MANAGERIAL JUDGMENT IN MARKETING: THE CONCEPT OF EXPERTISE

by

JEAN-CLAUDE LARRECHE AND REZA MOINPOUR

N° 82/19

Director of Publication :

Edith Penrose
Associate Dean: Research and Development
INSEAD

Printed by INSEAD, Fontainebleau
France

MANAGERIAL JUDGMENT IN MARKETING:

THE CONCEPT OF EXPERTISE

Jean-Claude Larréché

Reza Moinpour

revised

September 1982

and

December 1982

Jean-Claude Larréché is Professor of Marketing at INSEAD, Fontainebleau, France, and Reza Moinpour is Professor of Marketing at University of Washington, Seattle.

The authors wish to acknowledge the valuable comments of Robin Hogarth, Shelby McIntyre, Colleagues at INSEAD and Washington, and JMR reviewers.

Good for publication in Journal of Marketing Research.

MANAGERIAL JUDGMENT IN MARKETING: THE CONCEPT OF EXPERTISE

ABSTRACT

Judgment-based estimations are often used in marketing decision making and as inputs to decision calculus models. An evaluation of different approaches to integrate individual judgments is performed in the context of a marketing situation where the "true" values being estimated are known. Experts identified by a simple external measure of expertise are found to be more likely to provide better estimates, followed by groups participating in a Delphi process. Experts identified by self-rated confidence and interpersonal group consensus are not found likely to provide better estimates than the average of the initial judgments of the members of a group.

INTRODUCTION

Important efforts and progress have been made in market research to gather and analyze empirical data in order to better understand market phenomena and make more appropriate marketing decisions. Although market research is an essential part of marketing activity, it is unrealistic to assume that complete empirical data can be made available for most marketing decisions. More typically, such decisions are at least partially based on managerial judgments stemming from past experience and supplemented by market research data in specific areas where empirical information is perceived to be crucial. A multitude of reasons may explain this lack of complete availability of research data including time pressure, cost, or methodological difficulties in obtaining the desired data. Even when market research data is obtained, the decision to acquire these data is the result of a managerial judgment based on past experience and the potential contribution of the data to decision-making relative to its cost.

The analysis of managerial judgments has thus been increasingly recognized as an important area in marketing research (Best, 1974; Jolson and Rossow, 1971). This growing interest in managerial judgment has been fostered by the development of decision calculus models (Little, 1970) which explicitly consider such judgments as inputs in addition to empirical data (Chakravarti, et al., 1979, 1981; McIntyre, 1982).

There are two main issues which may be raised with regard to managerial judgments. The first one concerns the <u>validity</u> of these judgments in replicating or predicting observable phenomena and the second one concerns the methods to elicit better judgments. In psychology, a number of recent studies on clinical judgment and decision making have indicated that intuitive judgments often have a low predictive validity (Einhorn and Hogarth, 1978). In marketing, Chakravarti, Mitchell and Staelin (1979) have found that managerial judgments are subject to biases in estimating the parameters of a decision-calculus model, especially in areas where managers lack experience. They have further noted that if the decision environment is complex, managers may have more difficulty providing the necessary judgmental estimates (Chakravarti, Mitchell and Staelin, 1981). By contrast, McIntyre (1980) has shown that managers can make fairly accurate judgments in stable and simple decision environments. However, as Little and Lodish (1981) have observed, inadequate laboratory setting and limited access to information may have affected some of Chakravarti, et al's findings with regard to judgment-based models. Since it appears that the quality of judgment depends on the experience of managers with a specific situation, the evaluation of this kind of decision should be performed in appropriate field experiments; these tests, however, are often difficult to implement in practice. In the only controlled field experiment which appears to have been published on the subject, Fudge and Lodish (1977) have shown a significant improvement in the performance of a salesforce with judgments as inputs to a decision-calculus model.

In many situations, judgmental decisions have to be made on the basis of little or no empirical data within the prevailing time or budget constraints. The Bayesian approach to statistical inference and decision-making can be useful but it is difficult to apply. The most important practical problem is that of the assessment of the required prior distributions (Winkler, 1967). Furthermore, there is little doubt that when valid empirical information is available, it should be used by those making decisions. However, when data or objective evidence are not available, judgments constitute the only available source of information. In this context, the relevant research issue is to select the most appropriate method for eliciting judgment in order to provide the best estimate in a given situation. Although differences in estimates by various methods have been reported (see, for example, Gustafson, et al., 1973) the available means to improve them have not been fully discussed.

The purpose of the present research is to empirically investigate this particular issue. More specifically, we are interested in comparing several approaches for obtaining efficient judgmental estimates, including methods based on the selection of individual judgments as well as methods based on the integration of multiple judgments. The estimates obtained from these methods will be evaluated on the basis of known "true" values in order to compare the relative performance of the various approaches considered.

METHODS TO ELICIT JUDGMENT

The issue of human judgment in prediction has received considerable attention in both the psychological and managerial literature (see, e.g., Hogarth and Makridakis, 1981b for an extensive review). The use of various decision rules in complex (business) situations has also been discussed within the general context of the costs and values of managerial decisions (Beach and Mitchell, 1978). This has raised many conceptual and methodological questions concerning both the concept of expertise and methods to elicit judgment.

The Concept of Expertise

An important aspect of investigation of judgmental ability is the concept of expertise and the role of expert judgment in prediction. Most research in this area has raised questions with regard to the validity and accuracy of the predictions of experts. Moreover, it has generally been shown that human judgments add little to prediction by statistical analysis (Dawes, 1971). In a review of expert judgment, Camerer (1980) has examined various approaches that have been developed for improving predictions. Evidence supporting the superiority of models over expert prediction usually fall in two categories: a) least-squares regression and b) equal weighting of variables.

Linear regression. This method uses past outcomes thus eliminating the need for expert prediction. Past decision of experts are fitted through least-squares regression to build models for predicting future cases. It has been argued that substituting models for experts (which amounts to "bootstrapping" or "pulling-up" judges' decisions) can more consistently improve future predictions (Goldberg, 1970; Dawes, 1971). While that may be the case, it should be noted that the model is in fact based on the past decisions of experts. Therefore, it would be difficult indeed to ignore or minimize the role of the "expert" in the process.

Equal weighting of variables. This simple method of combining information has been found to be more appropriate with small samples (Dawes and Corrigan, 1974). It has been shown that equal weighting of variables can perform (predictive tasks) as well as ordinary least-squares weighting (Einhorn and Hogarth, 1975). However, merely suggesting a unit weighting scheme for decision making avoids the question of variable selection which ultimately determines performance. In this regard, the ability of the "expert" in selecting and defining variables of interest is most critical.

Although simple models may exhibit greater reliability than experts, their predictive performance should be regarded with some degree of caution. Their findings could be affected by the interpretation of relevant tasks and terms of reference. Clearly, for specific kinds of tasks and in appropriate situations, the existence of expertise cannot be rejected. Moreover, experts can often provide valuable benefits (funds of experience) beyond simple model prediction. Finally, adherence to mechanistic approaches toward prediction (which comes from the naive application of tools like linear regression) could ignore the role of expert judgment in guiding the process.

Despite the questions which have been raised concerning the concept of expertise and relative merits of human judgments compared to analytical models, a reliance on human judgment is often a necessity. The key problem is then to use methods which can elicit better judgments. Two types of approaches can be distinguished: integration methods and identification methods.

Integration Methods

These are based on integration of judgments of the individuals available. The traditional models for integrating individual judgments are: a) obtaining the concensus of interpersonal groups, b) collecting estimates from the Delphi process, and c) weighting (equally or differentially) the judgments of individuals in a group.

The Delphi procedure, developed by the Rand Corporation (Dalkey and Helmer, 1963; Dalkey, 1969), is based on an anonymous iterative process. Judgments are collected by anonymous questionnaires from several subjects who subsequently are asked to revise their judgments based on feedback they receive on the central tendency and dispersion measures of the group judgments. After several such iterations, the central tendency measure of the group judgments is expected to approach the true value of the entity being estimated. Generally, the Delphi process has been found to provide better estimates than the interpersonal group concensus as it retains the benefit of many judgments while removing the biasing effects of face-to-face interaction (Dalkey, 1969). This approach has been used extensively in managerial decision making (Milkovich, et al., 1972; Basu and Shroeder, 1977). In marketing decision making, the Delphi technique has been successfully used in estimation of demand (Jolson and Rossow, 1971; Best, 1974).

In addition to interpersonal and Delphi techniques other methods to integrate individual judgments have been considered in the psychological literature. These include: randomly picking an individual from a group, taking the average of individual judgments, and assigning different weights to different individual judgments according to various criteria.

Research has generally shown that simple schemes can be as accurate as more complex ones. The relative predictive efficiency of naive models has been demonstrated in investigations of quality of group judgment (Einhorn, et al., 1977) and forecasting accuracy (Makridakis and Hibon, 1979). More recently, Hogarth and Makridakis (1981a) have examined the use of arbitrary decision rules and their surprising performance in the context of a simulated competitive business situation.

