RABIKAR CHATTERJEE, JEHOSHUA ELIASHBERG, HUBERT GATIGNON, and LEONARD M. LODISH*

The authors describe a methodology and a personal-computer-based decision model for selecting optimal market testing strategies. A Bayesian decision theoretic framework is used that (1) considers continuously distributed payoffs, (2) allows for updating of information on market response to strategies that are not being tested directly (relaxing the assumption of independence among test outcomes), and (3) incorporates managers' attitude toward risk. The goals of the methodology are to bring managers a practical, usable tool that will help support their design of market tests and to obtain some insights into the market testing problem. An application of the methodology is reported to illustrate the potential of the model as a practical and easily implementable marketing decision aid. Using the analytical insights obtained from the model, the authors summarize the influence of the various characteristics of the alternative strategies on the value of a market test and the choice of the test strategy.

A Practical Bayesian Approach to Selection of Optimal Market Testing Strategies

The practice of product testing is widespread in the field of marketing (Paskowski 1984) and its value has been well recognized by practitioners as well as academics (Pringle, Wilson, and Brody 1982; Silk and Urban 1978; Urban and Katz 1983). Go/no-go decisions typically are based on the information from such tests (Charnes et al. 1966; Paskowski 1984; Urban and Katz 1983).

More important than the go/no-go decision may be determining the best marketing strategy for a product. The importance of diagnostic information from pre-test models has been pointed out by Narasimhan and Sen (1983). The authors of pre-test models also have discussed the potential of their models as diagnostic tools for developing a better marketing strategy (Blattberg and Golanty 1978; Pringle, Wilson, and Brody 1982; Urban 1970). The evaluation of alternative strategies—a secondary objective of these models—is done through simulations, after a pre-test market model is built, or on the basis of parameters available from other products (Blattberg and Golanty 1978).

Typically, a firm has a set of mutually exclusive marketing strategies under consideration, for example, a discrete set of acceptable price levels or different distribution or promotion options, either for a new product or as alternatives to the present strategy for a current product. Marketing managers are often uncertain of the precise market response to these alternative strategies. They can either select the strategy expected to yield the greatest payoff or conduct a market test before choosing the strategy. We use the term “market test” to refer generally to any means of obtaining more information about the market response to a strategy. It could be a traditional test market, a simulated test market, or any other type of strategy evaluation method.

The manager should conduct a market test if the expected value of the information exceeds the cost of the test. The information has value if there is a probability that it can change the manager’s prior choice of strategy.

*Rabikar Chatterjee is Assistant Professor of Management, Purdue University. Jehoshua Eliashberg is Professor of Marketing, Hubert Gatignon is Associate Professor of Marketing, and Leonard M. Lodish is Professor and Chairman, Marketing Department, The Wharton School, University of Pennsylvania.

The authors thank Bob Herbold of Procter & Gamble, Alan Markowitz and Jim Smith of Travelers Mortgage Services, and Bob Hoffs of Bell Atlantic for helping to lend realism to this work. Valuable comments and suggestions made by the editor and anonymous JMR reviewers are also gratefully acknowledged.
The information obtained from the test depends on which strategy (from the set of alternatives) is used in that test. The problem is to select a market test strategy that maximizes the expected value of information (adjusted for the cost of the test). The decision maker’s attitude toward risk can be incorporated via a nonlinear utility function.

We present a model that enables a manager to evaluate alternative market test strategies and select the optimal strategy, given the manager’s current state of knowledge. We focus on the selection of a single optimal market test. More generally, if multiple market tests are feasible, the problem is to design an optimal market test plan that specifies the number of tests to be conducted and the strategy for each test. We discuss this issue subsequently in connection with possible extensions of the model.

We first present the Bayesian decision analytic framework used to conceptualize the problem. We then discuss implementation issues in the context of several actual applications using a personal-computer-based, interactive model, TESTER. The results of one typical application are reported. Next, we employ the model to derive some analytical insights. Finally, limitations and extensions of the model are considered.

THE BAYESIAN DECISION ANALYTIC MODEL

Bayesian analysis has long been suggested as a useful tool to guide market research (Bass 1963; Day et al. 1973; Green 1963). Indeed, most basic marketing research texts (e.g., Aaker and Day 1986; Green and Tull 1978; Lehmann 1985) introduce the Bayesian decision analytic framework as a basis for evaluating market research. Blattberg (1979) demonstrates how the framework can be applied to the design and evaluation of advertising experiments. Ginter et al. (1981) provide a theoretical extension of Blattberg’s analysis to incorporate statistical risk. Most firms, however, have limited their use of formal Bayesian analysis in marketing applications. Possible reasons are that the analysis typically is computationally cumbersome and various judgmental inputs are required to calibrate the models. The increased use of personal computers may alleviate these problems by enabling managers to use such models interactively.

In applying the Bayesian framework, the researcher must make a tradeoff. An approach involving few or no assumptions about the functional forms of the model components requires the elicitation of many judgmental inputs from the decision maker, then the fitting of a curve through those data points. For the decision problem we consider, such an assumption-free approach results in an arduous elicitation procedure, typically leading to user fatigue and apathy. In contrast, by imposing more structure on the problem and invoking widely accepted assumptions about the various functional forms, one can minimize the amount of time needed to obtain the necessary inputs. We believe that relatively easy estimation of the value of information from alternative testing strategies is desirable. Hence, in calibrating our model with inputs obtained interactively by TESTER, we strive to simplify the respondent’s task, consistent with a reasonable model of the situation. TESTER provides graphs of the assumed functions and probability distributions with parameters based on the inputs. In more than 10 applications, decision makers have not found them to be problematic descriptors of their judgments.

