1-1-2004

# A Bayesian Network Estimation of the Service-Profit Chain for Transport Service Satisfaction

Ronald D. Anderson

Robert D. Mackoy
*Butler University*, rmackoy@butler.edu

# A Bayesian Network Estimation of the Service-Profit Chain for Transport Service Satisfaction

**Anderson Ronald**
**Robert Mackoy**
**Vincent B. Thompson**
**Gilber Harrell**

## Abstract

Bayesian network methodology is used to model key linkages of the service-profit chain within the context of transportation service satisfaction. Bayesian networks offer some advantages for implementing managerially focused models over other statistical techniques designed primarily for evaluating theoretical models. These advantages are (1) providing a causal explanation using observable variables within a single multivariate model, (2) analysis of nonlinear relationships contained in ordinal measurements, (3) accommodation of branching patterns that occur in data collection, and (4) the ability to conduct probabilistic inference for prediction and diagnostics with an output metric that can be understood by managers and academics. Sample data from 1,101 recent transport service customers are utilized to select and validate a Bayesian network and conduct probabilistic inference.

## INTRODUCTION

The Service-Profit Chain (SPC) is a conceptual framework, which describes a process for delivering superior service value (Heskett, Sasser, & Schlesinger, 1997). According to the framework, positive business outcomes result when service delivery activities lead to customer satisfaction and customer loyalty. However, managerial implementation of the SPC has remained problematic, primarily because of certain limitations in the modeling methods commonly employed. Providing managers with the ability to model this process in a way that is of practical value to decision makers would be a considerable contribution. Our objectives are to (1) present a mechanism for implementing the SPC model that is superior to other widely used models, and (2) to demonstrate the decision-enhancing capabilities of the implementation. We use transportation service satisfaction as the specific context for our presentation.

The SPC provides the framework for developing a conceptual model of causal relationships specific to transport service satisfaction. An implementation mechanism should provide a smooth translation of the conceptual model into a structure of actionable variables that can utilize typical survey data. The mechanism should have the capability to accommodate branching patterns in the data portrayed as decision points in a conceptual model, analy/e nonlinear relationships, and provide prediction and diagnostics over a variety of potential conditions. The Bayesian network methodology has these capabilities and is chosen as the implementation mechanism.

The article organization is first to explore issues from the literature concerning implementation of SPC models. Next, the decision to select a Bayesian network as the implementation mechanism is discussed and the fundamentals of a discrete Bayesian network are presented. Then, a conceptual model of transport service satisfaction is offered, the survey data used in the implementation are described, and the Bayesian network is used to conduct probabilistic inference. The article concludes with a brief summary that discusses the generalization of the process described in the article.

## SPC IMPLEMENTATION ISSUES

There have been only a few empirical studies of firms attempting to relate specific drivers of satisfaction to relevant business outcomes using service quality or satisfaction programs that are conceptualized within the SPC framework. These studies, which have tended to focus on statistical verification of selected SPC links, have identified multiple stumbling blocks. In an early investigation, Bolton and Drew (1991) examine key components of the SPC using a multistage model. Analyzing survey data collected from telephone customers, they estimate relationships among perceived performance, disconfirmation, service quality, and value constructs. The strength of this modeling effort is a confirmation of specific service component ratings as a cause of overall quality and value. Emphasizing the potential benefits of applying the SPC concept, they recommend that:

... the specification and operationalization of the model must he carefully tailored to the specific service context. This effort will be rewarded by the many managerial implications that can be derived from estimation results (p. 384).

Later, Bolton and Drew (1994, p. 179) comment that "... most companies are (as yet) unable to link engineering/operations attributes to customer perceptions of service through statistical models."

Mittal, Ross, and Baidasare (1998) investigate attribute-level performance and its impact on satisfaction and repurchase intentions, recognizing that managerial decisions most directly affect specific attributes. They find an asymmetric relationship in which negative attribute-level performance has a larger impact on satisfaction than does a similar magnitude of positive attribute-level performance. These relationships are found to be nonlinear, with positive performance yielding diminishing returns on satisfaction. Likewise, Anderson and Mittal (2000) in a review of empirical studies of key SPC linkages, describe a variety of asymmetric and nonlinear functional forms. They conclude that customer satisfaction models need to be able to capture the asymmetrical and nonlinear nature of the performance-satisfaction, satisfaction-retention, and retention-profitability linkages. Further, they address practical questions about how to provide information in a way "... that frontline managers can readily understand and implement the insights" (p. 111). One successful solution reported is based on emphasizing ordinal data characteristics ("top-2-box" and "bottom-2-box" scores) to portray asymmetrical relationships.

Mittal and Kamakura (2001 ) find in a study of satisfaction-repurchase intent and satisfaction-repurchase behavior relationships that these relationships (1) may vary by customer segment, (2) appear to be nonlinear, and (3) are subject to significant response bias. Further, the functional form of the satisfaction-intention relationship differs significantly from the functional form of the satisfaction-behavior relationship. Since most commercial surveys measure the former, managers may make decisions about managing satisfaction that are likely to be suboptimal given the true satisfaction-behavior relationship.

Loveman (1998) focuses on individual SPC relationships in examining six key linkages, one pair at a time, using data from a large bank. He reports correlational support for each hypothesized link, but explicitly notes that issues of causality cannot be addressed in this manner. Further, Loveman (1998, p. 30) speculates that "Future research will employ more sophisticated multiequation methods to improve tests of the service-profit chain model and better discriminate among competing explanations for equivocal results."

