as guides for future predictions and adjustments.

Despite this general consensus on the advantages of feedback systems, the survey data revealed that the ratio of their presence over all the participants was only a bit above 35%. This ratio seems to be quite low given all the comments made and the supporting evidence from the judgmental forecasting literature on its benefits (O'Connor and Lawrence, 1989; Benson and Önkal, 1992; Bolger and Wright, 1994; Önkal and Muradoğlu, 1995; Remus et al., 1996; Sanders, 1997; Bolger and Önkal, 2004; Goodwin et al., 2004; O'Connor et al., 2005). There appear to be barriers and difficulties for the widespread utilization of feedback systems. There is a strong need for research concentrating on these issues in order to expand their deployment and to improve their integration into routine forecasting practice in organizations.

However all these generalizations and recommendations should be considered taking into account the limitations of survey methodology. One of the drawbacks is related with the selection of the sample. Using a non-probability sampling method was thought to be the best alternative given the current population, the task and the complexity of accessing the individuals. The distribution of the responses with respect to firm size and industry demonstrated that the respondents can be considered as representative of the population. Therefore, for the current survey, having generalizability problems due to the sampling method seems highly unlikely.

Another limitation is about the general problems of survey research. As it has been frequently addressed in literature (Frankfort-Nachmias and Nachmias, 1996; Burns and Bush, 2000), there are many pertinent problems with the quality

of data obtained from surveys. One can rarely be sure of the sincerity, motivation, dedication and consistency of the respondents in answering the questions. There are also issues of giving socially desirable answers rather than the true answers. As precautions against these possible errors, filtering, monitoring and carefully checking the answers are among the recommended courses of action (Frankfort-Nachmias and Nachmias, 1996; Burns and Bush, 2000). In this regard, for minimizing these potential errors both the researchers and the hired supervisor performed regular controls on the answers, made random checks through phone/email and rigorously eliminated the forms in case any suspicion was present.

Even though the presence of all these potential problems have been acknowledged in the literature since 1950s (e.g. Deming, 1950), survey methodology has been and still is one of the most popular, widely-used and well-established research methods in the field. Therefore, if executed carefully and meticulously, the value of the results obtained through this technique is indisputable.

6.2. Judgmental Adjustments: Structural Characteristics of Explanations

Findings from the first laboratory-based study suggest that presenting an interval forecast together with a pertinent explanation has an influence on the adjustment, and therefore, on the acceptance of that forecast. However, these results appear to

be contingent on the structural characteristics of the explanations. In particular, long explanations expressed in a strongly confident style seem to be the most influential compared to others, leading to fewer and smaller adjustments. By long explanations, we mean explanations that are "adequately long", so that all points are clearly and openly presented. By strongly confident wording, we mean that an explanation should give a strong impression that the information provided is justifiable. In the forecasting context, a confident explanation should provide a strong impression that the providers of the forecasts believe that their forecasts are accurate, justifiable, and realizable. When these conditions are met, a long and confident explanation leads forecast users to introduce judgmental adjustments less often, and the magnitudes of the adjustments appear to be smaller as well. Similarly, if short explanations are demanded, it seems better to present them in weakly confident wording. Moreover, perceptions of the usefulness of long/confident or short/weakly confident explanations (i.e., their perceived informational value) seem to be higher than other explanations. In short, such explanations are likely to serve the purposes of a forecast provider better than others.

However, all of the above arguments are supported only for interval forecasts. The various structural characteristics of explanations have no statistically significant effect on point forecasts. The attributes of a provided explanation for a specific data set seems to affect the interval forecast adjustment on that specific data set; and yet, adjustment of the point forecast for that specific data set appears to be independent of that explanation's attributes. A potential

reason lies in the choice of experimental manipulations. It may be that the manipulations used were not particularly suited to obtain any direct effects on point forecasts. Point forecasts are solitary numbers meant to contain a variety of information, both internal and external, and it is generally very difficult to incorporate all the information within a single number. In this regard, trying to reach an accurate point forecast becomes a heavy judgmental burden, especially since point predictions appear to convey a false sense of certainty. Thus, people may find it difficult to incorporate the cues of trust or confidence they have received from the explanations into their point forecasts. They may have simply adjusted the point forecasts according to mental heuristics or shortcuts, which make them less responsive to manipulations conducted through the framing of explanations. It has been found that a lower percentage of point forecasts are adjusted relative to interval predictions. The size of adjustments made is also smaller. In short, it is highly likely that point forecasts may respond to different kinds of manipulations than interval forecasts.

Empirical findings indicate that the accuracy of adjusted forecasts, whether point or interval, is less than that of statistically generated forecasts. Individuals are not able to achieve better accuracy by making judgmental adjustments to point or interval forecasts, but instead have to compromise from the attainable performance were they to accept the external forecasts as presented. These findings confirm previous suggestions that judgmental adjustments to statistical forecasts may sometimes lead to decreased accuracy (Carbone et al., 1983; Remus et al., 1995; Lim and O'Connor, 1995). Depending on the quality of

forecast users' contextual information, explanations which are effective in persuading people not to make an adjustment may also be desirable from an accuracy perspective – since less adjustment to the provided forecasts also coincides with less deterioration in the accuracy of statistical forecasts. It may be suggested that long and strongly confident explanations appear to be promising candidates for this purpose, since they are found to be highly effective in reducing the amount of adjustments to interval forecasts.

Another important finding from the first study is related with the impact of the perceived informational value of an explanation on the adjustment and acceptance of its accompanying forecast. Independent of the state of their structural characteristics, explanations perceived as having high informational value were found to be much more persuasive than those seen to have low value, and this held true for the predictions both in point and interval formats. It may be the case that the informative content of an explanation is a more critical component than its structural attributes for inducing the acceptance of a provided forecast.

However, quite naturally, generalizing from these findings is prone to certain limitations. The most critical hindrance to generalization is related with the assumptions and shortcomings of the experimental designs (please see Montgomery, 2001 for a detailed discussion on these issues) and the subjects participating in the experiments. The ideal subjects to participate in these laboratory studies would be the real forecast users themselves. Nevertheless, the reluctance and lack of time/devotion of practitioners in such academic studies had

rendered this option unattainable. Therefore we had to revert to using student participants as proxies for real forecast users and decision makers. Even though this situation is less than ideal, the use of student participants is a common practice in psychological and judgmental forecasting research (Goodwin, 1998; Webby and O'Connor, 1996). There is plenty of evidence which shows that students are adequate surrogates for practitioners in decision-making (Ashton and Kramer, 1980; Remus, 1996) and other business contexts (Locke, 1986; Houghton and Hronsky, 1993).

Another limitation of the experiments is related with the use of artificially constructed time-series data that are framed as real stock prices. In order not to confound the results with the complicated characteristics of real time series and to attain controlled levels of trend and variability, we have used artificial data constructed by a simple process. At the same time, we also benefited from the fact that the artificial series were free from the influence of any extraneous effects that might have been present in the real stock-price data. In this sense, the series utilized were effective in achieving our objectives and minimizing any interference not related with the variables of interest. However, this condition also creates problems of experimental realism and generalization to real-time series data. Given this issue, the results obtained should be interpreted as initial steps in identifying the effects of the variables under minimal external interference and further work must build on them to elaborate by using real stock price data.