Our analysis considers (1) continuously distributed payoffs for each alternative marketing strategy, (2) non-independence of outcomes associated with the different strategies, implying that a market test employing one strategy may provide useful information not only about payoffs under that strategy, but also about payoffs under other (nontested) strategies, and (3) the decision maker’s attitude toward risk. For ease of exposition, we first present the model without considering attitude toward risk. This model can be viewed as a special case of the more general model, under the assumption that management is risk neutral. We then extend this model to incorporate the manager’s attitude toward risk.

Problem Definition

A firm is considering a set \( \{1, 2, \ldots, n\} \) of mutually exclusive marketing strategies either for a product launch or as alternatives to the present strategy for a current product. Examples include price levels, product profiles, advertising strategies, and distribution options. Some decision variables (such as price) are inherently continuous; however, managers commonly consider discrete levels of these variables when evaluating their strategic options. The strategy set includes the status quo option as the \( n^{th} \) alternative (i.e., “no go” for the product launch case or “present strategy” with often known market response for a current product), which serves as the baseline. The market response to each of these strategies (other than the status quo alternative) is uncertain.

Managers can either select a strategy without conducting a market test or conduct a test before choosing a strategy. The problem is to select which strategy (from the \( n-1 \) possibilities, excluding the status quo alternative), if any, should be tested before choosing the strategy for implementation. (If the market response to the current strategy is uncertain because of changed market conditions, all \( n \) options—including the current strategy—are candidates for testing.) Figure 1 is the decision tree corresponding to the problem. A complete list of notations is provided in the Appendix.

Because the manager is uncertain of the market response to the various strategies, the payoffs (profits) are uncertain and denoted by the random variable \( \pi_i \) (profit under strategy \( i \)).\(^1\) The prior probability density function of profit, \( f(\pi_i) \) \((i = 1, \ldots, n)\), represents the manager’s knowledge of the effectiveness of the various strategies.

\(^1\) The tilde denotes a random variable. For example, \( \tilde{\pi} \) is a random variable and \( \pi \) is a specific realization of that random variable.
before any test is conducted. Thus, if a strategy were to be chosen without further evaluation, the decision would be based on the manager’s prior distribution of $\tilde{\pi}_i$ for $i = 1, \ldots, n$.

We now consider the possibility of testing, and hence observing, additional information before the selection of a strategy for full-scale implementation. Figure 1 shows the chronological sequence of decisions (denoted by boxes) and events (circles). The manager first chooses a test strategy, say alternative $j$. Let $T_j$ denote the unbiased (or “full-scale”) projection of sales in units observed in the market test under strategy $j$. (The issue of systematic bias is discussed subsequently.) At this point, the outcome of the test is uncertain and hence denoted by the random variable $T_j$, with probability density function $f(T_j)$. Once the test is conducted and the outcome $T_j$ observed, the manager updates his or her knowledge about the effectiveness of various strategies, represented by the posterior probability density function $f(\tilde{\pi}|T_j)$, for $i = 1, \ldots, n$. This posterior distribution then becomes the basis for selecting the firm’s strategy.

To derive the posterior distribution of profit under alternative strategies, we must specify (1) the prior distribution of sales, (2) the reliability of the market test, and (3) the relationship (dependence) among outcomes under different strategies. These elements of the problem are discussed next.

Prior distribution of sales. Let $S_i$ denote the sales in units associated with strategy $i$. We capture the manager’s current knowledge about response to a particular strategy $i$ via the prior probability density function of $\tilde{S}_i$. More specifically, our analysis begins by assuming that the manager’s prior assessment of sales in response to strategy $i$ is represented by the normal distribution

$$
\tilde{S}_i \sim N(\mu_i, \sigma_i^2),
$$

where $\mu_i$ and $\sigma_i^2$ are the mean and variance, respectively, of the prior distribution of $\tilde{S}$.

The normal distribution assumption has been accepted widely as a robust model that can serve as a good approximation to many other distributions (Johnson and Kotz 1970; Winkler 1972). The normality assumption does imply that, conceptually, there is a nonzero probability of sales being negative, though in practice sales cannot be negative. However, the values of $\mu_i$ and $\sigma_i$ in actual applications are such that the coefficient of variation ($\sigma_i/\mu_i$) is sufficiently small that the possibility of $\tilde{S}_i$ being negative is negligible. For example, for the empirical application discussed in the next section, the order of the coefficients of variation implies that, under the normal distribution assumption, the probability of negative sales is less than $10^{-10}$. In practical situations where the coefficient of variation is not sufficiently small, the manager’s prior probability density function can be captured by a doubly symmetric truncated normal distribution (Johnson and Kotz 1970, p. 83) instead of the complete normal distribution. The analysis proceeds in an almost identical manner as with the complete normal distribu-
tion, except for the incorporation of adjustment factors for the variance and inflator factors to ensure proper probability densities. The normal distribution also implies symmetry about the mean. An extension of our framework to incorporate a more flexible (possibly skewed) distribution over non-negative values of sales is conceptually possible. However, it involves computational complexities and a more tedious assessment procedure, especially in our analysis of the impact of testing on the evaluation of nontested strategies (requiring consideration of correlations between outcomes under the different strategies). Consequently, we adhere to the decision calculus model-building criteria proposed by Little (1970) and strive for managerial acceptance.

Reliability of a Market Test. We view $T_j$ as an imperfectly reliable (but unbiased) measure of the “true” sales $S_j$. (Note that the “scaling up” of test sales ensures that $T_j$ and $S_j$ are specified on a comparable basis.) Thus, we have the familiar measurement error model

$$\hat{T}_j = \hat{S}_j + \hat{\epsilon}_j,$$

where $\hat{\epsilon}_j$ denotes the “measurement” or testing error. We assume that (1) the random error term is normally distributed, $\hat{\epsilon}_j \sim N(0, \delta_j^2)$, and (2) $\hat{S}_j$ and $\hat{\epsilon}_j$ are independent, as are $\hat{S}_j$ and $\hat{\epsilon}_j$ (for $i \neq j$).