Identification Methods

These attempt to identify "experts" in groups and use only the judgments of these experts. The key question is how well the more qualified group members (best judgments) can be identified. The standard approaches include: a) obtaining self-rated confidence scores and b) generating external measures of expertise.

Self-rated confidence scores are commonly used to identify the best group member although researchers have warned of the limitations (Einhorn and Hogarth, 1978). It has been noted that people are quite likely to exhibit confidence in fallible judgments (Kahneman and Tversky, 1973). In this context, an external measure of expertise may be more efficient in identifying the more appropriate experts for a given judgmental task. This measure could be used by the information seeker (researcher, analyst, or manager) in selecting experts from a given group or population. The difficulty is obviously in developing a measurement instrument which could be both simple to use in realistic research or management situation, and efficient in identifying experts for specific tasks. This form of "expert" identification is developed and tested in the present study.

The Identification of Experts

The basic considerations for eliciting and combining expert judgments are the criteria and methods for identification and selection of experts. Self ratings and qualifying tasks are methods encountered most often in the literature for separating experts from nonexperts. Noting the degree of expertise, for example, Best (1974) and Dalkey (1969) have shown that self-rated experts provide more accurate estimates than self-rated nonexperts. In some marketing decision-making situations (Jolson and Rossow, 1971), the expertise of the respondents has been verified by asking qualifying questions with known values.

We noted earlier that people tend to give more credence than is often warranted to marginal judgments. The more self-confident members of a group may not necessarily be the better-informed subjects. The opinions of individuals who believe that they are experts can often be affected by what amounts to an uncontrolled expert halo effect (Sackman, 1975). As for validation, confidence scores on seemingly related but specific topics may not extend to the main question and are not necessarily generalizable (Best, 1974). It is believed, therefore, that it is important to measure the respondents' expertise directly so that their knowledgeable opinions can receive the proper attention.

In attempting to evaluate various aspects of decision quality, McIntyre and Currim (forthcoming) have suggested developing multiple measures for a specific setting. In this context, explicit factors (e.g., knowledge, motivation, cognitive style, confidence) affecting a decision maker's experience can be identified. A reasonable definition of expertise involves recognition by others in the field in terms of some measurable criteria (Harmon and Press, 1975). In many situations, a combination of such measures could be used to identify experts. In this study, a screening procedure and a measure reflecting a wide range of factors are used for the identification and selection of experts.

Relative Merits of Integration and Identification Models

It should be apparent that integration methods are in facto substitutes for identification methods when there is no way of identifying experts. But it is certainly suboptimal. It is believed that any improvement in a group decision is not so much the result of a pooling of the members' judgments but the quality of the basic judgments themselves. Therefore, when expert judgment is available, identification methods should clearly be preferred. This in fact raises serious questions about the weighting method. Since this method is based on the relative ability of different individuals, it may require complex mechanisms for assessing conditional probabilities and differential weighting associated with various judgments. One possible alternative might be to appropriately adjust each individual judgment by the degree of expertise (or the variance in the confidence) of the subject. The criterion should clearly be to add the individual whose opinion increases group validity (Hogarth, 1978). It would seem reasonable, however, that stronger results could be obtained by giving exclusive weight to a more carefully arrived at expert opinion. Consequently, the weighting method will not be considered in this study.

Integration methods also tend to be mechanistic and provide less information than alternative methods. They avoid the key issue of defining and evaluating expertise which is at the core of the problem. We know, in fact, from classroom as well as business experience that individuals vary widely in their ability to understand the key determinants of a specific situation. The problem is to determine simple, manageable and realistic methods to identify the more "expert" individuals.

Methods Employed in the Study

In this research, five methods of processing individual judgments are compared in a marketing context:

1. Average judgment
2. Interpersonal group consensus
3. Delphi technique
4. Selecting the judgment of the "best expert" as identified by a self-rated confidence measure
5. Selecting the judgment of the "best expert" as identified by an externally-rated measure of expertise.

The first three are integration methods and the last two are identification methods to elicit individual judgments.

METHODOLOGY

Subjects

Subjects for the experiment were 60 MBA students at a leading International Businss Institute in Europe and 55 MBA students at a large American University on the West Coast of the United States. The students at the European institute represented most European countries and were all fluent in English. The students at the American university were predominantly North American with a heavier representation from the Western states. At the time of the experiment, all subjects were attending a marketing class which included several sessions on marketing communications. All subjects can consequently be considered to have a minimum level of expertise concerning market response to communications investments. However, as will be discussed later, one can still expect some degree of variation in the level of expertise within this subject population.[1]

No significant difference was observed between the European and the American samples on any of the judgmental issues considered. Consequently, the result of the study will be presented on the overall sample of 115 subjects. The stability of sample results over two different environments, however, should be viewed as further evidence of validity of the findings.

The Estimation Task

Historical data on a pharmaceutical product for 48 consecutive monthly periods were distributed to all subjects. This information consisted of market share as well as three types of communication expenditures: direct-mail, sampling, and journal advertising. This information was generated by a "true" distributed lag model. The specifications of this "true" model were taken from the estimates obtained by Montgomery and Silk (1972) based on time series data for an ethical drug, but did not contain an error term. The form of this model which specifies a multiplicative generating process is as follows:

$$(1) \quad \text{Log}[MS(t)] = a_o + \sum_{j=0}^{k_i} a_{ij} \, \text{Log}[X_{i,t-j}]$$

where: $MS(t)$ = market share

$X_{i,t}$ = communication expenditure for type i, i=1 (direct mail), 2 (samples), and 3 (journal advertising), in period t

k_i = maximum lagged effect for communication i, ($k_1 = 4$, $k_2 = 5$, $k_3 = 6$)

a_{ij} = response coefficient for variable X_i with a lag of j.

Historical data for each of the 48 months were generated by drawing at random direct mail, sampling and journal advertising expenditures within a certain range and computing market share using the above model formulation.

After reviewing these historical data, subjects were asked to estimate the market share that could be expected in the 49th month for eight alternate communications plans. These eight communications plans corresponded to all possible combinations of two expenditure levels for each element of the communications mix. Clearly, this estimation task is far from being an obvious one, particularly in view of the multiplicative form with lagged effects terms as specified by the model.

The Experimental Design

The experiment was carried over two 120 minute sessions (one at each institution) according to an identical procedure. The sessions were conducted in two parts. During the first part, all subjects were gathered in a large classroom. The nature of the estimation was described and the task presented as an exercise in the evaluation of market response to communications expenditures. The historical data (presented as real data provided by a pharmaceutical company) and a questionnaire were then distributed to each subject; 60 minutes were allowed for analyzing the historical data and performing the evaluation task individually. In addition to market share estimates for each of the 8 communications plans, subjects were asked to indicate the confidence they had in their estimates (e.g., probability of these estimates being within $\pm$ 25% of actual market shares obtained) and a list of the factors which they had considered in their analysis. At the end of this first part of the session, subjects returned a copy of the questionnaire to the experimenters and kept one copy for themselves.

In the second part of the session, subjects were randomly assigned to one of two experimental conditions: interpersonal or Delphi groups. Fifty five subjects were assigned to 11 interpersonal groups of 5 participants each. The remaining 60 subjects were assigned to 12 Delphi groups of the same size. The 11 interpersonal groups were asked to leave the large classroom and to meet in separate smaller conference rooms in order to reach a group consensus on market share estimates for each of the eight plans within 60 minutes. The subjects for the Delphi experimental condition remained in the large classroom and were assigned seats so that no members of the same group would be in contiguous positions. In addition, a given individual did not know the identity of other members of the group.

Each Delphi group was under the administrative responsibility of a group organizer who had been previously briefed about the mechanics of the technique and did not participate in the study. For each group, the organizer computed the average, lowest and highest estimates provided by the five members for each plan during the preliminary part of the session; this constituted the first iteration in the Delphi procedure. This information was then provided to the members of the group which were given 15 minutes to revise their previous estimates based on this feedback. In addition, they were asked to indicate the main reasons why their new estimates would be higher or lower than the previous group average.

Either the mean or the median can be used as a central tendency measure in returning summary information. The mean gives an equal weight to all inputs whereas the median deemphasizes the views of outliers. The results, however, may not differ significantly fron one another with respect to mean or median differences (Sackman 1975). The mean was selected as a feedback measure of central tendency to minimize the amount of time spent administering the Delphi process, as this was performed in real time while the subjects were captive. The median may be considered a more desirable alternative as it is less sensitive than the mean to extreme judgments. But, its computation takes longer and it is potentially less reliable under time pressure since it requires a rank ordering of data. An ex-post analysis of the implications of using the mean or the median will be discussed in the conclusion section.