Having mentioned these pertinent limitations, it is better to close the section by reporting some practical implications and recommendations. The

results obtained in this study have many practical ramifications in a variety of fields. The most apparent impact would be on institutions operating in the professional consulting, investment advice and forecasting areas. Supplying explanations carries a special significance for the financial interests of such firms, whose success is mostly dependent on their ability to convince their customers/users about the value of the information support they are providing. If the users are not satisfied with the presented advice/forecasts, they may always switch to competing information providers. In this respect, if the forecast suppliers wish to improve their predictions' acceptance, then accompanying their forecasts with informative explanations in long and confident style seems to be a good strategy.

Moreover, our findings also offer important suggestions for explanation facilities integrated into decision-support systems. Although there has been extensive research on designing these facilities, considerably less attention has been given to the structural characteristics of the explanations they generate. The results of current work propose directions for promising future work in this area.

6.3. Nested Judgmental Adjustments: Explanations and Original Forecasts

In many organizations, it is quite common for forecast users to receive predictions that have previously been adjusted by the providers or other users of the forecasts. To investigate some of the factors that may influence the size and propensity of

further adjustments on already-adjusted forecasts, two laboratory-based studies were reported. These studies focused on the potential effects of adjustment framing (Study 2), and the availability of explanations and/or original forecasts alongside adjusted forecasts (Study 3).

Overall, the findings suggest that the effects may be contingent on the forecast format. For point forecasts, whether the user receives an original or adjusted forecast may not make a significant difference. However, in the case of interval forecasts, users who assume they have been provided with already-adjusted predictions tend to introduce fewer and smaller adjustments (in comparison to users who assume they have been given unmodified/original predictions). Our exit interviews with the participants also confirm that forecast users appear to show some reluctance to introduce further modifications on a forecast that has already been worked on and thought about. Also, participants given original predictions indicated that they would adjust less if they were to receive already-adjusted forecasts instead. As might be expected, the opposite comment was made by those led to believe that they had been given already-adjusted forecasts – i.e., they indicated that they would introduce more adjustments if they received predictions that had not been worked on. These comments may reflect implicit strategies of forecast users and their efforts to refrain themselves from the cognitive burden of judgmental adjustments; or it may be that they implicitly attach a higher value to others' judgments. Irrespective of their reasons, however, users still adjust even already-adjusted forecasts, but they do it to a lesser extent.

Provision of explanations for adjustments seems to affect both the size and frequency of pending modifications. Users who are given explanations accompanying adjusted forecasts appear to have an increased acceptance of the conveyed predictions, hence introducing smaller and fewer further modifications. Our participants particularly emphasized how seeing the explanations behind the initial adjustments contributed to their understanding of the behaviour of the series, the forecasts given, and the adjustments conducted on them. They stressed that the explanations provided insights and highlighted specific points that they would not have been particularly aware of on their own.

In addition to the presence of an explanation, the informational value attributed to it also has an impact on the adjustment/acceptance behaviour, paralleling the findings of the first study. The more informative an explanation is perceived to be, the more influence it has on the users – as reflected in the frequency and size of adjustments. In the case of a seemingly contradictory explanation, the reverse behaviour in adjustments occurs – some participants commented that they especially adjusted more if they thought the explanation was somewhat contradictory or misleading.

The presence of original forecasts (in addition to their adjusted versions) appears to have similar effects to those observed with provided explanations. Having access to the original/unadjusted forecasts appears to give an informational edge, leading to a better understanding of previous adjustments. This appreciation/understanding apparently facilitates forecasts' acceptance and

leads to fewer modifications with smaller magnitudes. Participant comments indicated that the existence of original forecasts provided a different perspective or insight into the given forecasts, enabling better comparisons, and thus facilitating future modification decisions. There was a strong indication that none of the participants viewed the original forecasts as a further cognitive load or as a barrier. They seemed to prefer having access to this additional perspective. Furthermore, some participants made modifications to carry the forecasts to somewhere between the original and adjusted predictions. This adjustment direction may hint at an important motive in the context of multiple/nested adjustments. Such potentially influential factors may alter the nested adjustment process, presenting a promising direction for further research.

Overall, participants appeared to alter their forecasting performance by introducing further adjustments. In the second study, when participant-adjusted forecasts were compared with the accuracy of provided forecasts, adjusted point forecasts showed higher errors relative to those of provided point forecasts. The interval format seemed more robust to this effect, with no accompanying deterioration in hit rates. In the third study, when expert-adjusted forecasts were given with or without explanations and/or original forecasts, the portrait somewhat changed. Further revisions on already-adjusted forecasts did not lead to significant deterioration in point forecasting accuracy over expert-adjusted forecasts. As in the second study, both the expert's and the participants' interval forecasts showed equally successful hit rates when original forecasts were given. However, if participants did not have access to original forecasts, they clearly

showed a tendency to reduce the interval widths, leading to lower hit rates. These findings once again highlight the importance of forecasting format, confirming earlier results on differences in user perceptions of predictive formats (Önkal and Bolger, 2004; Yates, et al., 1996). As is also apparent from the first study, point and interval forecasts respond to different criteria and there are significant differences in their respective adjustments. Future work is needed to understand the different cognitive processes commanding these two widely-used expressions of forecast communication.

The findings from the second and third studies may have important repercussions for the forecasting process in organizations. Forecasts that are generated in a particular unit and transferred through other units may undergo known or unknown adjustments. Our results suggest that even though the forecasts are identical, believing that a particular forecast has undergone a previous adjustment makes a difference for the latter forecast users. Moreover, this process responds to the presence of explanations about revisions and whether initial/unadjusted forecasts are or are not made accessible.

Such differences in adjustment tendencies may in fact be utilized to an organization's advantage. If the former forecast users do not want to restrain any subsequent adjustments on the predictions they are disseminating to other users, it is better for them to disclose no information about whether the forecasts have been subject to previous adjustments. If further revisions are welcome, former users are also better off not presenting any explanations or initial forecasts. On the other hand, if the former forecast users wish to restrict or limit the amount of

adjustments that may be conducted by the latter users, they are encouraged to reveal that the forecasts have already been worked on and adjusted. Moreover, it is strongly advisable to attach a set of explanations about the adjustments they have conducted in addition to the original/unadjusted forecasts. Such a 'forecast trail' can be expected to result in improved acceptance, translating to less frequent and smaller subsequent adjustments.

The studies reported here provide preliminary investigations of user adjustment. Future work with practitioners in organizational settings will enhance our understanding of adjustment processes under the impending constraints of organizational politics, motivational contingencies, and informational externalities. Designing effective support systems for multi-tier or nested forecasts relies on confronting and synchronizing the intricate processes behind forecast adjustment and communication. Further work in these venues will be imperative for improving organizational forecasting performance.

BIBLIOGRAPHY

Ackert, L.F., B. K. Church and M. Shehata 1997 "The Experimental Examination of the Effects of Forecast Bias on Individuals' Use of Forecasted Information," *Journal of Accounting Research* 35(1): 25-42.

Adam, E.E. and R.J. Ebert. 1976 "A Comparison of Human and Statistical Forecasting," *AIIE Transactions* 8: 120–127.

Andersson, P., J. Edman and M. Ekman. 2005. "Predicting the World Cup 2002 in Soccer: Performance and Confidence of Experts and Non-experts," *International Journal of Forecasting* 21: 565-576.

Angus-Leppan, P. and V. Fatseas. 1986. "The Forecasting Accuracy of Trainee Accountants Using Judgmental and Statistical Techniques," *Accounting and Business Research* 16: 179–188.

Arkes, H.R. 2001. "Overconfidence in Judgmental Forecasting" In S.Armstrong, ed., *Principles of Forecasting*. Mass:Kluwer Academic Publishers.