Hence, the full-scale unbiased projection of test sales, given an “actual” sales level $S_j$, is normally distributed with mean $\hat{S}_j$ and variance $\delta_j^2$. The variance $\delta_j^2$ (or the standard deviation $\delta_j$) can be interpreted as an inverse measure of the reliability of the market test outcome as a predictor of actual sales.

In practice, the market test outcome may provide a biased estimate of sales; for example, tests often overproject sales, possibly as a result of a Hawthorne effect. Other factors (such as unusual levels of competitive activity) also may affect the test results. To the extent the manager can assess a systematic bias attributable to these factors, the actual test results can—and should—be adjusted for that bias. Thus, if $T_j^*=T_j+b_j$ is the observed market test outcome and $b_j$ is the assessed systematic bias, then

$$T_j = T_j^* - b_j.$$

Note that the uncertain (or “random”) element of the impact of these factors becomes a component of the random error term, $\hat{\epsilon}_j$, and reduces the reliability of the test.

Non-independence of outcomes. In general, information about response to a particular strategy may influence the state of knowledge about the market response to that strategy, as well as alternative strategies. For example, given information about sales response to a $2.99 per unit price level, managers would be likely to revise their assessment of sales under a $1.49 per unit pricing strategy. The implication is that, for any pair of strategies, $\{i,j\}$, the corresponding (uncertain) sales, $\hat{S}_i$ and $\hat{S}_j$, may be correlated. Let $\rho_{ij}$ denote the correlation between $\hat{S}_i$ and $\hat{S}_j$, for $i, j = 1, \ldots, n$. In practice, we would usually expect $\rho_{ij}$ to be non-negative, though conceptually the model places no restrictions on the sign of $\rho_{ij}$ (for $i \neq j$). We assume that the joint distribution of $\hat{S}_i$ and $\hat{S}_j$ (for $i \neq j$) and the joint distribution of $\hat{T}_j$ (recall that $\hat{T}_j = \hat{S}_j + \hat{\epsilon}_j$) are bivariate normal.

The parameters of the prior distribution of sales ($\mu_i, \sigma_i^2$), the reliability of the market test ($1/\delta_j^2$), and the pairwise correlations between sales associated with different strategies ($\rho_{ij}$) are either obtained directly or derived from inputs elicited from managers.

Evaluation of Alternative Test Strategies

Given our specification of the problem, we can now evaluate the alternative test strategies represented by the branches of the decision tree in Figure 1.

Consider first the “no-test” option (node 1A in Figure 1). If a strategy were to be selected without additional evaluation, the strategy $i^*$ with the highest expected profit should be selected. That is, the decision is to choose $i^*$ such that

$$E[\hat{\pi}_i] = \max_i E[\hat{\pi}_i], \quad i = 1, \ldots, n,$$

with

$$E[\hat{\pi}_i] = P_i E[\hat{S}_i] - A_i = \mu_i P_i - A_i.$$

where:

- $P_i$ = the unit margin associated with strategy $i$ and
- $A_i$ = the fixed cost associated with strategy $i$.

Note that the margins and fixed costs are assumed to be known. In practice, managers are more likely to be uncertain about sales response than about costs. Uncertainty about the cost structure, however, can be accommodated by the framework. We next turn to the “test” option. To determine which strategy (if any) should be tested, we must evaluate the branches representing the test alternatives by “folding back” the decision tree.

Posterior analysis. We begin at node 2 of the decision tree in Figure 1. Assume that some strategy $j$ has been tested and the market test outcome, $T_j$, observed. In this case, the optimal decision (at node 2) is to select strategy $i^*$ that maximizes expected profit given the test outcome:

$$E[\hat{\pi}_i|T_j] = \max_i E[\hat{\pi}_i|T_j], \quad i = 1, \ldots, n.$$

The moments of the distribution of the posterior profit, $E[\hat{\pi}_i|T_j]$, are derived by Chatterjee et al. (1986, Appendix A). In particular, we note that $E[\hat{\pi}_i|T_j]$ is a linear function of $T_j$ and can be expressed as

$$E[\hat{\pi}_i|T_j] = \alpha_i + \beta_i T_j,$$

with

$$\alpha_i = \mu_i P_i - \frac{\rho_{ij} \sigma_i \sigma_j P_i}{\delta_j^2 + \sigma_j^2} - A_i.$$


and
\[
\beta_j = \frac{\rho_j \sigma_i \sigma_j P_j}{\sigma_i^2 + \sigma_j^2}.
\]
Note that when \( i = j, \rho_j = \rho_j = 1.\)

**Preposterior analysis.** We now can evaluate the branches corresponding to the alternative strategies. The expected profit given that strategy \( j \) will be tested (and thus, \( T_j \) observed prior to choosing the strategy) is given by
\[
E[\pi|\text{Test}_j] = \int_0^\infty [\max\{a_j + \beta_j T_j\}]f(T_j)dT_j
\]
where \( f(T_j) \) is the probability density function of \( T_j. \) Note that in the computation of the expected profit one assumes that the optimal strategy will be implemented once \( T_j \) is observed.

In general, the optimal strategy for full-scale implementation depends on the specific outcome of the test. We therefore must identify which strategy is best for different values of \( T_j \) over its entire range. More formally, we divide the range of \( T_j \) into a set of mutually exclusive and collectively exhaustive segments such that, for each segment, one particular strategy (identified by equation 7) is best over the entire segment.