In the following iterations, the subjects were given a list of the qualitative arguments provided by all members of the group, as well as the group average and lowest and highest estimates for each plan. A total of three iterations were performed in addition to the plenary session. The total time spent on the Delphi process, including response and analysis for each iteration, was 60 minutes, the same as for the interpersonal group condition.[2]

Following the session, the objectives of the experiment and past research on judgmental estimation were discussed with all subjects. It was clear from this discussion that subjects had not ascertained that the data had been generated by a model. As can be expected, none of the subjects had correctly estimated the form of the underlying model, a practically impossible task in the time available without access to computerized data analysis packages. The results of the study were presented to the participants several weeks later. No uncontrolled interaction took place between subjects as they were under the supervision of experimenters at all times.

At the end of the experiment, data were available on five different methods of obtaining estimates from individual judgments (see Figure 1):

INSERT FIGURE 1 ABOUT HERE

1. Group average. This approach consists of taking the average of the initial estimates provided by the five subjects in each of the 23 groups. It represents an equal weighting of each member of a given group. In order to compare these initial averages with the final estimates obtained under the interpersonal and Delphi treatments, initial interpersonal group averages (IIA) and initial Delphi group averages (IDA) are considered separately. As expected from the random assignments of subjects to the interpersonal and Delphi conditions, no significant differences were observed between the initial averages of these groups.

2. Interpersonal group consensus. These are the final interpersonal group consensus (FIC) estimates obtained by the 11 interpersonal groups.

3. Delphi consensus. These are the final Delphi group consensus (FDC) estimates obtained by the 12 Delphi groups after three iterations.

4. Selection of "experts" on the basis of self-rated confidence. A common method of identifying "experts" is to rely on their self-rated confidence in their judgments. Self-rated experts (SRE) were identified as those who indicated a confidence level equal or above 95% in their initial estimates. In order to appraise the validity of this measure of expertise, a group of self-rated nonexperts (SRN) was identified as comprising of the individuals who indicated a confidence level equal

or below 10% in their initial estimates. The group of self-rated experts and the group of self-rated nonexperts consisted of 11 and 8 subjects respectively.[3]

5. Selection of "experts" on the basis of externally-rated expertise. Based on the list of factors indicated as relevant in the estimation task by each subject in the initial plenary part of the experiment, a measure of externally-rated expertise was developed. It consisted of giving one point to a subject for each factor considered, eliminating possible redundancies, but irrespective of the nature of this factor (e.g., credit was given for factors which were not included in the underlying "true" model, such as seasonality or interaction effects).[4] This measure corresponds to the richness of the cognitive process undertaken by the individual in this specific problem area and is closely related to the concept of integrative cognitive complexity (Schroder, Driver and Streufert, 1967; Larréché, 1979). Also, it closely reflects the concept of expertise without attempting to impose possibly biasing value judgments on different factors indicated by respondents. In addition, it is simple and can be easily applied in practical situations. This measure was used to identify two groups of 9 subjects each: externally-rated experts (ERE) and externally-rated nonexperts (ERN).[5]

Since MBA students constitute a knowledgeable population, they can be considered to have a minimum level of expertise in this case. However, individual variations in the degree of expertise within this population may be expected. The method described here can be used for selecting a panel of appropriate experts. By using a procedure that accurately screens each subject by means of both general background and individual differences, we are in effect generating measures that may permit supportive validation.

Hypotheses

From the 7 basic estimation approaches investigated in the experiment (IIA/IDA, FIC, FDC, SRE, SRN, ERE and ERN), 21 pairwise comparisons of the relative quality of these approaches can be analyzed. A priori, however, only eight hypotheses could be formulated on the basis of past research or of commonly shared expectations of assessment performance for each approach.

H1: FIC > IIA. The consensus reached by the members of an interpersonal group is more likely to be of better quality (i.e., closer to the "true" value) than the average of the initial individual estimates of the group members. This is based on the expectation that the group discussion will provide subjects with additional information which will help them in improving their estimates, and that improvement of the judgments at the individual level will result in a group consensus of better quality than the average of the initial individual estimates.

H2: FDC > IDA. The consensus reached in an anonymous fashion through a Delphi process is more likely to be of better quality than the average of the initial individual estimates of the group members. This is a basic expectation underlying the Delphi technique.

H3: FDC > FIC. The consensus reached through a Delphi process is more likely to be of better quality than the consensus reached in an interpersonal group. This is also a basic expectation underlying the Delphi technique. It is based on the fact that discussion in interpersonal groups may be hampered by factors such as motivation to comply or pressure to conform, and that an anonymous process removes the biasing effects of face-to-face interaction.

H4: SRE > SRN. The estimates provided by "experts" identified by self-rated confidence are more likely to be of better quality than those provided by "nonexperts" identified by the same measure.

H5: SRE > IIA, IDA. The estimates provided by "experts" identified by self-rated confidence are more likely to be of better quality than the average of the initial individual estimates of the group members.

H6: ERE > ERN. The estimates provided by "experts" identified by a simple external measure of expertise are more likely to be of better quality than those provided by "nonexperts" identified by the same measure.

H7: ERE > IIA, IDA. The estimates provided by "experts" identified by a simple external measure of expertise are more likely to be of better quality than the average of the initial individual estimates of the group members.

H8: ERE > SRE. The estimates provided by "experts" identified by a simple external measure of expertise are more likely to be of better quality than those provided by "experts" identified by self-rated confidence.

RESULTS

The measure selected for assessing the relative performance of alternative estimation methods is the Mean Absolute Percentage Error (MAPE) between the judgments obtained and the "true" market share for the eight communications plans considered. The MAPE was selected since there is no justification for a nonlinear loss function and since it does standardize the error by the magnitude of the true value. In order to investigate the convergence of alternative measures of performance, further analysis on the basis of the Root Mean Square (RMS) is also presented.

The objective of the analysis was to compare the performance of alternative judgmental methods. In this study, we have a sample of observations for each of these methods. When selecting a method in a practical situation, however, one does not usually have a sample of observations but a single measure (i.e., a single interpersonal or Delphi group). Consequently, the correct analysis should not be based on the central tendency (such as the mean) of the judgments under various methods as it will not usually be a relevant measure. Rather, the proper approach is to identify which method will have a <u>greater likelihood</u> of providing a better single estimate, without making any assumption about the nature of the distribution of judgmental estimates.

This important research issue is illustrated in Figure 2. Method A may provide, on the average, a better estimate than method B; however, a single estimate from B may be more likely of being closer to the true value than a single estimate from A. In estimation research in particular, the measure of average performance may be expected to be significantly influenced by judgments which deviate highly from the central tendency but have a low likelihood of occurrence.

INSERT FIGURE 2 ABOUT HERE

Traditional parametric tests used in an analysis of variance framework, such as the t or F tests, would not be appropriate as they are based on the difference between the central tendencies of two populations and assume specific distributional properties. By contrast, the Mann-Whitney U test which is based on the probability that a score from a population is larger than a score from another population (see Siegel, 1956, p. 116) corresponds exactly to the type of hypotheses being tested. Moreover, it is particularly appropriate for small sample sizes and is consequently used here.

The results of the experiment are summarized in Table 1. The $\overline{\text{MAPE}}$ measure is the average of the Mean Absolute Percentage Error across all groups or all individuals for a given estimation method. It is given here as a reference but is not used for the computation of the Mann-Whitney U statistic which is based on a case by case analysis rather than on averages. From this measure, it appears that the initial group averages for the interpersonal and Delphi groups are similar; this is to be expected since subjects were assigned at random to the two conditions. The interpersonal group consensus appears on the average to have provided estimates of similar quality to those of the initial group averages, while an improvement was achieved with the Delphi process.

INSERT TABLE 1 ABOUT HERE

An important difference can be observed between the "experts" and "nonexperts" identified by self-rated confidence; this difference is even higher when they are identified by an external measure of expertise. The quality of the judgments is comparable for the experts identified by self-rated confidence and for the Delphi technique; although it is higher for the experts identified by an external measure of expertise.

The analysis of the Mann-Whitney U statistic would indicate if one method provides significantly better estimates than another in the context of the previously formulated hypotheses. The conclusions drawn are summarized below.

Hypothesis 1 is not confirmed.[6] Interpersonal group consensus does not appear to provide significantly better estimates than the average of the group members' initial judgments. Although additional information was exchanged between group members during the discussion, this information did not result in better estimates. An important reason for this phenomenon is that the most influential members of a group in a face-to-face interaction may not be the best judge--an issue often noted in evaluation of the performance of interpersonal groups.

Hypotheses 2 and 3 are confirmed. The Delphi process appears to provide significantly better estimates than the average of the group members' initial judgments and than the consensus reached by comparable interpersonal groups. These results are consistent with the rationale underlying the Delphi technique.

Hypotheses 4 and 5 are not confirmed. Self-rated confidence does appear to discriminate on average between "experts" and "non-experts," but "experts" identified in this fashion are not more likely to provide significantly better estimates than the average of the group members' initial judgments, or than "nonexperts."[7]

Hypotheses 6, 7 and 8 are confirmed. "Experts" identified by a simple external measure of expertise appear to provide significantly better estimates than "nonexperts" identified by the same measure, than the average of the group members' initial judgments, and than "experts" identified by self-rated confidence.