Ashton, R.H. and S.S. Kramer. 1980. "Students as Surrogates in Behavioral Accounting Research: Some Evidence," *Journal of Accounting Research* 18: 1-15.

Aukutsionek, S.P. and A.V. Belianin. 2001. "Quality of Forecasts and Business Performance: A Survey Study of Russian Managers*," Journal of Economic Psychology* 22(5): 661-692.

Ayton, P. and A.G.R. McClelland. 1997. "How Real is Overconfidence?," *Journal of Behavioral Decision Making* 10(3): 279 - 285.

Ayton, P. and G. Wright. 1994. "Subjective Probability: What Should We Believe," In G.Wright and P.Ayton, eds., *Subjective Probability*, Chichester: John Wiley & Sons.

Batchelor, R. and P. Dua. 1990. "Forecaster Ideology, Forecasting Technique, and the Accuracy of Economic Forecasts," *International Journal of Forecasting* 6: 3-10.

Benson, P.G. and D. Önkal. 1992. "The Effects of Feedback and Training on the Performance of Probability Forecasters," *International Journal of Forecasting* 8(4): 559-573.

Bolger, F. and N. Harvey. 1993. "Context-Sensitive Heuristics in Statistical Reasoning ," *The Quarterly Journal of Experimental Psychology* 46A: 779- 811.

Bolger, F. and N. Harvey. 1995. "Judging the Probability That the Next Point in an Observed Time-Series Will Be Below, or Above, a Given Value," *Journal of Forecasting* 14(7): 597- 607.

Bolger, F. and D. Önkal. 2004. "The Effects of Feedback on Judgmental Interval Predictions," *International Journal of Forecasting* 20: 29 – 39.

Bolger, F. and G. Wright. 1994. "Assessing the Quality of Expert Judgment," *Decision Support Systems* 11(1): 1- 24.

Bunn, D.W. 1996. "Non-traditional Methods of Forecasting," *European Journal of Operational Research* 92(3): 528-536.

Bunn, D.W. and A.A. Salo. 1996. "Adjustment of Forecasts with Model Consistent Expectations," *International Journal of Forecasting* 12(1): 163-170.

Bunn, D.W. and G. Wright. 1991. "Interaction of Judgmental and Statistical Forecasting Methods: Issues and Analysis," *Management Science* 37: 501-518.

Burns, A.C. and R.F. Bush. 2000. *Marketing Research, Third Ed.,* USA: Prentice Hall.

Carbone, R., and S.J. Armstrong. 1982. "Evaluation of Extrapolative Forecasting Methods: Results of a Survey of Academicians and Practitioners," *Journal of Forecasting* 1: 215-217.

Carbone, R., A. Andersen, Y. Corriveau, P.P. Corson. 1983. "Comparing the Different Time Series Methods: The Value of Technical Expertise, Individualized Analysis and Judgmental Adjustment," *Management Science* 29: 559-566.

Carbone, R. and W. Gorr. 1985. "Accuracy of Judgmental Forecasting of Time Series," *Decision Sciences* 16: 153-160.

Chatfield, C. 2001. "Prediction Intervals for Time-Series Forecasting," In S.Armstrong, ed., *Principles of Forecasting*. Mass:Kluwer Academic Publishers.

Clements, M.P. 1995. "Rationality and the Role of Judgement in Macroeconomic Forecasting," *The Economic Journal* 105: 410-420.

Cochran, W.G. and G.M. Cox. 1957. *Experimental Designs*. New York: Wiley.

Dalrymple, D.J. 1987. "Sales Forecasting Practices: Results from a United States Survey," *International Journal of Forecasting* 3: 379-392.

Davis, F.D., G.L. Lohse and J.E. Kottemann. 1994. "Harmful Effects of Seemingly Helpful Information on Forecasts of Stock Earnings," *Journal of Economic Psychology* 15(2): 253-267.

Davis, F. and J. Kottemann. 1995. "Determinants of Decision Rule Use in a Production Planning Task," *Organizational Behavior and Human Decision Processes* 63: 145-157.

Deming, W.E. 1950. *Some Theory of Sampling*, New York: Wiley.

Dhaliwal, J.S. and I. Benbasat. 1996. "The Use and Effects of Knowledge-Based System Explanations: Theoretical Foundations and a Framework for Empirical Evaluation," *Information Systems Research* 7(3): 342-362.

Diamantopoulos, A. and B. Mathews. 1989. "Factors Affecting the Nature and Effectiveness of Subjective Revision in Sales Forecasting: An Empirical Study," *Managerial and Decision Economics* 10: 51-59.

Dickson, G.W., G. DeSanctis and D.J. McBride. 1986. "Understanding the Effectiveness of Computer Graphics for Decision Support: A Cumulative Experimental Approach," *Communications of the ACM* 29: 40–47.

DiFonzo, N. 1997. "Cueing Rumor and Prediction: Making Sense (but Losing Dollars) in the Stock Market," *Organizational Behavior and Human Decision Processes* 71(3): 329-353.

Donihue, M.R. 1993. "Evaluating the Role Judgment Plays in Forecast Accuracy," *Journal of Forecasting* 12: 81–92.

Edmundson, R., M. Lawrence and M. O'Connor. 1988. "The Use of Non-Time Series Information in Sales Forecasting: A Case Study," *Journal of Forecasting* 7: 201-211.

Eggleton, I.R.C. 1982. "Intuitive Time-Series Extrapolation," *Journal of Accounting Research* 20: 68-102.

Ehrman, C.M. and S.M. Shugan. 1995. "The Forecaster's Dilemma," *Marketing Science* 14(2): 123-147.

Fildes, R., P. Goodwin and M. Lawrence. 2005. "The Design Features of Forecasting Support Systems and Their Effectiveness," *Decision Support Systems* 12(1): 37– 53.

Fischhoff, B. 1988. "Judgmental Aspects of Forecasting : Needs and Possible Trends," *International Journal of Forecasting* 4: 331-339.

Fischhoff, B. 1994. "What Forecasts (Seem To) Mean," *International Journal of Forecasting* 10(3): 387-403.

Flores, B.E., D.L. Olson and C. Wolfe. 1992. "Judgmental Adjustment of Forecasts: A Comparison of Methods," *International Journal of Forecasting* 7(4): 421-433.

Frankfort-Nachmias, C. and D. Nachmias. 1996. *Research Methods in the Social Sciences, Fifth Ed.,* London: St. Martin's Press Inc.

Ghalia, M.B. and P.P. Wang. 2000. "Intelligent System to Support Judgmental Business Forecasting: The Case of Estimating Hotel Room Demand," *IEEE Transactions on Fuzzy Systems* 8(4): 380– 397.

Gilovich, T. 1991. *How We Know What Isn't So*, New York: Free Press.

Goodwin, P. 1998. "Enhancing Judgmental Sales Forecasting: The Role of Laboratory Research," In: G. Wright and P.Goodwin, eds., *Forecasting with Judgment*, Chichester: John Wiley & Sons.

Goodwin, P. 2000. "Improving the Voluntary Integration of Statistical Forecasts and Judgment," *International Journal of Forecasting* 16: 85–99.

Goodwin, P. 2004. "Providing Support for Decisions Based on Time Series Information under Conditions of Asymmetric Loss," *European Journal of Operational Research* article in press.

Goodwin, P., D. Önkal-Atay, M.E. Thomson, A.C. Pollock, A. Macaulay. 2004. "Feedback-Labelling Synergies in Judgmental Stock Price Forecasting," *Decision Support Systems* 37: 175-186.