**Determining the optimal market test strategy.** Once the expected profits have been determined, we can compare the test options. The expected profit in each case must be adjusted for the cost of testing. Thus, the optimal decision (at node 1B of the decision tree) is to select strategy \( j^* \) for the market test such that
\[
E[\pi|\text{Test}_j] - C_j = \max_j [E[\pi|\text{Test}_j] - C_j],
\]
where \( C_j \) is the cost of test \( j. \)

Note that the expected profit in the absence of testing is given by \( \pi \) (see equation 5). Thus, the optimal decision (at node 1) is to conduct a market test (employing strategy \( j^* \)) if and only if
\[
E[\pi|\text{Test}_j] - \max_i E[\pi_i] > C_{j^*},
\]
otherwise, the firm is better off not conducting a market test at all. Equivalently, the strategies can be evaluated by setting the “do not test” alternative as the datum and computing the expected value of information obtained from a test employing strategy \( j, \) given by
\[
EVI(\text{Test}_j) = E[\pi|\text{Test}_j] - \max_i E[\pi_i],
\]
for each \( j, \) \( EVI(\text{Test}_j), \) which captures the benefit of testing strategy \( j, \) can then be compared with the cost of the test, \( C_j. \)

**Incorporating Risk**

In our analysis to this point we have ignored the effect of risk on the manager’s preference for different alternatives. In practice, managers are typically risk averse in decision making under uncertainty (Swalm 1966). We employ von Neumann-Morgenstern utility theory (von Neumann and Morgenstern 1947) to model the effect of risk. Specifically, we assume the utility function
\[
u(\pi) = 1 - r \exp(-r\pi)
\]
where \( \pi \) denotes uncertain payoffs (profits) and \( r \) is the coefficient of risk aversion (Pratt 1964). The behavioral implication of the assumed form of the utility function is that the decision maker displays constant risk aversion over the range of payoffs. The exponential utility function has the virtues of empirical support as a reasonable representation of decision makers’ preference under uncertainty (Howard 1971), analytical tractability, and ease of calibration.

If \( \pi \) is normally distributed, the expected utility is
\[
E[u(\pi)] = 1 - \exp\left[-rE[\pi] + \frac{r^2}{2}\text{Var}[\pi]\right].
\]

The **certainty equivalent** is defined as the value of a certain payoff that makes the decision maker indifferent between that payoff and the strategy with uncertain payoffs. For the exponential utility function (equation 15), the certainty equivalent is given by
\[
CE[\pi] = -\frac{1}{r} \ln(1 - E[u(\pi)]) = E[\pi] - \frac{r}{2} \text{Var}[\pi].
\]
where the second equality follows from equation 16.

When preference (or value) under certainty is measured in terms of profit, the corresponding measure on the same dollar scale of preference under uncertainty is the certainty equivalent (CE). (Note that expected utility monotonically increases in CE; that is, the preference ordering of alternatives is identical under the expected utility and certainty equivalent criteria.) Note also, from equation 17, that the certainty equivalent incorporates a downward adjustment to the expected profit to account for the effect of risk. The extent of the downward adjustment \( \left(\frac{r}{2}\text{Var}[\pi]\right) \) depends on the coefficient of risk aversion, \( r \) (larger values of \( r \) imply greater risk aversion), and the amount of uncertainty in the payoff, captured by \( \text{Var}[\pi]. \) In evaluating strategies under risk aversion, we consider the **certainty equivalents** of uncertain profits instead of the expected values.

Under the “no-test” option, the strategy \( i^* \) with the highest certainty equivalent should be selected. That is, the decision is to choose \( i^* \) such that

\[\text{Note that the profits represent periodic (e.g., annual) inflows whereas the market test costs typically are incurred as one-time expenditures. If managers have a multiyear planning horizon, they must consider the net present value of the expected profit stream, given the length of the planning horizon and the discount rate.}\]
A market test (employing strategy $j^*$) should be conducted if and only if
\[ CE[\hat{\pi}_{j^*}] - \text{Max}_j \{ CE[\hat{\pi}_j] \} > C_r. \]

If that condition is not met, the firm is better off not conducting a market test. Corresponding to the expected value of information in the risk neutral case (equation 14) is the "cash equivalent of information" obtained from a test of strategy $j$, defined as
\[ CE(T_{test}) = CE[\hat{\pi}_{j_{test}}] - \text{Max}_j \{ CE[\hat{\pi}_j] \}. \]

For each $j$, $CE(T_{test})$ can then be compared with the cost of the test, $C_r$.

**ILLUSTRATIVE APPLICATIONS OF THE MODEL**

We now examine issues in implementing the TESTER computer-based model, focusing on one application in detail. The major inputs are managers' prior judgments about the distribution of sales and the reliability of the various test markets. There is a vast body of literature on assessment of subjective probabilities (see Wallsten and Budescu 1983 for a review). In marketing, Woodruff (1972) and Gatignon (1984) have implemented methods for direct assessment of probability distributions. The approach taken here is consistent with the decision calculus philosophy (Little 1970; Little and Lodish 1981) and probability distribution assessment for Bayesian analysis (Leamer 1978; Winkler 1967a,b). Managers' judgments are assessed and sensitivity analysis is performed on the parameters for which the assessment uncertainty may be great.

To assess the prior probability distribution, two points of the distribution (fractiles) are assessed and the parameters of the normal distribution probability density function can be determined. For example, the decision maker may be asked:

1. What is an optimistic (only 2.5% chance of exceeding) national sales projection in units in the second year of implementation of the strategy? Answer: $X_{0.025}$.
2. What is a pessimistic (97.5% chance of exceeding) national sales projection in units in the second year of implementation of the strategy? Answer: $X_{0.975}$.

The mean $\mu_i$ is estimated by $(X_{0.025} + X_{0.975})/2$ and the standard deviation $\sigma_i$ by $(X_{0.975} - X_{0.025})/4$. Note that .025 and .975 fractiles are illustrative; other values can be used. Further, the mean $\mu_i$ can be assessed directly by asking: "What is your expectation of the national sales projection in units in the second year of implementation of the strategy?" The answer provides a third data point, which can be used to assess the validity of the symmetry assumption. (We discuss this issue more fully in the Conclusions section.)