In addition to the above hypotheses, the experiment also provides interesting results on the relative performance of "experts" identified through self-rated confidence or through an external measure compared to the final consensus estimates of the interpersonal and Delphi groups. The "experts" identified by self-rated confidence do not provide significantly better estimates than the interpersonal or Delphi groups, while the "experts" identified by an external measure of expertise do. The identification of "experts" by an external measure of expertise thus appears to be significantly superior to any of the other alternative methods considered.

When using the Root Mean Square (RMS) as a measure of predictive validity, the results (summarized in Table 2) are globally similar to the ones obtained with the Mean Absolute Percentage Error (MAPE).

INSERT TABLE 2 ABOUT HERE

As can be expected, the results are more unstable with RMS than with MAPE because of the greater importance given to outliers. In particular, the average Root Mean Square ($\overline{RMS}$) is substantially different for the initial interpersonal (IIA) and Delphi (IDA) group averages although subjects were assigned randomly to these two conditions. The distribution of the two groups, however, is not statistically significant as tested by the Mann-Whitney U test.

In terms of the eight hypotheses tested in this study, two differences (with regard to H2 and H7) appear when using RMS instead of MAPE:

a. the consensuses reached through the Delphi process (FDC) do not provide statistically significant improvement over the initial Delphi group averages (IDA) -- Hypothesis 2.

b. the estimates given by experts identified by an external measure of expertise (ERE) do not provide statistically significant improvement over the initial Delphi group averages (IDA) -- Hypothesis 7.

Both of these instances involve a comparison with the initial Delphi group averages (IDA) which appear to provide a rather favorable set of estimates in conjunction with RMS. The same comparisons performed with the initial interpersonal group averages (IIA) are, however, statistically significant. These two differences between MAPE and RMS measures are due to the fact that the initial interpersonal (IIA) and Delphi (IDA) group averages are substantially different with the use of the RMS measure despite the random assignment of subjects to the two groups. A single outlier judgment in the initial estimates may indeed be given a disproportionate importance when the RMS measure is employed.

Apart from the above two differences, the MAPE and RMS measures give similar results in terms of the various hypotheses tested in this study and show a high degree of convergent validity. In particular, the use of the RMS measure confirms the better predictive validity of estimates by externally-rated experts (ERE) compared to other methods of eliciting judgments.

To give a basis of comparison for the $\overline{\text{MAPE}}$ and $\overline{\text{RMS}}$ measures, a simple log linear model expressing market share as a function of the three elements of the communications mix was estimated on the 48-month historical data, without lagged effects. This model was then used to predict market share for the eight communications plans considered. The $\overline{\text{MAPE}}$ and $\overline{\text{RMS}}$ for the predicted value of market share were 5.07 and 4.35 respectively. Most of the estimates obtained by the alternative judgmental approaches are of lesser quality which largely reflects the complexity of the estimation task. The only estimates which are of better quality are those obtained from the "experts" identified by an external measure of expertise when $\overline{\text{MAPE}}$ is used as a measure of predictive validity. As already noted, the purpose of this research was not to compare judgmental with analytical methods, but rather to evaluate the relative merits of alternative judgmental methods. In many real world situations, these judgments are the only estimates available.

CONCLUSIONS

The approaches considered in this experiment may be classified in three groups, in increasing order of likelihood of providing better judgments:

1. Averaging individual judgments, reaching an interpersonal group consensus, or identifying "experts" on the basis of self-rated confidence.
2. Reaching a group consensus through a Delphi process.
3. Identifying "experts" on the basis of an external measure of expertise.

Although one should proceed with caution in attempting to generalize these results to other types of estimation tasks, the findings do provide a basis for further research under different experimental conditions and for a comparison of the relative merits of different methods. The choice of a specific method in practice will depend on the time, cost and implementation aspects of various methods.

In this study it was found that averaging individual judgments (possibly the easiest method) provides estimates of comparable quality to those of interpersonal group consensus. Although an interpersonal group meeting may be desired in a given situation (e.g., to facilitate implementation or to reduce the perceived risk of missing some important factors in the estimation task), taking the average of the initial individual judgments can provide an equally valuable reference basis.

Identifying "experts" on the basis of self-rated confidence is not likely to provide better estimates than the above two approaches. It is possible, however, that self-rated confidence may be a more appropriate measure of expertise when subjects can actually evaluate their confidence in terms of a specific problem area to which they are regularly exposed. In typical laboratory experiments on judgment, subjects are often asked to evaluate their confidence for a task which is unrelated to their normal activities (extreme examples are the estimation of the number of marbles in an urn or the population of country X). In this case, it may be argued that it is not confidence in a specific task which is measured but rather a generalized construct of confidence; this, in turn, could explain the poor results usually obtained with self-rated confidence measures under those conditions. Here, however, self-rated confidence has been defined in terms of a specific task on a topic in which subjects had acquired a minimum level of knowledge in marketing classes. This experiment consequently raises additional doubts on the validity of self-rated confidence as a measure of expertise.

The Delphi process is more likely to provide better estimates than the interpersonal group consensus or the average of the group members' initial judgments. It does not, however, significantly improve on the judgments provided by "experts" identified by self-rated confidence. This may illustrate the fact that the highly confident individuals may be less willing to update their initial judgments and consequently pull the other members of the group towards their own points of view even in the absence of interpersonal interactions. Indeed, if the Delphi technique successfully eliminates biases due to face-to-face interactions, it is still influenced by the willingness of individual judges to revise their estimates in the light of new information. This in fact may be a desirable feature when confidence is actually based on a specific body of knowledge as indicated earlier. Moreover, the value of the Delphi technique may be higher when "unexpected" information may be obtained from other members of the group--a condition which is more likely in long range forecasting applications for which the technique was originally designed (Linstone and Turoff, 1975).

Finally, the results of the experiment indicate that a simple external measure of expertise can be developed which successfully identifies "experts" who provide significantly better estimates than the other approaches. The real relevant issue is, of course, identifying "experts." And, we have shown that this is possible with a simple external measure. This would suggest that simple "unobtrusive" measures a) can be used to identify the "best expert" in a given subject and b) are preferable to measures which tap on more general traits such as self-rated confidence.[8] Accordingly, it can be argued that the strong results of Fudge and Lodish (1977) field survey reflects subjects' experience (judgments made by salespeople on sales problems with familiar products). On the other hand, laboratory experiments including Chakavarti, Mitchell and Staelin (1979) study tend to show less-informed judgments by subjects in situations in which they are not experts.

Methodological Issues

This study provides evidence that selecting experts on the basis of a measure that is simple, area-specific, and externally qualified can be an effective means of identifying the most accurate judgments. However, some difficulties and limitations, both conceptually and methodologically, remain.

First, the participants in the study were knowledgeable MBA students. While there is no evidence that this affected the results (comparison of various methods), the ideal subjects for experiments designed to elicit managerial judgments would be experienced decision makers (i.e., actual managers). Laboratory experiments and field applications are different approaches used to address issues with regard to judgmental estimation. While laboratory studies offer some control advantages, they have external validity problems. The use of actual decision makers in real world situations would provide external validation thus making the results more generalizable (Little and Lodish, 1981).

Second, although all subjects in the study had sufficient training to perform the required task, it is anticipated that removal of nonexperts from all groups could reduce the judgmental error. In fact, a reexamination of Delphi results indicated that, as expected, groups with two or more experts (externally qualified and identified ex-post) provided better estimates than did groups with nonexperts. There were significant differences in the MAPE and RMS measures for Delphi groups with and without experts (Mann-Whitney U Test, $p < .05$). It should be clear that using an external measure of expertise to select a priori members of Delphi groups will capitalize on the benefits of these two identification and integration methods and will provide better results.

Third, there is a debate in the literature on the Delphi process regarding the use of mean or median values as measures of central tendency. It would be interesting to investigate ex-post the potential impact that the choice of a feedback measure of central tendency may have on the performance of the Delphi groups. Short of testing the use of both measures in two separate and controlled groups, one could see which of two measures gave feedback closer to the true value at the end of the first iteration. This first feedback is believed to be important as it will influence the future evolution of the group judgments. Out of 96 cases (8 estimates for each of 12 Delphi groups), the mean and median gave identical feedback measures in two instances; the median was closer to the true value than the mean in 53 instances; and the mean was closer to the true value than the median in the remaining 41 instances. At the end of the first iteration, the average of the mean absolute percentage error ($\overline{\text{MAPE}}$) for all the Delphi groups combined would have been 6.57 compared to 7.86 on the basis of the median and the mean respectively. This would seem to indicate that the median is preferable to the mean as a measure of central tendency in Delphi studies. It is also likely on the basis of this evidence that the performance of the Delphi groups could have been further improved by using the median as a measure of central tendency.