Goodwin, P. and R. Fildes. 1999. "Judgmental Forecasts of Time Series Affected by Special Events: Does Providing a Statistical Forecast Improve Accuracy?," *Journal of Behavioral Decision Making* 12(1): 37– 53.

Goodwin, P. and G. Wright. 1993. "Improving Judgmental Time Series Forecasting: A Review of Guidance Provided by Research," *International Journal of Forecasting* 9: 147– 161.

Goodwin, P. and G. Wright. 1994. "Heuristics, Biases and Improvement Strategies in Judgmental Time Series Forecasting," *Omega* 22(6): 553– 568.

Granger, C.W.J. 1996. "Can We Improve the Perceived Quality of Economic Forecasts?," *Journal of Applied Econometrics* 11(5): 455– 473.

Gregor, S. 2001. "Explanations from Knowledge-Based Systems and Cooperative Problem Solving: An Empirical Study," *International Journal of Human-Computer Studies* 54: 81-105.

Gregor, S. and I. Benbasat. 1999. "Explanations from Intelligent Systems: Theoretical Foundations and Implications for Practice," *MIS Quarterly* 23(4): 497-530.

Gross, C. and R.T. Peterson. 1978. "Some Human Problems in Industrial Sales Forecasting," *Industrial Marketing Management* 7: 367– 368.

Harries, C. and N. Harvey. 2000. "Taking Advice, Using Information and Knowing What You Are Doing," *Acta Psychologica* 104: 399-416.

Harvey, N. 1995. "Why Are Judgments Less Consistent in Less Predictable Task Situations," *Organizational Behavior and Human Decision Processes* 63(3): 247-263.

Harvey, N., F. Bolger. and A.G.R. McClelland. 1994. "On the Nature of Expectations," *British Journal of Psychology* 85: 203-229.

Harvey, N. and F. Bolger. 1996. “Graphs versus Tables: Effects of Data Presentation Format on Judgemental Forecasting,” *International Journal of Forecasting* 12(1): 119-137.

Harvey, N. and I. Fischer. 1997. “Accepting Help, Improving Judgment, and Sharing Responsibility,” *Organizational Behavior and Human Decision Processes* 70(2): 117-133.

Harvey, N., C. Harries and I. Fischer. 2000. “Using Advice and Assessing Its Quality,” *Organizational Behavior and Human Decision Processes* 81(2): 252-273.

Harvey, N. and C. Harries. 2004. “Effects of Judges’ Forecasting on Their Later Combination of Forecasts for the Same Outcomes,” *International Journal of Forecasting* 20: 391-409.

Houghton, K.A. and J.J.F. Hronsky. 1993. “The Sharing of Meaning between Accounting Students and Members of the Accounting Profession,” *Accounting and Finance*, November: 131-147.

Irandoust, H. 2002. “Attitudes for Achieving User Acceptance Explaining, Arguing, Critiquing,” Paper presented at the Conference on “7th International Command & Control Research and Technology Symposium”, held in Quebec City-Canada.

Johnson, W.B. 1982. “The Impact of Confidence Interval Information on Probability Judgements,” *Accounting, Organizations and Society* 7(4): 349-367.

Jones, V.D., S. Bretschneider and W.L. Gorr. 1997. "Organizational Pressures on Forecast Evaluation: Managerial, Political, and Procedural Influences," *Journal of Forecasting* 16(4): 241-254.

Keren, G. 1991. "Calibration and Probability Judgments: Conceptual and Methodological Issues" *Acta Psychologica* 77: 217-273.

Keren, G. 1997. "On the Calibration of Probability Judgments: Some Critical Comments and Alternative Perspectives," *Journal of Behavioral Decision Making* 10(3): 269 - 278.

Kleinmuntz, B. 1990. "Why We Still Use Our Heads Instead of Formulas: Toward an Integrative Approach," *Psychological Bulletin.* 107: 296-310.

Koning, A.J., P.H. Franses, M. Hibon, H.O. Stekler. 2005. "The M3 Competition: Statistical Tests of the Results," *International Journal of Forecasting* 21: 397–409.

Lawrence, M., R.H. Edmundson and M. O'Connor. 1985. "An examination of the accuracy of judgmental extrapolation of time series," *International Journal of Forecasting* 1: 25–36.

Lawrence, M., L. Davies, M. O'Connor and P. Goodwin. 2001. "Improving Forecast Utilization by Providing Explanations," Paper presented at the Conference on "21st International Symposium on Forecasting," held in Atlanta-USA.

Lawrence, M., P. Goodwin, M. O'Connor and D. Önkal. 2006. "Judgmental Forecasting: A Review of Progress Over the Last 25 Years," *International Journal of Forecasting* 22: 493–518.

Lawrence, M. and S. Makridakis. 1989. "Factors Affecting Judgmental Forecasts and Confidence Intervals," *Organizational Behavior and Human Decision Processes* 43: 172–187.

Lawrence, M. and M. O'Connor. 1992. "Exploring Judgemental Forecasting," *International Journal of Forecasting* 8(1): 15–26.

Lawrence, M. and M. O'Connor. 1993. "Scale, Variability, and the Calibration of Judgmental Prediction Intervals," *Organizational Behavior and Human Decision Processes* 56(3): 441–458.

Lawrence, M. and M. O'Connor. 1995. "The Anchor and Adjustment Heuristic in Time-Series Forecasting," *Journal of Forecasting* 14(5): 443–451.

Lawrence, M. and M. O'Connor. 2005. "Judgmental Forecasting in the Presence of Loss Functions," *International Journal of Forecasting* 21(1): 3-14.

Lee, J.K., S.B. Oh and J.C. Shin. 1990. "UNIK-FCST: Knowledge-Assisted Adjustment of Statistical Forecasts", *Expert Systems with Applications* 1(1): 39–49.

Lee, J.K. and C.S. Yum. 1998. "Judgmental Adjustment in Time Series Forecasting Using Neural Networks", *Decision Support Systems* 22: 135–154.

Lichtenstein S., B. Fischhoff and L.D. Philips. 1982. "Calibration of Probabilities: The State of the Art to 1980," In D.Kahneman, P.Slovic and A.Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Lim, J. S. and M. O'Connor. 1995 "Judgmental Adjustment of Initial Forecasts - Its Effectiveness and Biases", *Journal of Behavioral Decision* Making 8(3): 149–168.

Lim, J. S. and M. O'Connor. 1996a "Judgmental Forecasting with Time Series and Causal Information", *International Journal of Forecasting* 12(1): 139–153.

Lim, J. S. and M. O'Connor. 1996b "Judgmental Forecasting with Interactive Forecasting Support Systems", *Decision Support Systems* 16(4): 339–357.

Locke E.A. 1986. *Generalizing from Laboratory to Field Settings*, Lexington Books.

Lories, G., M. Dubois and J. Gaussin. 1997. "Judgmental Forecasting and Anticipation in Human Process Control," *Travail Humain* 60(1): 87–101.

Mahmoud, E., G. Rice and N. Malhotra. 1988. "Emerging Issues in Sales Forecasting and Decision Support Systems," *Journal of the Academy of Marketing Science* 16: 47-61.

Mahmoud, E., R. DeRoeck, R.G. Brown, G. Rice. 1992. "Bridging the Gap between Theory and Practice in Forecasting," *International Journal of Forecasting* 8(2): 251-267

Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen and R. Winkler. 1982. "The Accuracy of Extrapolation Methods: Results of a Forecasting Competition," *Journal of Forecasting* 1(2): 111–153.