To assess the reliability of the test market, managers are asked to assume that the level of sales in units would be the mean of the prior distribution and that, if the test were perfectly reliable, the nationally projected test outcome would exactly equal this value. Then they are asked...
to provide an optimistic and a pessimistic estimate of the test market sales outcome, projected to the national (or full-scale) level, in units. The standard deviation $\delta_j$, an inverse measure of the test reliability, is determined similarly to $\sigma_j$.  

TESTER has been used by decision makers in several industries including pharmaceuticals, mortgage services, and telecommunications. For expository purposes, we report a typical application in which TESTER evaluated two possible advertising programs for a well-established consumer packaged good. The user, the director of marketing research, had recently been a brand manager for this product, which we call “Old Reliable.” Figure 2 shows the input data needed to evaluate the two strategic options: an advertising budget of $38 million versus a lower budget of $29 million (all data have been multiplied by a constant for confidentiality). This manager had greater uncertainty about his sales estimate for the lower advertising strategy. The tests considered for evaluation were instrumented markets. In these markets, scanner panel members would be exposed to different advertising levels for a 1-year period. The statistical reliability of the tests depended on the sample size of the panels and the market share of Old Reliable. Reliability estimates given by the firm providing the instrumented markets were based on analysis of the various experimental results in their experience base.

To assess the correlations $\rho_{ij}$, the following question is asked.

Assume you marketed the product under strategy $i$ and found that the sales were 10% higher than your current expectation. To what extent would you revise your expectation of sales under strategy $j$?

The conditional expectation of $S_j|S_i$ is given by (Mood, Graybill, and Boes 1974, p. 167):

$$E[S_j|S_i] = \mu_j + \rho_{ij} \frac{\sigma_j}{\sigma_i} (S_i - \mu_i).$$

From the preceding question, $S_i = 1.1\mu_i$. Therefore,

$$E[S_j|S_i = 1.1\mu_i] = \mu_j + \rho_{ij} \frac{\sigma_j}{\sigma_i} (0.1\mu_i).$$

If the manager’s response to the question is $R(%)$, the correlation $\rho_{ij}$ is given by:

$$\rho_{ij} = \frac{R \sigma_j \mu_j}{10 \sigma_i \mu_i}.$$

The correlation derived from the responses of the marketing research director by equation 29 is .32. All units are cases of Old Reliable. For this first run, risk attitude was not considered.

The TESTER evaluation and recommendations (see Table 1) show that only a test of the lower advertising strategy is worth the $250,000 total test cost. The expected value of a test of this strategy is $473,446 or $223,446 net of the cost. Note that the test market cost does not include the potential adverse impact of competitive response. If this impact could be quantified in dollar terms, it could be incorporated in the cost. The possibility of competitive interference would adversely affect the test reliability and thereby discount the net value of the test information.

TESTER concluded that if the test projected national sales greater than 9,848,321 units, the lower advertising level should be adopted; otherwise, the higher level should be used.

In the next analysis, the manager’s attitude toward risk was assessed with a procedure proposed by Elishberg and Hauser (1985). A series of 10 simple binary choice problems is presented sequentially to the manager, who chooses the preferred alternative in each case. The profit associated with either alternative is uncertain. In the first alternative, there is a probability $\alpha$ of obtaining a profit level $x_1$ and a probability $1 - \alpha$ of obtaining zero profit. For the second alternative, there is a probability $\beta$ of obtaining a profit level $x_2$ and a probability $1 - \beta$ of obtaining zero profit. The probabilities and profit levels are specified in each choice problem. The parameters ($\alpha$, $\beta$, $x_1$, and $x_2$) are varied across the problems and their values chosen so that the following relationship holds:

$$\beta \exp(-r_x x_2) - \alpha \exp(-r_x x_1) = \beta - \alpha,$$

where $r_x$ is a positive constant and $\beta > \alpha$.

The risk aversion parameter $r$ is estimated by (Elishberg and Hauser 1985):

$$r = \frac{r_x}{\ln(1 - n_i/10)},$$

where $n_i$ is the number of times the first alternative is chosen in preference to the second. The utility function (equation 15) for the decision maker therefore is estimated as:

$$u(\bar{r}) = 1 - \exp(-P\bar{r}).$$

In this application, the estimated parameter $\bar{r}$ indicated a high degree of risk aversion.

When the manager’s risk aversion was incorporated, the recommendation (to test the lower advertising strategy) remained unchanged (see Table 2). However, the value (cash equivalent) of the market test increased to $511,058 (in comparison with $473,446 when risk attitude was not considered). In this case, TESTER recommended that the lower advertising strategy be adopted only if the national sales projection of the test outcome exceeds 9,832,588 units (in comparison with 9,848,321 units).

These results illustrate the impact of risk attitude on decision making. Once the test outcome is observed, the
Alternative 1: High Advertising Expenditure

<table>
<thead>
<tr>
<th>NATIONAL LEVEL OF advertising</th>
<th>$38000000.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>NATIONAL LEVEL OF PRICE IS</td>
<td>$29.00</td>
</tr>
</tbody>
</table>

OPTIMISTIC (ONLY 2.5% CHANCE OF EXCEEDING) NATIONAL SALES PROJECTION in units
FOR old.reliable1 FOR SECOND YEAR OF IMPLEMENTATION OF STRATEGY WITH
advertising AT $38000000.00 AND PRICE AT $29.00? 10900000
PESSIMISTIC (97.5% CHANCE OF EXCEEDING) NATIONAL SALES PROJECTION in units
FOR old.reliable1 FOR SECOND YEAR OF IMPLEMENTATION OF STRATEGY WITH
advertising AT $38000000.00 AND PRICE AT $29.00 9900000

ASSUME THAT THE NATIONAL SALES in units ACTUALLY ARE 10400000.

Now, we shall try to elicit your feelings concerning the reliability of the test marketing.

If the test were perfectly reliable, we would observe local results that would be nationally projected to exactly 10400000 units. Given that the test represents only a sample of the total market, the national projection of the test will not necessarily be the same as what is assumed.