Directions for Future Research

Although marketing education and marketing management are based on the acquisition and exploitation of marketing expertise, research on the issue of identifying experts for specific marketing situations is conspicuously non-existent. In particular, the relative success of the decision-calculus type of marketing models is based in great part on the utilization of managerial judgment in addition to available empirical data. In this situation "experts" are identified as the managers who are most familiar with the problem addressed by the model and who eventually will have to make the decision. This raises three important issues for further research. The first issue concerns the development of tools to more systematically identify "experts" for different inputs from a large pool of individuals (including, for example, product managers, advertising staff, and market researchers). This becomes more important as marketing gets increasingly involved with decisions at a more strategic level (Day, 1977; Larréché and Srinivasan, (1981). Empirical data on many strategic issues, such as competitive reactions, are not easily available and managerial judgment becomes a crucial component of decision making. Moreover, decisions at a strategic level usually draw on a broader basis of managerial judgment than at a tactical level, making it even more important to select the most appropriate judgment for specific tasks.

The second implication concerns the validity of the decision-calculus approach. Since it relies heavily on judgments and the quality of these judgments may vary significantly with expertise in a specific product and market area, it can be adequately tested only in difficult field situations. The study by Fudge and Lodish (1977) is an excellent illustration of this research approach which hopefully will develop in the future.

The third issue stemming from this study is how we can increase the level of expertise of an individual in a specific subject area. This is obviously the objective of education in general and one of our missions as marketing educators. The attempt by McIntyre and Currim (forthcoming) to identify factors that might affect decision quality in a particular situation is of interest in this context. Research on measuring the effectiveness of our pedagogical methods in reaching this objective is, however, in its infancy and remains an important challenge for the future.

FOOTNOTES

1. In the real world situation, experienced managers may prove better decision makers because of their rich knowledge of the marketing environment (Little and Lodish, 1981). It is often claimed that business students, particularly MBAs, represent a population with sufficient interest and ability to act as surrogates for actual managers (McIntyre, 1982). When the objective of judgmental research is to evaluate the quality of human judgment (for instance, compared to estimates generated by analytical methods), using MBA students as surrogates for managers still leaves some external validity issues unresolved. Research on the relative efficiency of alternative judgmental approaches which is the issue addressed in the current paper should, however, be less affected by sample characteristics.

2. Out of this time, the members of the Delphi groups spent only 45 minutes actively revising their judgments and a total of 15 minutes were required to administer the feedback process. During this "unproductive" time, subjects could, however, still think about the estimation task. The idle time could be reduced with the use of interactive computer data processing; however, the proper equipment was not available in the current experiment.

3. The extreme ranges (as opposed to upper half and lower half) of self-rated experts were selected for an effective comparison of the most self-confident and the least self-confident members. We have selected the expert and nonexpert groups to represent the extreme deciles of our sample. The cutoff point on the rating scales was selected so that the actual size of the extreme groups was closest to 10% of the sample (11.5).

4. To illustrate, let's consider a subject indicating the following response: "(a) market share affected by total investment; (b) journal advertising as most determinant, with (c) effects lasting for several months; (d) seasonal variations accounted for; and (e) all significant changes in market share resulting from important changes in advertising." This response would receive a score of 4 as the subject accounted for total investment, importance of advertising, lagged effects, and seasonal variations; despite the fact that some of these observations do not reflect factors in the underlying model. Note that the fifth item is not considered since it is redundant with the second item concerning the relative importance of advertising.

5. There is evidence of supportive validation for the specific measure of expertise employed in this study. Significant Spearman rank correlation coefficients for the sample ($n = 115$) were found between externally-rated expertise and self-rated confidence ($p < .05$) and MAPE ($p < .025$) respectively. Moreoever, externally qualified experts outperformed other groups at the extreme ends of the expertise range. Also, independent ratings of subjects' expertise by the two investigators yielded a highly significant coefficient of inter-rater reliability (Spearman correlation .69, $p < .005$).

6. By this, we mean that the alternative null hypothesis could not be rejected at a 0.10 level of significance.

7. Note that a different conclusion could have been drawn by considering the average performance, rather than the likelihood of a better performance, as the $\overline{\text{MAPE}}$ for the "experts" identified by self rated confidence is much lower than for the initial group averages (see Table 1).

8. It should be noted that the concept of expertise, when it is defined in terms of a specific task, may be different from intelligence, experience, or other general attributes associated with judgment. For example, a CEO of a corporation may be brighter and more experienced than his subordinates but not an expert on many tasks such as media planning. A distinction should be made between expertise and more general characteristics or traits which are not task-specific.

Table 1

Results of the Experiment--with MAPE

	n	$\overline{MAPE}$ [a]	Mann-Whitney signficance levels[b] IIA	IDA	FIC	FDC	SRN	SRE	ERN	ERE
IIA	11	7.32	-	-	-	-	-	-	-	-
IDA	12	7.86	NS	-	-	-	-	-	-	-
FIC	11	7.27	NS	NS	-	-	-	-	-	-
FDC	12	6.52	NS	.05	.05	-	-	-	-	-
SRN	8	13.42	NS	NS	NS	-.10	-	-	-	-
SRE	11	6.42	NS	NS	NS	NS	NS	-	-	-
ERN	9	13.63	NS	-.10	-.10	-.05	NS	-.05	-	-
ERE	9	4.65	.005	.001	.005	.05	.01	.05	.005	-

[a]MAPE (Mean Absolute Percentage Error) is the average absolute percentage error over the 8 estimates of market share for a given group or for a given individual (in the case of "experts" and "nonexperts"). $\overline{MAPE}$ is the average of MAPE across all groups or all individuals for a given estimation method.

[b]One-tailed test. A negative sign is shown when the row method is more likely to provide worse estimates than the column method. The significance levels were computed with the statistical package developed by Coleman (1972a, 1972b). NS = not significant at the 0.10 level.

Table 2

Results of the Experiment--with RMS

	n	$\overline{RMS}$[a]	Mann-Whitney signficance levels[b] IIA	IDA	FIC	FDC	SRN	SRE	ERN	ERE
IIA	11	10.90	-	-	-	-	-	-	-	-
IDA	12	8.28	NS	-	-	-	-	-	-	-
FIC	11	11.13	NS	NS	-	-	-	-	-	-
FDC	12	11.19	.025	NS	.01	-	-	-	-	-
SRN	8	43.44	NS	-.025	NS	-.01	-	-	-	-
SRE	11	13.83	NS	-.05	NS	-.01	NS	-	-	-
ERN	9	46.09	-.025	-.01	-.025	-.001	NS	NS	-	-
ERE	9	6.24	.05	NS	.025	NS	.025	.025	.001	-

[a]RMS (Root Mean Square) is the average squared error over the 8 estimates of market share for a given group or for a given individual (in the case of "experts" and "nonexperts"). $\overline{RMS}$ is the average of RMS across all groups or all individuals for a given estimation method.

[b]One-tailed test. A negative sign is shown when the row method is more likely to provide worse estimates than the column method. The significance levels were computed with the statistical package developed by Coleman (1972a, 1972b). NS = not significant at the 0.10 level.