Makridakis, S., M. Hibon, E. Lusk and M. Belhadjali. 1987. "Confidence intervals: An Empirical Investigation of the Series in the M-Competition," *International Journal of Forecasting* 3: 489–508.

Makridakis, S., C. Chatfield, M. Hibon, M. Lawrence, T. Mills, K. Ord, F.L. Simmons. 1993. "The M2-Competition - A Real-Time Judgmentally Based Forecasting Study," *International Journal of Forecasting* 9(1): 5–22.

Makridakis, S. and M. Hibon. 2000. "The M3-Competition: Results, Conclusions and Implications," *International Journal of Forecasting* 16: 451–476.

Makridakis, S., S.C. Wheelwright and R.J. Hyndman. 1998. *Forecasting: Methods and Applications* (Third Edition) USA: John Wiley & Sons Inc.

Mao, J.Y. and I. Benbasat. 2000. "The Use of Explanations in Knowledge-Based Systems: Cognitive Perspectives and a Process-Tracing Analysis," *Journal of Management Information Systems* 17(2): 153-179.

Mathews, B.P. and A. Diamantopoulos. 1986. "Managerial Intervention in Forecasting. An Empirical Investigation of Forecast Manipulation," *International Journal of Research in Marketing* 3: 3-10.

Mathews, B.P. and A. Diamantopoulos. 1989. "Judgemental Revision of Sales Forecasts: A Longitudinal Extension," *Journal of Forecasting* 8: 129-140.

Mathews, B.P. and A. Diamantopoulos. 1990. “Judgemental Revision of Sales Forecasts: Effectiveness of Forecast Selection,” *Journal of Forecasting* 9: 407-415.

McClelland, A.G.R. and F. Bolger. 1994. "The Calibration of Subjective Probabilities: Theories and Models 1980-94," In G.Wright and P.Ayton, eds., *Subjective Probability*, Chichester: John Wiley & Sons.

McNees, S.K. 1990. “The Role of Judgment in Macroeconomic Forecasting Accuracy,” *International Journal of Forecasting* 6(3): 287-299

Mentzer, J.T. and J.E. Cox. 1984. “Familiarity, Application of Sales Forecasting Techniques,” *Journal of Forecasting* 3: 27-36.

Montgomery, D.C. 2001. *Design and Analysis of Experiments, Fifth Ed.,* USA: John Wiley & Sons Inc.

Mulligan, E.J. and R. Hastie. 2005. “Explanations Determine the Impact of Information on Financial Investment Judgments,” *Journal of Behavioral Decision Making* 18: 145– 156.

Murphy, A.H. 1993. “What is A Good Forecast – An Essay on the Nature of Goodness in Weather Forecasting,” *Weather and Forecasting* 8(2): 281-293.

Murphy, A.H. and D.S. Wilks. 1998. “A Case Study of the Use of Statistical Models in Forecast Verification: Precipitation Probability Forecasts,” *Weather and Forecasting* 13(3): 795-810.

O’Connor, M. and M. Lawrence. 1989. “An Examination of the Accuracy of Judgmental Confidence Intervals in Time Series Forecasting,” *International Journal of Forecasting* 8: 141–155.

O’Connor, M. and M. Lawrence. 1992. “Time Series Characteristics and the Widths of Judgemental Confidence Intervals,” *International Journal of Forecasting* 7(4): 413–420.

O'Connor, M., W. Remus and K. Griggs. 1993. "Judgmental Forecasting in Times of Change," *International Journal of Forecasting* 9(2): 163–172.

O'Connor, M., W. Remus and K. Griggs. 1997. "Going up Going down: How Good Are People at Forecasting Trends and Changes in Trends?," *Journal of Forecasting* 16(3): 165–176.

O'Connor, M., W. Remus and K. Griggs. 2001. "The Asymmetry of Judgmental Confidence Intervals in Time Series Forecasting," *International Journal of Forecasting* 17: 623–633.

O'Connor, M., W. Remus and K. Lim. 2005. "Improving Judgmental Forecasts with Judgmental Bootstrapping and Task Feedback Support" *Journal of Behavioral Decision Making* 18: 247–260.

Önkal-Atay, D. 1998. "Financial Forecasting with Judgment" In G.Wright and P. Goodwin, eds., *Forecasting with Judgment*. Chichester: John Wiley & Sons.

Önkal, D. and F. Bolger. 2004. "Provider–User Differences in Perceived Usefulness of Forecasting Formats," *Omega* 32: 31 – 39.

Önkal, D., J.F. Yates, C. Şimga-Muğan and Ş. Öztin. 2003. "Professional vs. Amateur Judgment Accuracy: The Case of Foreign Exchange Rates," *Organizational Behavior and Human Decision Processes* 91: 169–185.

Önkal, D. and M.S. Gönül. 2005. "Judgmental Adjustment: A Challenge for Providers and Users of Forecasts," *Foresight: The International Journal of Applied Forecasting* 1(1): 13-17.

Önkal, D. and G. Muradoğlu. 1994. "Evaluating Probabilistic Forecasts of Stock Prices in a Developing Stock Market," *European Journal of Operational Research* 74(2): 350-358.

Önkal, D. and G. Muradoğlu. 1995. “Effects of Feedback on Probabilistic Forecasts of Stock Prices,” *International Journal of Forecasting* 11(2): 307–319.

Önkal, D. and G. Muradoğlu. 1996. “Effects of Task Format on Probabilistic Forecasting of Stock Prices,” *International Journal of Forecasting* 12(1): 9–24.

Önkal-Atay, D., M.E. Thomson, and A.C. Pollock. 2002. "Judgmental Forecasting" In M.P.Clements and D.Hendry, eds., *A Companion to Economic Forecasting*. Blackwell Publ, 133151.

Papamichail, K.N. and S. French. 2003. “Explaining and Justifying the Advice of a Decision Support System: A Natural Language Generation Approach,” *Expert Systems with Applications* 24: 35-48.

Pons-Novell, J. 2003. “Strategic Bias, Herding Behaviour and Economic Forecasts,” Journal of Forecasting 22: 67-77.

Price P.C. and E. R. Stone. 2004. "Intuitive Evaluation of Likelihood Judgment Producers: Evidence for a Confidence Heuristic," *Journal of Behavioral Decision Making* 17(1): 39 - 57.

Remus, W. 1996. ”Will Behavioral Research on Managerial Decision Making Generalize to Managers?,” *Managerial and Decision Economics* 17: 93-101.

Remus, W., M. O’Connor and K. Griggs. 1995. “Does Reliable Information Improve the Accuracy of Judgmental Forecasts?,” *International Journal of Forecasting* 11(2): 285–293.

Remus, W., M. O’Connor and K. Griggs. 1996. “Does Feedback Improve the Accuracy of Recurrent Judgmental Forecasts,” *Organizational Behavior and Human Decision Processes* 66(1): 22-30.

Remus, W., M. O’Connor and K. Griggs. 1998. “The Impact of Information of Unknown Correctness on the Judgmental Forecasting Process,” *International Journal of Forecasting* 14: 313–322.

Sanders, N.R. 1992. “Accuracy of Judgmental Forecasts: A Comparison,” *Omega*: *The International Journal of Management* 20: 353–364.

Sanders, N.R. 1997. “The Impact of Task Properties Feedback on Time Series Judgmental Forecasting Tasks,” *Omega* 25(2): 135–144.

Sanders, N.R. and K.B. Manrodt. 1994. “Forecasting Practices in U.S. Corporations: Survey Results,” *Interfaces* 24: 92– 100.