PROVIDE AN OPTIMISTIC AND PESSIMISTIC ESTIMATE OF WHAT THE TEST MARKET SALES COULD CONCEIVABLY PROJECT NATIONALLY in units UNDER THE ASSUMPTION STATED ABOVE:
OPTIMISTIC? 10825000
PESSIMISTIC? 9975000

ENTER THE COST OF THE TEST MARKET? 250000

Alternative 2: Low Advertising Expenditure

<table>
<thead>
<tr>
<th>NATIONAL LEVEL OF advertising</th>
<th>$29000000.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>NATIONAL LEVEL OF PRICE IS</td>
<td>$29.00</td>
</tr>
</tbody>
</table>

OPTIMISTIC (ONLY 2.5% CHANCE OF EXCEEDING) NATIONAL SALES PROJECTION in units
FOR old.reliable1 FOR SECOND YEAR OF IMPLEMENTATION OF STRATEGY WITH
advertising AT $29000000.00 AND PRICE AT $29.00? 10800000
PESSIMISTIC (97.5% CHANCE OF EXCEEDING) NATIONAL SALES PROJECTION in units
FOR old.reliable1 FOR SECOND YEAR OF IMPLEMENTATION OF STRATEGY WITH
advertising AT $29000000.00 AND PRICE AT $29.00 9550000

ASSUME THAT THE NATIONAL SALES in units ACTUALLY ARE 10175000.

Now, we shall try to elicit your feelings concerning the reliability of the test marketing.

If the test were perfectly reliable, we would observe local results that would be nationally projected to exactly 10175000 units. Given that the test represents only a sample of the total market, the national projection of the test will not necessarily be the same as what is assumed.

PROVIDE AN OPTIMISTIC AND PESSIMISTIC ESTIMATE OF WHAT THE TEST MARKET SALES COULD CONCEIVABLY PROJECT NATIONALLY in units UNDER THE ASSUMPTION STATED ABOVE:
OPTIMISTIC? 10600000
PESSIMISTIC? 9750000

ENTER THE COST OF THE TEST MARKET? 250000
risk averse manager should be more likely to select the less risky strategy. After the test of the lower advertising strategy, the (posterior) uncertainty associated with this strategy becomes lower than that of the other strategy. Therefore, under risk aversion, the “critical value” of the test outcome (above which adoption of the lower advertising strategy is recommended) is lower in relation to the risk neutral case.

Other uses of TESTER include evaluations of a new form of an orange drink, different price levels for a pharmaceutical product, and distribution systems for a new mortgage service. In these applications the value of information was higher once the managers’ risk attitude was considered. Incorporating risk attitude also changed the cutoff point(s) of the test outcome (“critical value(s)”) determining the optimal strategy to be implemented once the test result is observed. Overall, our experience with TESTER suggests that managers are comfortable with the model calibration procedure and the recommendations obtained from the analysis.

**THE VALUE OF A TEST—SOME ANALYTICAL INSIGHTS**

The Bayesian model has several analytical implications in terms of the expected value or cash equivalent of information obtained from testing a particular strategy. For ease of exposition, we consider two strategies, each with uncertain market response. Our objective is to obtain some managerially useful insights into the impact of the following key elements of our model on the value of information: the, prior expectation of profitability as well as the uncertainty associated with either strategy, the test reliability, and the correlation between sales under the two strategies. These insights suggest conditions under which a market test may or may not be worthwhile.

In the risk neutral case, strategies are evaluated in terms of expected profit; when risk aversion is explicitly considered, the criterion for evaluation is the certainty equivalent of profit. Similarly, under risk aversion, we evaluate information from a test in terms of the cash equivalent of information instead of the expected value of information used in the risk neutral case. The cash equivalent of information represents the maximum a manager should pay for the market test. (Note that the certainty equivalent of profit and the cash equivalent of information reduce to the expected profit and the expected value of information criteria, respectively, in the risk neutral situation, i.e., when the manager’s coefficient of risk aversion is zero.) Because the implications under the risk aversion and risk neutral cases are similar, our discussion applies to both. The following results are derived formally by Chatterjee et al. (1986).

**Result 1.** The expected value (cash equivalent) of information from a test of either strategy is greater when (1) the prior expected profits (certainty equivalent of profits) under the two strategies are closer to each other, (2) the test is more reliable, and (3) the prior uncertainty of sales under the strategy to be tested is high.

This result is intuitive. A market test is most valuable when the two strategies are roughly equally attractive (in terms of prior expected profit or, under risk aversion, prior certainty equivalent of profit). When one strategy is clearly superior to the other, a market test is expected to be less worthwhile because the optimal strategy for implementation is unlikely to change. This implication is consistent with recommendations on testing in the literature (e.g., Urban and Hauser 1980). We also would expect to learn more if the test were more reliable or if the prior uncertainty of the strategy being tested was higher.

Result 1 pertains to the value of testing a particular strategy. To decide which of the two strategies (if either) should be tested, we would like to know which strategy has the higher expected value (cash equivalent) of information associated with its test. The answer depends

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Expected value without a test</th>
<th>Expected value with test before test cost</th>
<th>Expected value with test after test cost</th>
<th>Expected value of information after test cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Higher level</td>
<td>151,200,000</td>
<td>155,111,330</td>
<td>154,861,329</td>
<td>-163,670</td>
</tr>
<tr>
<td>2 Lower level</td>
<td>155,025,000</td>
<td>155,498,446</td>
<td>155,248,446</td>
<td>223,446</td>
</tr>
</tbody>
</table>

Recommended strategy to test: alternative 2 (lower level).
Maximal value of information (expected value) = 473,446.