Figure 1

Alternative Methods to Elicit Judgments

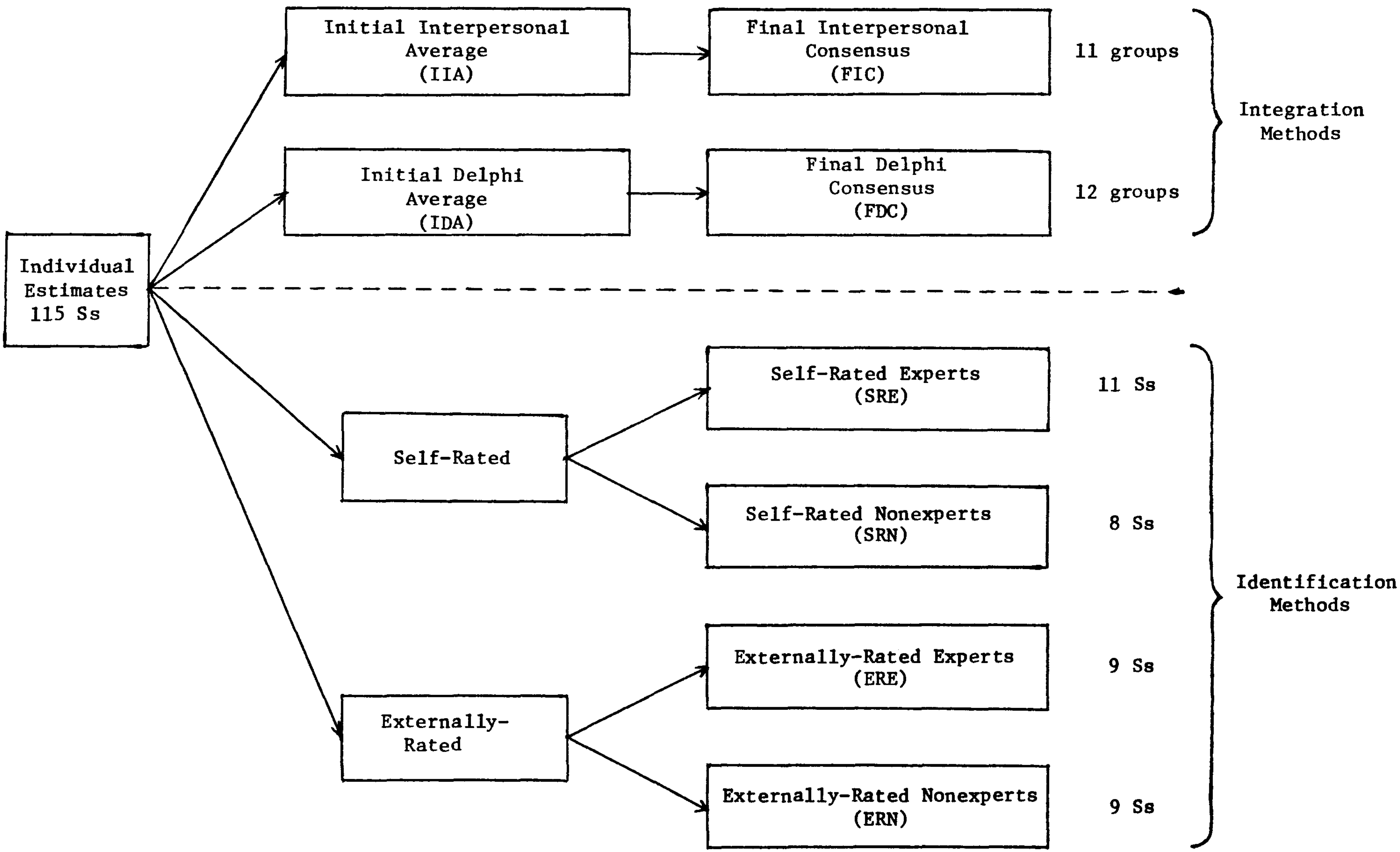
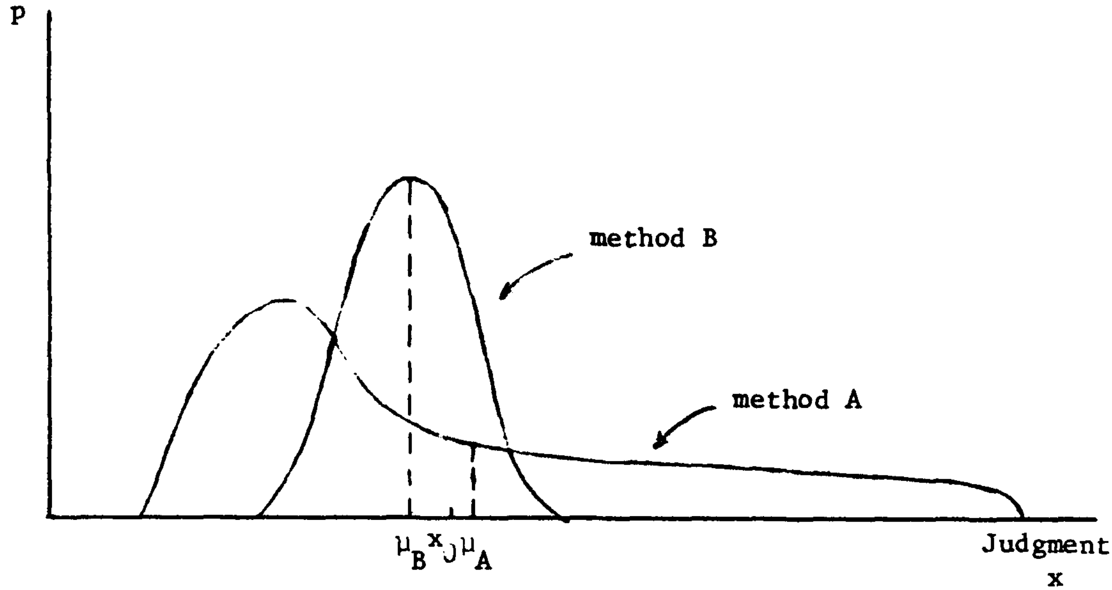


Figure 2

Likelihood of Better Performance vs. Better Average Performance



μ_A = mean judgment under method A

μ_B = mean judgment under method B

x_0 = true value

$|\mu_A - x_0| < |\mu_B - x_0|$ but

$P(|x_B - x_0| < |x_A - x_0|) > P(|x_A - x_0| < |x_B - x_0|)$

REFERENCES

Basu, Shankar, and Roger G. Schroeder (1977), "Incorporating Judgments in Sales Forecasts: Application of the Delphi Method at American Hoist & Derrick," Interfaces, 7 (May), 18-27.

Beach, Lee R., and Terence R. Mitchell (1978), "A Contingency Model for the Selection of Decision Strategies," Academy of Management Review, 3 (July), 439-449.

Best, Roger J. (1974), "An Experiment in Delphi Estimation in Marketing Decision Making," Journal of Marketing Research, 11 (November), 447-468.

Camerer, Colin (1980), "The Validity and Utility of Expert Judgment," Research Paper, Graduate School of Business, University of Chicago.

Chakravarti, Dipankar, Andrew Mitchell, and Richard Staelin (1979), "Judgment Based Marketing Decision Models: An Experimental Investigation of the Decision Calculus Approach," Management Science, 25 (March), 251-263.

Chakravarti, Dipankar, Andrew Mitchell, and Richard Staelin (1981), "Judgment Based Marketing Decision Models: Problems and Possible Solutions," Journal of Marketing, 45 (Fall), 13-23.

Coleman, Denis R. (1972a), "The Stanford Graduate School of Business Data Management and Analysis System: Part I--The Data Management System," Technical Report, Graduate School of Business, Stanford University.

Coleman, Denis R. (1972b), "Non Parametric Statistical Package," Technical Report, Graduate School of Business, Stanford University.

Dalkey, Norman (1969), The Delphi Method: An Experimental Study of Group Opinion, Santa Monica: The Rand Corporation, RM-5888-PR.

Dalkey, Norman, and Olaf Helmer (1963), "An Experimental Application of the Delphi Method to the Use of Experts," Management Science, 9 (April), 458-467.

Dawes, Robyn M. (1971), "A Case Study of Graduate Admissions: Application of Three Principles of Human Decision Making," American Psychologist, 26 (February), 180-188.

Dawes, Robyn M., and Bernard Corrigan (1974), "Linear Models in Decision Making," Psychological Bulletin, 81 (February), 95-106.

Day, George S. (1977), "Diagnosing the Product Portfolio," Journal of Marketing, 41 (April), 29-38.

Einhorn, Hillel J., and Robin M. Hogarth (1975), "Unit Weighting Schemes for Decision Making," Organizational Behavior and Human Performance, 13 (April), 171-192.

Einhorn, Hillel J., and Robin M. Hogarth (1978), "Confidence in Judgment: Persistence of the Illusion of Validity," Psychological Review, 85(5), 395-416.

Einhorn, Hillel J., Robin M. Hogarth, and Eric Klempner (1977), "Quality of Group Judgment," Psychological Bulletin, 84(1), 158-172.

Fudge, William K., and Leonard M. Lodish (1977), "Evaluation of the Effectiveness of a Salesman's Planning System by Field Experimentation," Interfaces, 8 (November), 97-106.

Goldberg, Lewis R. (1970), "Man Versus Model of Man: A Rationale, Plus Some Evidence for a Method of Improving on Clinical Inferences," Psychological Bulletin, 73 (June), 422-432.

Gustafson, David H., Ramesh K. Shukla, André Delbecq, and G. William Walster (1973), "A Comparative Study of Differences in Subjective Likelihood Estimates Made by Individuals, Interacting Groups, Delphi Groups and Nominal Groups," Organizational Behavior and Human Performance, 9 (April), 280-291.

Harman, Alvin J., and S. James Press (1975), Collecting and Analyzing Expert Group Judgment Data. Santa Monica: The Rand Corporation P-5467.

Hogarth, Robin M. (1978), "A Note on Aggregating Opinions," Organization Behavior and Human Performance, 21 (February), 40-46.

Hogarth, Robin M., and Spyros Makridakis (1981a), "The Value of Decision Making in a Complex Environment: An experimental Approach," Management Science, 27 (January), 93-107.

Hogarth, Robin M., and Spyros Makridakis (1981b), "Forecasting and Planning: An Evaluation," Management Science, 27 (February), 115-138.

Jolson, Marvin A., and Gerald L. Rossow (1971), "The Delphi Proeess in Marketing Decision Making," Journal of Marketing Research, 8 (November), 443-448.

Kahneman, Daniel, and Amos Tversky (1973), "On the Psychology of Prediction," Psychological Review, 80 (July), 237-251.

Larréché, Jean-Claude (1979), "Integrative Complexity and the Use of Marketing Models," in R. Doktor, R. L. Schulz and D. P. Slevin, eds., The Implementation of Management Science, New York: North-Holland.

Larréché, Jean-Claude, and V. Srinivasan (1981), "STRATPORT: A Decision Support System for Strategic Planning," Journal of Marketing, 45 (Fall), 39-52.