Sanders, N.R. and K.B. Manrodt. 2003. “The Efficacy of Using Judgmental versus Quantitative Forecasting Methods in Practice,” *Omega* 31: 511–522.

Sanders, N.R. and L.P. Ritzman. 1992. “The Need for Contextual and Technical Knowledge in Judgmental Forecasting” *Journal of Behavioral Decision Making* 5: 39-52.

Sanders, N.R. and L.P. Ritzman. 2001. "Judgmental Adjustment of Statistical Forecasts" In J. S. Armstrong, ed., *Principles of Forecasting: A Handbook for Researchers and Practitoners*. Mass:Kluwer Academic Publishers.

Schrah, G.E., R.S. Dalal and J.A. Sniezek. 2006. “No Decision-Maker is an Island: Integrating Expert Advice with Information Acquisition” *Journal of Behavioral Decision Making* 5: 39-52.

Shapiro, D.L., E.H. Buttner and B. Barry. 1991. “Explanations: When Are They Judged Adequate?,” *Academy of Management Proceedings*: 395-399.

Shapiro, D.L., E.H. Buttner and B. Barry. 1994. “Explanations: What Factors Enhance Their Perceived Adequacy?,” *Organizational Behavior and Human Decision Processes* 58: 346-368.

Shortliffe, E.H. 1976. *Computer-Based Medical Consultations: MYCIN*. New York: Elsevier Computer Science Library.

Sniezek, J.A. and T. Buckley. 1995. “Cueing and Cognitive Conflict in Judge-Advisor Decision Making,” *Organizational Behavior and Human Decision Processes* 62(2): 159-174.

Sniezek, J.A., G.E. Schrah and R.S. Dalal. 2004. “Improving Judgement with Prepaid Expert Advice” *Journal of Behavioral Decision Making* 17: 173-190.

Sniezek, J.A. and L.M. Swol. 2001. “Trust, Confidence, and Expertise in a Judge-Advisor System,” *Organizational Behavior and Human Decision Processes* 84(2): 288-307.

Sparkes, J.R. and A.K. McHugh. 1984. “Awareness and Use of Forecasting Techniques in British Industry,” *Journal of Forecasting* 3: 37–42.

Swol, L.M. and J.A. Sniezek. 2005. “Factors Affecting the Acceptance of Expert Advice,” *British Journal of Social Psychology* 44: 443-461.

Thomson, M.E., D. Önkal-Atay, A.C. Pollock and A. Macaulay. 2003. “The Influence of Trend Strength on Directional Probabilistic Currency Predictions,” *International Journal of Forecasting* 19: 241–256.

Turban, E. and J.E. Aronson. 2001. *Decision Support Systems and Intelligent Systems 6th ed.* Upper Saddle River, NJ: Prentice Hall.

Turner, D.S. 1990. “The Role of Judgement in Macroeconomic Forecasting,” *Journal of Forecasting* 9: 315–345.

Tversky, A. and D. Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases," *Science* 185: 1124-1131.

Tversky, A. and D. Kahneman. 1982. "Judgment under Uncertainty: Heuristics and Biases", In D.Kahneman, P.Slovic and A.Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases* Cambridge: Cambridge University Press.

Wacker, J.G. and R.R. Lummus. 2002. "Sales Forecasting for Strategic Resource Planning," *International Journal of Operations & Production Management* 22(9-10): 1014-1031.

Wagenaar, W.A. and S.D. Sagaria. 1975. " Misperception of Exponential Growth," *Perception and Psychophysics* 18: 416-422.

Webby, R. and M. O'Connor. 1996. "Judgemental and Statistical Time Series Forecasting: A Review of the Literature," *International Journal of Forecasting* 12(1): 91-118.

Webby, R., M. O'Connor and B. Edmundson. 2005. "Forecasting Support Systems for the Incorporation of Event Information: An Empirical Investigation," *International Journal of Forecasting* 21: 411-423.

Webby, R., M. O'Connor and M. Lawrence. 2001. "Judgmental Time-Series Forecasting Using Domain Knowledge" In J. S. Armstrong, ed., *Principles of Forecasting: A Handbook for Researchers and Practitoners*. Mass:Kluwer Academic Publishers.

Wheelwright, S.C. and D.G. Clarke. 1976. "Corporate Forecasting: Promise and Reality," *Harvard Business Review* 54: 40–47.

Whitecotton, S.M. 1996. "The Effects of Experience and a Decision Aid on the Slope, Scatter, and Bias of Earnings Forecasts," *Organizational Behavior and Human Decision Processes* 66(1): 111-121.

Wilkie-Thomson, M.E., D. Önkal-Atay and A.C. Pollock. 1997. "Currency Forecasting: An Investigation of Extrapolative Judgement," *International Journal of Forecasting* 13(4): 509-526.

Willemain, T.R. 1989. "Graphical Adjustment of Statistical Forecasts," *International Journal of Forecasting* 5: 179-185.

Willemain, T.R. 1991. "The Effect of Graphical Adjustment on Forecast Accuracy," *International Journal of Forecasting* 7(2): 151-154.

Winklhofer, H., A. Diamantopoulos and S.F. Witt. 1996. “Forecasting Practice: A Review of the Empirical Literature and an Agenda for Future Research,” *International Journal of Forecasting* 12: 193-221.

Winklhofer, H. and A. Diamantopoulos. 2002. “Managerial Evaluation of Sales Forecasting Effectiveness: A MIMIC Modeling Approach,” *International Journal of Research in Marketing* 19(2): 151-166.

Wolfe, C. and B. Flores. 1990. “Judgmental Adjustment of Earnings Forecasts,” *Journal of Forecasting* 9: 389-405.

Yaniv, I. 2004a. “Receiving Other People’s Advice: Influence and Benefit,” *Organizational Behavior and Human Decision Processes* 93: 1-13.

Yaniv, I. 2004b. “The Benefit of Additional Opinions,” *Current Directions in Psychological Science* 13(2): 75-78.

Yaniv, I. and D.P. Foster. 1995. "Graininess of Judgment Under Uncertainty: An Accuracy-Informativeness Trade-Off", *Journal of Experimental Psychology: General* 124(4): 424-432.

Yaniv, I. and D.P. Foster. 1997. "Precision and Accuracy of Judgmental Estimation", *Journal of Behavioral Decision Making* 10: 21-32.

Yaniv, I. and E. Kleinberger. 2000. “Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation,” *Organizational Behavior and Human Decision Processes* 83(2): 260-281.

Yates, J.F., L.S. McDaniel and E.S. Brown. 1991. "Probabilistic Forecasts of Stock Prices and Earnings: The Hazards of Nascent Expertise," *Organizational Behavior and Human Decision Processes* 49: 60-79.

Yates, J.F., P.C. Price, J.W. Lee and J. Ramirez. 1996. "Good Probabilistic Forecasters: The `Consumer's' Perspective," *International Journal of Forecasting* 12: 41-56.

Ye, L.R. and P.E. Johnson. 1995. “The Impact of Explanation Facilities on User Acceptance of Expert Systems Advice,” *MIS Quarterly* 19(2): 157-172.

Yokum J.T. and J.S. Armstrong. 1995. “Beyond Accuracy: Comparison of Criteria Used to Select Forecasting Methods,” *International Journal of Forecasting* 11: 591-597.

Young, R.M. 1982. “Forecasting with an Econometric Model: The Issue of Judgemental Adjustment,” *Journal of Forecasting* 1: 189–204.