**TESTER recommendations**

TESTER recommends testing strategy 2 (lower level). Then:
If the market test sales are greater than 9,848,321, use strategy 2 (lower level).
If the market sales are less than 9,848,321, use strategy 1 (higher level).
Table 2
EVALUATION OF ALTERNATIVE TESTS AND RECOMMENDATIONS CONSIDERING ATTITUDE TOWARD RISK

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Certainty equivalent without a test</th>
<th>Certainty equivalent with test before test cost</th>
<th>Certainty equivalent with test after test cost</th>
<th>Cash equivalent of information after test cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Higher level</td>
<td>150,777,955</td>
<td>154,485,292</td>
<td>154,235,292</td>
<td>−130,263</td>
</tr>
<tr>
<td>Expected utility</td>
<td>.9787070425</td>
<td>.9806299581</td>
<td>.98050933</td>
<td></td>
</tr>
<tr>
<td>2 Lower level</td>
<td>154,365,555</td>
<td>154,876,613</td>
<td>154,626,613</td>
<td>261,058</td>
</tr>
<tr>
<td>Expected utility</td>
<td>.9805706552</td>
<td>.9808225106</td>
<td>.980699718</td>
<td></td>
</tr>
</tbody>
</table>

Recommended strategy to test: alternative 2 (lower level). Maximal value of information (cash equivalent) = 511,058.

**TESTER recommendations**

If the market test sales are greater than 9,832,598, use strategy 2 (lower level).

If the market sales are less than 9,832,598, use strategy 1 (higher level).

on (1) the prior uncertainty in market response (sales), (2) the unit contributions, (3) the market test reliabilities, and (4) the correlation between sales under the two strategies. In particular, we have the following intuitive result.

Result 2. If one strategy is more uncertain in terms of both sales and profit and the test of that strategy is at least as reliable as the test of the other strategy, a test of the more uncertain strategy will have the greater expected value (cash equivalent) of information.

Note that the prior uncertainty in terms of profit depends on the prior uncertainty in sales and the unit contribution. Managers often are faced with the problem of choosing between strategies roughly similar in terms of prior expected profit (prior certainty equivalent of profit under risk aversion), but where one strategy has a higher unit contribution as well as higher fixed costs. If the strategies are about equally uncertain in terms of market response (sales), the high contribution strategy (which is riskier in terms of profitability) should be tested. In other situations (where the conditions in result 2 do not hold), the choice of the optimal test strategy involves a tradeoff and the criterion in footnote 4 can be employed.

It is the prior uncertainty, not the prior expectation (certainty equivalent) of profit, that determines which strategy (if any) should be tested, though the difference between the prior expectations (certainty equivalents) affects the expected value (cash equivalent) of information of both tests. If this difference is very large, neither test may be worthwhile. If the more uncertain strategy is considered for testing, the impact of the prior uncertainty of the other strategy and of the correlation between sales under the two strategies on the value of the test (in addition to the factors considered in result 1) is:

Result 3. If the strategy having the higher prior uncertainty of profit is considered for testing, the expected value (cash equivalent) of information from the test is greater when (1) the prior uncertainty of profit associated with the other strategy is low and (2) the correlation between sales under the two strategies is low.

The intuition behind this result is less obvious. The value of a test is related directly to the likelihood that the test will change the optimal strategy selection decision or, equivalently, the likelihood that the order of the strategies in terms of expected (certainty equivalent of) profits will be reversed when the test result is incorporated to update the priors. This likelihood depends on the differential impact of the test outcome on the posterior expectations (certainty equivalents) of profits under the two strategies. The differential impact is the greatest when the prior uncertainty of profit under the nontested strategy and the correlation are low.

Thus, an “ideal” scenario, in which the expected value (cash equivalent) of information is likely to be very high, is one in which the two strategies are very similar in terms of prior expectation (certainty equivalent) of profit, but very different in terms of prior uncertainty of profit. In this situation, a test of the more risky strategy should be very worthwhile.

The manager’s attitude toward risk may have a less significant role in determining a priori the choice of the optimal testing strategy than in determining the choice...
of the optimal strategy for full-scale implementation (possibly after a test). In other decision-making contexts (see, e.g., Eliashberg and Winkler 1978, 1981 for competitive and cooperative decision-making contexts), the effects of nonlinear utility functions have been found to be significant in altering overall strategic behavior.

CONCLUSIONS

We present a Bayesian model to evaluate market tests of alternative strategies and identify which strategy (if any) should be tested. Our model incorporates continuous distributions of uncertain outcomes that represent the problem realistically and minimize the amount of information researchers must obtain from managers. The model also explicitly takes into account the non-independence of market responses to various alternative strategies—an important consideration because the market responds to variables that are not tested explicitly and are common across strategies. For example, when one is considering alternative pricing strategies for a new product, the product itself is common across strategies. In such situations, it is imperative that the prior distributions of payoffs associated with nontested strategies be revised on the basis of the outcome of a test of a particular strategy. Finally, the model explicitly considers the decision maker’s attitude toward risk.

The empirical application report illustrates the approach to eliciting responses from the manager to calibrate the model and provides an example of the information that can be gained from TESTER. This application, and others not reported in detail here, demonstrate the feasibility of using such a decision aid, based on a Bayesian decision analytic approach, with an interactive program on a personal computer. Managers have found the experience of interacting with TESTER useful. In particular, the interactive process triggers questions that yield new insights into the decision problem.

Future Research

Further empirical experience with TESTER and feedback from managers may suggest refinements in the elicitation procedures currently used for model calibration. To keep the respondent’s task as simple as possible, we chose to seek only the minimal essential inputs for model calibration. The normal distributions for prior sales and market test outcomes under different strategies, based on the inputs, are presented to the manager by the interactive model for an assessment of their face validity. Furthermore, sensitivity analysis can be performed to determine the sensitivity of TESTER’s recommendations to the input parameters. It is possible to elicit additional points on the distribution and thereby obtain a goodness-of-fit measure to assess the validity of the normal distribution assumption and/or the reliability of the parameters. In practical terms, however, we believe the face validity check and sensitivity analysis are more appropriate in our context. In our applications, the normal distribution assumption appears robust and the potential limitations (allowance of negative values and symmetry about the mean) have not posed a problem.