Linstone, Harold A., and Murray Turoff, eds. (1975), The Delphi Method: Techniques and Applications, New York: Addison-Wesley.

Little, John D. C. (1970), "Models and Managers: The Concept of a Decision Calculus," Management Science, 16 (April), 446-485.

Little, John D. C., and Leonard M. Lodish (1981), "Commentary on Judgment Based Marketing Decision Models," Journal of Marketing, 45 (Fall), 24-29.

Makridakis, Spyros, and Michéle Hibon (1979), "Accuracy of Forecasting: An Empirical Investigation," Journal of the Royal Statistical Society, A, 142(2), 97-125.

McIntyre, Shelby H. (1980), "The Leverage Impact of Judgment Based Marketing Models," in Market Measurement and Analysis, Proceedings of the First ORSA/TIMS Special Interest Conference, D. B. Montgomery and D. R. Wittink, eds., Cambridge, MA: Marketing Science Institute, 551-567.

McIntyre, Shelby H. (1982), "An Experimental Study of the Impact of Judgment Based Marketing Models," Management Science, 28 (January), 17-33.

McIntyre, Shelby H., and Imran S. Currim (forthcoming), "Evaluating Judgment Based Marketing Models: Multiple Measures, Comparisons and Findings," in A. Zoltners, ed., Marketing Decision Models, TIMS Studies in the Management Sciences, New York: North-Holland.

Milkovich, George T., Anthony J. Annoni, and Thomas A. Mahoney (1972), "The Use of the Delphi Procedures in Manpower Forecasting," Management Science, 19 (December), 381-388.

Montgomery, David B., and Alvin J. Silk (1972), "Estimating Dynamic Effects of Market Communications Expenditures," Management Science, 18 (June), 485-501.

Sackman, Harold (1975), Delphi Critique ÷ Expert Opinion, Forecasting, and Group Process, Lexington, Mass: Lexington Books.

Schroder, Harold M., Michael J. Driver, and Siegfried Streufert (1967), Human Information Processing, New York: Holt, Rinehart and Winston.

Siegel, Sidney (1956), Nonparametric Statistics for the Behavioral Sciences, New York: McGraw-Hill.

Winkler, Robert L. (1967), "The Assessment of Prior Distributions in Bayesian Analysis," Journal of the American Statistical Association, 62 (September), 776-781.

LIST OF INSEAD RESEARCH WORKING PAPERS

80/01 "Identifying cognitive style determinants of retail patronage, by Christian PINSON, Arun K. JAIN and Naresh K. MALHOTRA, January 1980.

80/02 "Dimensions culturelles des conceptions de management - une analyse comparative internationale", par André LAURENT Février 1980.

80/03 "Cognitive style and effective communication", by Arun K. JAIN, Naresh K. MALHOTRA and Christian PINSON, Dec. 1979.

80/04 "Accomodative cognitive style differences in consumer reduction of alternatives", by Naresh K. MALHOTRA, Christian PINSON and Arun K. JAIN, October 1979.

80/05 "Stability and reliability of Part-Worth utility in conjoint analysis : a longitudinal investigation", by Arun K. JAIN, Naresh K. MALHOTRA and Christian PINSON, September 1979.

80/06 "The expected future spot exchange rate, the forward rate, and the trade balance", by Charles A. WYPLOSZ, March 1980.

80/07 "Decline and adjustment: Public intervention strategies in the European clothing industries", by José de la TORRE, July 1980.

80/08 "The uncommon market: European policies towards a crisis industry - clothing in the 1970's", by José de la TORRE and Michel BACCHETTA, May 1980.

80/09 "Stratport: a decision support system for strategic planning", by Jean-Claude LARRECHE and V. SRINIVASAN, April 1980, Revised October 1980.

80/10 "A new approach to market segmentation strategy: a banking application", by Arun K. JAIN, Christian PINSON and Naresh K. MALHOTRA, March 1980.

80/11 "The exchange and interest rate term structure under risk aversion and rational expectations", by Charles A. WYPLOSZ, Revised Version, September 1980.

80/12 "Individual cognitive differences in MDS analysis of perceptions", by Arun K. JAIN, Naresh K. MALHOTRA and Christian PINSON, July 6-12, 1980.

80/13 "STRATPORT: A Model for the evaluation and formulation of Business Portfolio Strategies", by Jean-Claude LARRECHE and V. SRINIVASAN, April 1980, Revised November 1980.

80/14 "Les styles cognitifs : une nouvelle approche de la segmentation des marchés, by Christian PINSON, Naresh K. MALHOTRA and Arun K. JAIN, Septembre 1980.

80/15 "Cognitive styles: A new approach to market segmentation", by Christian PINSON, Naresh K. MALHOTRA and Arun K. JAIN, March 1980.

81/01 "Eurobanking, open market operations and the monetary base" by Herwig LANGOHR, August 1980.

81/02 "Alternative approaches to the theory of the banking firm: a note" by Herwig LANGOHR, September 1980.

81/03 "Why does beta shift when the length of securities returns varies?" by Gabriel HAWAWINI, December 1980.

81/04 "Forward market and the cooperative firm" by Gabriel HAWAWINI, January 1981.

81/05 "On some propositions regarding the behavior of the labor-managed firm under uncertainty" by Gabriel HAWAWINI, Jan. 1981.

81/06 "Impact of the investment horizon on the association between securities' risk and return: theory and tests" by Gabriel HAWAWINI and Ashok VORA, February 1981.

81/07 "New evidence on beta stationarity and forecast for Belgian common stocks" by Gabriel A. HAWAWINI and Pierre A. MICHEL, February 1981.

81/08 "Industrial policy in the European economic community : Crisis and change", by Kenneth S. COURTIS, June 1980.

81/09 "Dogmatism as a moderator of banking behavior and attitudes", by Arun K. JAIN, Christian PINSON and Naresh K. MALHOTRA, March 1981.

81/10 "Investment horizon, diversification, and the efficiency of alternative beta forecasts", by Gabriel A. HAWAWINI and Ashok VORA, March 1981.

81/11 "Organizational Development & Change", by Claude FAUCHEUX, Gilles AMADO and André LAURENT, April 1981.

81/12 "The Pricing of Risky Assets on the Belgian Stock Market" by Gabriel HAWAWINI and Pierre A. MICHEL, May 1981.

81/13 "A Test of the Generalized Capital Asset Pricing Model" by Gabriel HAWAWINI and Ashok VORA, May 1981.

81/14 "On the History of Yield Approximations" by Gabriel HAWAWINI and Ashok VORA, May 1981.

81/15 "Pitfalls of the Global Product Structure" by William H. DAVIDSON and Philippe C. HASPESLAGH, May 1981.

81/16 "Adjusting Beta Estimates: Real Gains or Illusions?" by Gabriel A. HAWAWINI and Ashok VORA, June 1981.

81/17 "Do European Industrial Marketers Budget Differently? an International Comparison via the Advisor Model" by David WEINSTEIN and Gary L. LILIEN, June 1981.

81/18 "The Internationalisation of Manufacturing in the Automobile Industry - Some Recent Trends" by Yves L. DOZ, April 1981.

81/19 "Portfolio Planning: Use and Usefulness" by Philippe HASPESLAGH, May 1981.

81/20 "Production Decisions in the Mixed Firm" by Claude VIALLET, Octobre 1981.

81/21 "Foreign Investment and Economic Development: Conflict and Negotiation", by José de la TORRE, April 1981.

81/22 "Forecasting Country Political Risk" by José de la TORRE and David H. NECKAR, March 1981.

81/23 "The Impact of Inflation on Real Variables: A General Equilibrium Approach", by Antonio M. BORGES, November 1981.

81/24 "An Assessment of the Risk and Return of French Common Stocks", by Gabriel A. HAWAWINI, Pierre A. MICHEL and Claude J. VIALLET, November 1981.

81/25 "Mode de vie et style de vie : Quatre observations sur le fonctionnement des termes" par Jean-François BERNARD-BECHARIES et Christian PINSON.

81/26 "Simulating an Oil Shock with Sticky Prices" by Francesco GIAVAZZI, Mehmet ODEKON and Charles WYPLOSZ, November 1981.

81/27 "Decomposing the Impact of Higher Energy Prices on Long-Term Growth" by Antonio M. BORGES and Lawrence H. Goulder.

81/28 "Forecasting for Industrial Products" by David WEINSTEIN.

82/01 "Estimating and Adjusting for the Intervalling-Effect Bias in Beta" by Kalman J. COHEN, Gabriel A. HAWAWINI, Steven F. MAIER, Robert A. SCHWARTZ and David K. WHITCOMB. February 1980, Revised October 1981.

82/02 "Friction in the Trading Process and the Estimation of Systematic Risk" by Kalman J. COHEN, Gabriel A. HAWAWINI Steven F. MAIER, Robert A. SCHWARTZ and David K. WHITCOMB. November 1981.

82/03 "On the Mathematics of Macaulay's Duration: A Note" by Gabriel A. HAWAWINI. December 1981.