APPENDICES

APPENDIX A
The Questionnaire Form

Dear participant,
This survey is a part of a research conducted in **Bilkent University Faculty of Business Administration** for the aim of understanding the use of financial forecasts, expectations from those forecasts and the adjustments made on those forecasts. There are no right or wrong answers. The important point is answering the questions according to what you really think or believe. After reading carefully, please attend to all questions in a way that will best reflect your observations and thoughts. No personal information and names will be disclosed. The results will be used in a generalized manner and only for research purposes. Your time and effort for completing this survey is extremely valuable for us.

Your name: .....

Phone number: .....

1- What is the name of the company you are working for?
.....

2- What is your position?
.....

3- How long have you been in this position?
☐ Less than 1 year ☐ 1-3 years ☐ 4-6 years ☐ 7-9 years ☐ More than 9 years

4- Do you use financial forecasts in your company?
☐ Yes
☐ No (PLEASE ANSWER QUESTIONS 7, 8 & 9 AND THEN KINDLY SKIP TO QUESTION 32.)

5- What are the types of financial forecasts that are used in your company? (You may select more than one.)
☐ Currency/parity ☐ Stock price ☐ Investment Funds ☐ Interest/bond ☐ Macroeconomic
☐ Other .....

6- How long have you been using financial forecasts?
☐ Less than 1 year ☐ 1-3 years ☐ 4-6 years ☐ 7-9 years ☐ More than 9 years

7- What are your expectations from financial forecasts? (What do you expect them to have and how should they be?)
.....
.....
.....
.....
.....

8- Some **expectations** from financial forecasts are provided below. Please indicate the **degree of your agreement** by giving an appropriate value between 1-7.

a) A forecast should have the smallest amount of error possible.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

b) A forecast should have a tolerable (not necessarily minimal) amount of error.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

c) A forecast should have a plausible and justifiable basis and assumptions.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

d) A forecast should include scenarios and alternative forecasts covering a variety of circumstances.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

e) A forecast should have a reasonable cost.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

f) A forecast should be timely.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

9-As long as a firm reaches its targets, some error in forecasts is tolerable.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

10- Please select the most suitable answer.

☐ The financial forecasts that we use are produced in **our department**.

☐ The financial forecasts that we use are **acquired externally** (from banks, specialized firms or other departments in the company).

IF THE MAJORITY OF FORECASTS ARE PRODUCED IN YOUR DEPARTMENT PLEASE PROVIDE ANSWERS TO THE FOLLOWING QUESTIONS BY CONSIDERING THE CASES YOU HAVE ACQUIRED EXTERNAL FORECASTS.

11- Do you acquire forecasts from different sources?

☐ Yes. (We acquire forecasts from more than one source.)

☐ No. (We acquire forecasts from a single source only.) (PLEASE SKIP TO QUESTION 13.)

12- When you acquire forecasts from more than one source, which of the following do you usually employ?

☐ I choose a single source among them and use only that source.

☐ The forecasts from these sources are distinct and are not suitable for combination.

☐ I combine various forecasts and use the combination. I utilize a statistical method (simple average, weighted average etc.) for combination.

☐ I combine various forecasts and use the combination. I do not utilize a statistical method, instead I combine according to my judgment and experience.

13- Do externally acquired forecasts include scenarios and alternative forecasts?

Almost never 1 2 3 4 5 6 7 Almost always

14- Do externally acquired forecasts include explanations about forecasts?

Almost never 1 2 3 4 5 6 7 Almost always

15- Do you apply adjustments/revisions on the externally acquired forecasts?

Almost never 1 2 3 4 5 6 7 Almost always

16- What is the most important reason for you when you choose **not to adjust** an acquired forecast?

.....
.....
.....
.....
.....

17- What is the most important reason for you to **apply an adjustment** on an acquired forecast?

.....
.....
.....
.....
.....

18- Some situations where externally acquired forecasts are **not adjusted** are provided below. Please indicate how **frequently** they occur by giving each statement an appropriate value between 1-7.

a) My knowledge and experience on the subject is not adequate to make an adjustment.

Very rarely 1 2 3 4 5 6 7 Very frequently

b) The source providing the forecasts is well-known and famous.

Very rarely 1 2 3 4 5 6 7 Very frequently

c) I believe that the source providing the forecasts is unbiased and objective.

Very rarely 1 2 3 4 5 6 7 Very frequently

d) The methods and analysis used in the acquired forecasts are highly persuasive.

Very rarely 1 2 3 4 5 6 7 Very frequently

e) The presentation and the style of language used in the acquired forecasts are highly persuasive.

Very rarely 1 2 3 4 5 6 7 Very frequently

f) The explanations provided with the acquired forecasts are highly persuasive.

Very rarely 1 2 3 4 5 6 7 Very frequently

g) I do not want to be responsible for the consequences of the adjustment.

Very rarely 1 2 3 4 5 6 7 Very frequently

h) I'm not authorized to make an adjustment.

Very rarely 1 2 3 4 5 6 7 Very frequently

19- Some situations where externally acquired forecasts are **not adjusted** are provided below. Please indicate how **important** they are by giving each statement an appropriate value between 1-7.

a) My knowledge and experience on the subject is not adequate to make an adjustment.

Not important at all 1 2 3 4 5 6 7 Very important

b) The source providing the forecasts is well-known and famous.

Not important at all 1 2 3 4 5 6 7 Very important

c) I believe that the source providing the forecasts is unbiased and objective.

Not important at all 1 2 3 4 5 6 7 Very important

d) The methods and analysis used in the acquired forecasts are highly persuasive.

Not important at all 1 2 3 4 5 6 7 Very important

e) The presentation and the style of language used in the acquired forecasts are highly persuasive.

Not important at all 1 2 3 4 5 6 7 Very important

f) The explanations provided with the acquired forecasts are highly persuasive.

Not important at all 1 2 3 4 5 6 7 Very important

g) I do not want to be responsible for the consequences of the adjustment.

Not important at all 1 2 3 4 5 6 7 Very important

h) I'm not authorized to make an adjustment.

Not important at all 1 2 3 4 5 6 7 Very important

20- Some situations where externally acquired forecasts are **adjusted** are provided below. Please indicate how **frequently** they occur by giving each statement an appropriate value between 1-7.

a) To integrate my knowledge, experience and initiative into the forecasts.

Very rarely 1 2 3 4 5 6 7 Very frequently

b) To reflect the unexpected events and new information into the forecasts.

Very rarely 1 2 3 4 5 6 7 Very frequently

c) The source providing the forecasts is small and barely-known.

Very rarely 1 2 3 4 5 6 7 Very frequently

d) I believe that the source providing the forecasts is biased and leading.

Very rarely 1 2 3 4 5 6 7 Very frequently

e) There are extreme forecasts present.

Very rarely 1 2 3 4 5 6 7 Very frequently

f) To gain control of and take responsibility for externally acquired forecasts.

Very rarely 1 2 3 4 5 6 7 Very frequently

g) The methods and analysis used in the acquired forecasts are not persuasive enough.

Very rarely 1 2 3 4 5 6 7 Very frequently

h) The presentation and the style of language used in the acquired forecasts are not persuasive enough.

Very rarely 1 2 3 4 5 6 7 Very frequently

i) The explanations provided with the acquired forecasts are not persuasive enough.

Very rarely 1 2 3 4 5 6 7 Very frequently

21- Some situations where externally acquired forecasts are **adjusted** are provided below. Please indicate how **important** they are by giving each statement an appropriate value between 1-7.

a) To integrate my knowledge, experience and initiative into the forecasts.