An approach to enhancing the model would be to develop an option that relaxes some of the assumptions about the functional forms. This option would require deeper questions for calibration, but would be used only when sensitivity analysis indicates more detailed data are needed. It would leverage the methodology without imposing excessive respondent burden.

Further extensions of the model include the analysis of multiple market tests (under different strategies) and the incorporation of uncertain cost structures. The model currently focuses on the selection of a single optimal market test. In practice, it may be possible to conduct market tests of several strategies. Such multiple tests are frequently observed, with companies experimenting, for example, with different price levels or advertising strategies in different test markets. Thus the problem can be generalized to one of designing an optimal market test plan that specifies the number of tests to be conducted simultaneously and the strategy to be tested. The problem also can be modified to consider sequential testing plans.

Our model currently considers a finite set of strategies, with discrete levels of marketing mix variables. The model can be extended to incorporate a continuous strategy space for variables such as price or advertising expenditure. Such a model would employ a sales response function whose parameters (e.g., price sensitivity) are uncertain. Identification of the optimal strategy would be based on maximization of the (continuous) objective function incorporating these parameters.

Finally, future research could consider a formal evaluation of the benefits of TESTER from a normative perspective. Fudge and Lodish (1977) provide one of the few examples of a formal test of a decision calculus model, demonstrating how the use of the model can improve profitability. Similar research—applicable to any normative model—would provide a rigorous assessment of the value of TESTER in improving management decisions.

APPENDIX

LIST OF NOTATIONS

Variables

\[
\begin{align*}
S_i & : \text{sales in units associated with strategy } i (i = 1, \ldots, n, \text{ where } n \text{ denotes the status quo option}). \\
\pi_i & : \text{profit in dollars associated with strategy } i. \\
T_j & : \text{outcome of the market test employing strategy } j, \text{ expressed as projected sales in units at the national (or full-scale) level } (j = 1, \ldots, n - 1, \text{ assuming that the status quo option is not considered for testing}).
\end{align*}
\]

Random variables are denoted by the tilde sign; for
example, $S_i$ is a random variable whereas $S_j$ is a specific realization of that random variable.

**Parameters**

Parameter values are obtained/derived from inputs provided by the manager(s).

**Price/cost parameters.**

$P_i$: unit margin associated with strategy $i$.

$A_i$: fixed costs associated with strategy $i$.

$C_i$: cost of market test employing strategy $j$.

**Judgmental parameters.**

\[ \mu_i, \sigma_i^2 \]: mean and variance of prior distribution of $S_i | S_j \sim N(\mu_i, \sigma_i^2)$. \[ \delta_i^2 \]: variance of the conditional distribution of $T_i | S_j$: $1/\delta_i$ is a measure of the reliability of $T_i | S_j \sim N(S_j, \delta_i^2)$. \[ \rho_{ij} \]: correlation between $S_i$ and $S_j$ ($\rho_{ii} = 1$).

**Risk preference parameter.**

$r$: coefficient of risk aversion (measure of the decision maker’s risk attitude).

**Functions/Operations**

$f(\cdot)$: probability density function (p.d.f.).

$\phi(\cdot)$, $\Phi(\cdot)$: standard normal p.d.f. and c.d.f.

$F(\cdot)$: utility function.

$E[\cdot]$: expected value.

$\text{Var}[\cdot]$: variance.

$CE[\cdot]$: certainty equivalent.

**Value of Information**

$\text{EVI(}T\text{est)}$: expected value of information from a test of strategy $j$.

$\text{CEI(}T\text{est)}$: cash equivalent of information from a test of strategy $j$.

REFERENCES


Reprint No. JMR254103

AMERICAN MARKETING ASSOCIATION

REPRINT POLICY AND PROCEDURE FOR JOURNAL ARTICLES

The Prices below apply to each title ordered. An assortment of titles may be combined, however, to reach the minimum order amount of $5.00.

<table>
<thead>
<tr>
<th>Minimum order is $5.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–4 reprints (same article or an assortment)</td>
</tr>
<tr>
<td>5–99 reprints (same article)</td>
</tr>
<tr>
<td>First 100 reprints (same article)</td>
</tr>
<tr>
<td>Additional reprints up to 499 copies</td>
</tr>
<tr>
<td>Example: 101 reprints = $100.75</td>
</tr>
<tr>
<td>First 500 reprints (same article)</td>
</tr>
<tr>
<td>Additional reprints up to 999 copies</td>
</tr>
<tr>
<td>Example: 501 reprints = $405.50</td>
</tr>
<tr>
<td>First 1,000 reprints (same article)</td>
</tr>
<tr>
<td>More than 1,000 reprints (same article)</td>
</tr>
<tr>
<td>Orders over $1,000 - no charge for delivery</td>
</tr>
</tbody>
</table>

Prices subject to change.

There is NO RETURN and NO EXCHANGE on reprints.

Under the "fair use" provision of the copyright law effective January 1978, anyone may make a photocopy of a copyrighted article for his or her own use without seeking permission.

Also, a single copy reprint or an order of less than 50 copies may be obtained from University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Articles are priced prepaid at $17.50 plus $2.25 for each additional copy of the same article. Those who have a standing account with UMI pay $7.50 per article plus $2.25 for each additional copy of the same article. Complete issues are obtainable at $35.00 each.

To obtain permission to reproduce one’s own reprints in quantity, please contact the Permissions Department, American Marketing Association, 250 S. Wacker Drive, Chicago, IL 60606.

This content downloaded from 128.91.110.146 on Mon, 15 Apr 2013 10:25:22 AM
All use subject to JSTOR Terms and Conditions