82/04 "Systematic Risk, the Investment Horizon, and the Market Index: an Analytical Examination" by Gabriel A. HAWAWINI and Ashok VORA, December 1981.

82/05 "Why Beta Estimates Depend upon the Measurement Interval" by Gabriel A. HAWAWINI. January 1982.

82/06 "Nationalization, Compensation and Wealth Transfer: an Empirical note about the French Experience" by Herwig LANGOHR and Claude VIALLET, 1981/1982.

82/07 "The Keynesian and Classical Determination of the Exchange Rate" by Emil-Maria CLAASSEN, May 1982.

82/08 "The Real Exchange Rate, the Current Account and the Speed of Adjustment" by Francesco GIAVAZZI and Charles WYPLOSZ, April 1982.

82/09 "Simulation: A Complementary Method for Research on Strategic Decision Making Processes" by Danielle NEES, May 1982.

82/10 "The Zero-Root Problem: Dynamic Determination of the Stationary Equilibrium in Linear Models" by Francesco GIAVAZZI and Charles WYPLOSZ, August 1982.

82/11 "The Theory of Risk Aversion and Liquidity Preference: A Geometric Exposition" by Gabriel A. HAWAWINI.

82/12 "The Effect of Production Uncertainty on the Labor-Managed Firm" by Gabriel A. HAWAWINI and Pierre A. MICHEL.

82/13 "On the Independence Between Deposit and Credit Rates" by Jean DERMINE, September 1982.

82/14 "Entrepreneurial Activities of INSEAD MBA Graduates" by Lister VICKERY, October 1982.

82/15 "Proportional VS. Logarithmic Models of Asset Pricing" by Gabriel A. HAWAWINI, July 1982.

82/16 "Capital Controls: Some Principles and the French Experience" by Emil-Maria CLAASSEN and Charles WYPLOSZ, October 1982.

82/17 "The Third World's Campaign for a new International Economic Order" by Jan MURRAY, October 1982.

82/18 "Extremity of Judgment and Personality Variables: Two Empirical Investigations" by Naresh K. MALHOTRA, Arun K. JAIN and Christian PINSON, April 1982. Revised July 1982.

EURO-ASIA CENTRE

INSEAD

Institut Européen d'Administration des Affaires
European Institute of Business Administration
Europäisches Institut für Unternehmensführung
Institut Privé d'Enseignement Supérieur

Boulevard de Constance
77305 Fontainebleau Cedex, France
Telephone (6) 422 48 27 Telex 690389F

EAC Publications List
Update September 1982

E A C BRIEFING PAPERS

N°1. Strategies and Practices of Transfer of Technology from European to Asean Enterprises.

Philippe LASSERRE and Max BOISOT. April 1980. 30 p.

N°2. The Partners of Foreign Investors in Indonesia : the Tip of the Ethnic Iceberg. (working draft)

Stephen C. HEADLEY. December 1980. 21 p.

N°3. Foreword to Government-Business Relations in Indonesia. (working draft)

Stephen C. HEADLEY. December 1980. 17 p.

N°4. Personnel Management in Indonesia : How ? (working draft)

Stephen C. HEADLEY. December 1980. iv, 16 p.

N°5. Can you work for Honda and remain yourself ? The Cultural Dimensions of Indonesian Management. (working draft)

Stephen C. HEADLEY. December 1980. 17 p.

N°6. The Context of Management Development in Malaysia.

Bryony CONWAY. December 1980. 17 p.

N°7. Racial Balance in Management in Malaysia.

Bryony CONWAY. December 1980. 13 p.

N°8. Appropriate Education for Management in Malaysia.

Bryony CONWAY. December 1981. 10 p.

N°9. Foreign Enterprise and Management Development in Malaysia.

Bryony CONWAY. November 1981. 8 p.

N°10 The Chinese Malaysian Enterprise.

Bryony CONWAY. June 1982. 12p.

<u>E A C RESEARCH PAPERS</u>

N°1. A Contribution to the Study of Entrepreneurship Development in Indonesia.

Philippe LASSERRE. April 1979 (corrected version 1980).
72, 7 p. (limited distribution)

N°2. The Transfer of Technology from European to Asean Enterprises : Strategies and Practices in the Chemical and Pharmaceutical Sectors.

Philippe LASSERRE and Max BOISOT. February 1980. 109, VI p.

N°3. Possibilité d'un transfert à l'étranger des techniques japonaises de gestion : le cas français.

Tetsuo AMAKO. July 1982. 145 p.

E A C REPRINTS

N°1. Japanese Organizational Behaviour : A Psychocultural Approach.

Henri-Claude de BETTIGNIES. February 1981.
Reproduced from : Management Research: A Cross-Cultural Perspective.
Edited by Desmond Graves. Amsterdam, London, New York: Elsevier Scientific Publishing Company, 1973. pp. 75-93.

N°2. The Transfer of Management Know-How in Asia : An Unlearning Process.

Henri-Claude de BETTIGNIES. February 1981.
Reproduced from : Breaking down Barriers: Practice and Priorities for International Management Education.
Edited by Bob Garratt and John Stopford. London: Gower for the Association of Teachers of Management, 1980. pp. 293-310.

N°3. Korean Management in the 1980's : The International Challenge.

Henri-Claude de BETTIGNIES. February 1981.
Reproduced from : The Korean Journal of International Business. Vol. 1. International Management Institute, Korea University, Seoul, July 1980. pp. 119-125.

N°4. La Sociologie des organisations : Le cas du Japon

Henri-Claude de BETTIGNIES. February 1981.
Reproduced from : Les Etudes Japonaises en France.
Colloque, oct. 1979.
Paris : Association pour l'Etude de la Langue et la Civilisation Japonaise, 1980. pp. 118-130.

N°5. Analyse des craintes françaises.

Henri-Claude de BETTIGNIES. February 1981.
Reproduced from : Revue Française de Gestion. N° 27-28, sept-oct. 1980.
Numéro spécial : Le Japon Mode ou Modèle ? pp. 16-23.

N°6. L'Indonésie et ses Potentiels

Philippe LASSERRE. May 1981.
Reproduced from : Marchés Internationaux N° 51, mai 1981. pp. 83-98.

N°7. Transferts de Technologie : des mariages difficiles.

Philippe LASSERRE. May 1981.
Reproduced from : Revue Française de Gestion, N° 30 mars-avril 1981. pp. 97-103

N°8. The Industrialising Countries of Asia : Perspectives and Opportunities.

Philippe LASSERRE. July 1981.
Reproduced from : Long Range Planning Vol 14 N° 3, June 1981. pp. 36-43.

N°9. Le Japon prépare ses managers de demain à l'école de l'Occident.

Jacques de RUGY. July 1981.
Reproduced from : France Japon éco N°9, automne 1981, pp. 10-21.

N° 10. Quand les entreprises japonaises intègrent la gestion américaine.

Tetsuo AMAKO. July 1982.
Reproduced from : Revue Française de Gestion N° 35, mars-avril 1982, pp 59-63 + 10 p. annexes.

N° 11. Training : key to technological transfer.

Philippe LASSERRE. July 1982.
Reproduced from : Long Range Planning, vol 15 N° 3, June 1982. pp. 51-60

EURO-ASIA CENTRE

INSEAD

Institut Européen d'Administration des Affaires
European Institute of Business Administration
Europäisches Institut für Unternehmensführung
Institut Privé d'Enseignement Supérieur

Boulevard de Constance
77305 Fontainebleau Cedex, France
Telephone (6) 422 48 27 Telex 690389F

Update September 1982

EAC DOCUMENTATION SERIES

N°1. A bibliography on Japan: holding list of the Euro-Asia Centre and INSEAD library. May 1980.- 55 p.-

N°1 bis. A bibliography on Japan: resources of the Euro-Asia Centre and INSEAD library: new additions.- July 1981.- 28 p.-

N°2. A bibliography on ASEAN countries: holding list of the Euro-Asia Centre and INSEAD library. July 1980.- 79 p.-

N°3. A bibliography on East and Northeast Asia: holding list of the Euro-Asia Centre and INSEAD library. July 1980.- 30 p.-

N°4. A bibliography on the People's Republic of China: Resources of the Euro-Asia Centre and INSEAD library. October 1981.- 15 p.-

N°5. A bibliography on ASEAN and ASEAN countries : Resources of the Euro-Asia Centre and INSEAD Library. October 1981.- 77 p.-

N°6. A bibliography on South Korea, Hong Kong and Taiwan : Resources of the Euro-Asia Centre and INSEAD Library. January 1982.- 22 p.-

N°7. A bibliography on Asia : Resources of the Euro-Asia Centre and INSEAD Library. February 1982.- 25 p.-

N°8. A bibliography on Japan : Resources of the Euro-Asia Centre and INSEAD Library. July 1982.- 108 p.-

N°9. A bibliography on the People's Republic of China : Resources of the Euro-Asia Centre and INSEAD Library. August 1982.- 18 p.