Not important at all 1 2 3 4 5 6 7 Very important

b) To reflect the unexpected events and new information into the forecasts.

Not important at all 1 2 3 4 5 6 7 Very important

c) The source providing the forecasts is small and barely-known.

Not important at all 1 2 3 4 5 6 7 Very important

d) I believe that the source providing the forecasts is biased and leading.

Not important at all 1 2 3 4 5 6 7 Very important

e) There are extreme forecasts present.

Not important at all 1 2 3 4 5 6 7 Very important

f) To gain control of and take responsibility for externally acquired forecasts.

Not important at all 1 2 3 4 5 6 7 Very important

g) The methods and analysis used in the acquired forecasts are not persuasive enough.

Not important at all 1 2 3 4 5 6 7 Very important

h) The presentation and the style of language used in the acquired forecasts are not persuasive enough.

Not important at all 1 2 3 4 5 6 7 Very important

i) The explanations provided with the acquired forecasts are not persuasive enough.

Not important at all 1 2 3 4 5 6 7 Very important

22- Adjustments made to the forecasts improve their accuracy.

Almost never 1 2 3 4 5 6 7 Almost always

23- Adjustments applied done to the forecasts make them more persuasive.

Almost never 1 2 3 4 5 6 7 Almost always

24- Does anyone check your forecasts after you have applied your adjustments?

☐ Yes
☐ No (PLEASE SKIP TO QUESTION 30.)

25- Does this person introduce further adjustments on those forecasts?

Almost never 1 2 3 4 5 6 7 Almost always

26- Does this person expect explanations about the adjustments you have made?

Almost never 1 2 3 4 5 6 7 Almost always

27- Does this person access the original (unadjusted) forecasts?

Almost never 1 2 3 4 5 6 7 Almost always

28- I believe that further adjustments made on already adjusted forecasts improve their accuracy.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

29- I believe that further adjustments applied on already adjusted forecasts make them more persuasive.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

30- Is there a feedback mechanism about the accuracy of forecasts in your company?

☐ Yes, there is a systematic feedback mechanism.
☐ No, there is no systematic feedback, however, some feedback applications are observed at personal level.
☐ No, there are neither systematic nor personal feedback applications in the firm. (PLEASE SKIP TO QUESTION 32.)

31- Do you believe that those feedback applications are useful? How?

.....
.....
.....
.....

32- According to you, what is the meaning of a “high-quality forecast”?

.....
.....
.....
.....

33- The only criterion for appraising the quality of a forecast should be accuracy.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

34- A forecast satisfying my expectations can be considered to be a high quality forecast.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

35- Some statements about the **quality** of forecasts are provided below. Please indicate the **degree of your agreement** by giving an appropriate value between 1-7.

a) A high quality forecast has a plausible and justifiable basis and assumptions.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

b) A high quality forecast includes scenarios and alternative forecasts covering a variety of circumstances.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

c) A high quality forecast has the smallest amount of error possible.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

d) A high quality forecast has a tolerable (not necessarily minimal) amount of error.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

e) A high quality forecast has a reasonable cost.

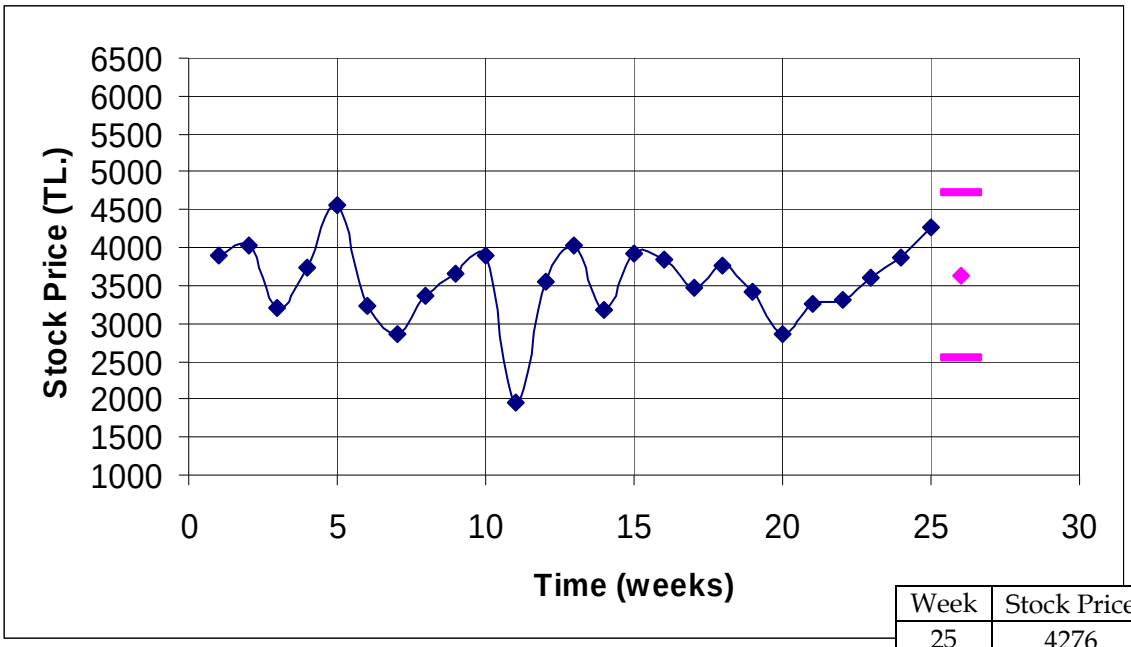
Strongly disagree 1 2 3 4 5 6 7 Strongly agree

f) A high quality forecast is timely.

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

APPENDIX B
Examples for the First Study

SAMPLE FORM GIVEN TO PARTICIPANTS FOR EACH TIME SERIES



Week	Stock Price
25	4276
24	3857
23	3597
22	3308

OUR FORECAST FOR WEEK 26:

Point Forecast : **3630**

95% Interval Forecast : **[2530 – 4730]**

OUR EXPLANATION :
New government subsidy is certainly responsible for the upward trend starting from week 20. We strongly believe that this is excessive and there will be a drop.

WOULD YOU LIKE TO MODIFY THE GIVEN POINT FORECAST? ☐ YES ☐ NO

IF YES,
Your point forecast : …….

WOULD YOU LIKE TO MODIFY THE GIVEN INTERVAL FORECAST? ☐ YES ☐ NO

IF YES,
Your 95% interval forecast : **[**…….. – …….**]**

SAMPLE EXPLANATIONS

Short Explanation, Strong Confidence
New government subsidy is certainly responsible for the upward trend starting from week 20. We strongly believe that this is excessive and there will be a drop.

Short Explanation, Weak Confidence
New government subsidy may have led to the upward trend starting from week 20. This seems to be excessive and there may be a drop.

Long Explanation, Strong Confidence
The government has decided to provide a new subsidy for the some of the company's products. This was announced in week 20, and it immediately caused the upward trend in the stock prices starting from that week. This upward trend can be distinctly observed from the time-series data. However, we find this increase to be very excessive. We, therefore, strongly believe that there will be a drop in the stock prices in the 26^{th} week.

Long Explanation, Weak Confidence
The government has decided to provide a new subsidy for the some of the company's products. This was announced in week 20, and it may have led to the upward trend in the stock prices starting from that week. This upward trend may be observed from the time-series data. However, this increase may appear to be excessive so that there may be a drop in the stock prices in the 26^{th} week.