More recently, Kamakura, Mittal, De Rosa, and Mazzon (2002) employ sophisticated multiequation methods to simultaneously assess SPC links at the strategic and operational levels. Their study uses a two-stage analysis approach by first developing a strategic model using structural equation modeling, then applying data envelopment analysis to perform an efficiency analysis at the operational level of the firm (bank branches). In this comprehensive SPC modeling effort, a structural equation model (SEM) features five construct equations to describe and test relationships between operational efforts, consumer perceptions, and business outcomes. Among the hypothesized causal relationships encoded by the construct equations are the simultaneous positive and negative impacts of service quality expenditures and investments on profitability, as proposed in the Return on Quality framework by Rust, Zahorik, and Keiningham (1995). The resulting model, based on rank order correlations, is presented as an acceptable representation of the strategic aspects of the SPC concept.

Kamakura et al. (2002) conclude that a strategic model created with structural equation-modeling methodology can provide a blueprint for the top management. This strategic blueprint is presented in the form of a path diagram with measured operational variables and latent strategic variables. The diagram represents all variables in standardized form and expresses the strength of relationships in standard deviation units. We speculate from a managerial perspective that there are obvious weaknesses in adopting this approach. Although the notion that a model of latent strategic variables causing observed operational variables has considerable theoretical appeal and provides an elegant representation of a process, there is no path to practical implementation. Specifically, the portrayal of operational variables caused by theoretical strategic constructs is not actionable, since the constructs cannot be directly observed or measured.

Collectively, the above studies suggest the issue is no longer whether hypothesized SPC linkages exist, but that the focus should be on selecting an implementation mechanism that meets the requirements for producing managerially relevant information based on a conceptual model. As noted above, these requirements include the capability to accommodate branching patterns in the data, and estimating nonlinear relationships from typical nominal and ordinal level measurements (Bolton & Drew, 1994; Mittal et al., 1998; Anderson & Mittal, 2000; Soteriou & Zenios, 1999). More specifically, a practical issue that has been ignored in past studies is that survey instruments commonly employ branching sequences in data collection. If consistency is to be maintained between the conceptual model and the implementation mechanism, branching should be portrayed in a single comprehensive model. Researchers investigating service recovery comment on the difficulty of conducting empirical work since recovery is always triggered by a service failure (Smith, Bolton, & Wagner, 1999). That is, one data collection sequence tracks the service failure experience while a separate branch traces the no service failure experience. The two branching paths result in missing respondent data on one sequence, and most methods cannot provide analysis for representation of service failure versus no service failure in a single model.

## ALTERNATIVE IMPLEMENTATION TECHNIQUES

Heskett et al. (1997) recommend the cause-effect methodologies of service mapping and Fishbone diagrams to summarize survey data. Service mapping provides a blueprinting of service processes in a flowchart format, and Fishbone diagrams provide a descriptive portrayal of assumed causal relationships that emphasize the more important causes. While both methods are very useful for a qualitative viewpoint, they lack desirable quantitative features necessary for prediction and diagnostics. To address these issues, most previous researchers have turned to regression analysis and to structural equation modeling. We propose that Bayesian network analysis is an attractive alternative quantitative modeling technique.

A comparison of regression models, SEMs, and Bayesian networks as potential implementation mechanisms is presented in Table 1. The comparison focuses on model building, statistical assumptions, goodness-of-fit measures, inferential capabilities, and interpretability. The type of variables a method can accommodate, branching capabilities, and an evaluation of overall complexity are the comparative elements in model building.

The statistical assumptions are functional and distribution forms. Goodnessof-fit components are the available significance tests and descriptive indices. Inferential capabilities include prediction and diagnostics. The elements of interpretability are the primary output metric and evaluations of theoretical and managerial explanation abilities.

Regression models offer a convenient simplification of a process by dichotomizing the variables into a single dependent variable and a group of predictors. Attributes of regression models include the flexibility to accommodate continuous and discrete variables, and relative low complexity. Further, regression models require only a minimal understanding of the process and few technical skills related to data setup and interpretation. After decades of classroom repetitions, together with accessible user-friendly computational programs and help screens (e.g., Excel), there is increased familiarity with the interpretation of R^sup 2^, the uses of dummy variables, and the interpretation of regression coefficients. These characteristics make regression methods very useful for short-term forecasting. The primary weakness of regression models is the inability to explain independent variable relationships. The simplistic dichotomization into dependent and independent variables provides no improved understanding of the process or useful insight for decision making.

Latent variable SEM is currently a very popular method in the decision sciences. The method requires specification of a measurement model where latent variables are the common causes of observed measures, and of construct equations that postulate cause-effect relationships between the latent variables. Thus, the model-building task is complex, and SEM cannot accommodate a branching sequence. The computational programs yield linear regression parameterizations for the assumed multivariate normally distributed variables (e.g., LISREL, the acronym for Linear Structural Relations). Although there are recommendations for how to accommodate nonlinear relationships (Schumacker & Marcoulides, 1998), the required elaborate coding and resulting complex output are generally unacceptable for empirical research and managerial consumption.

Significance tests and a wide assortment of goodness-of-fit indices are available for SEM assessment. In applications, researchers tend to place more emphasis on descriptive fit indices than significance testing, because of the questionable tenability of the normality assumption and the sensitivity of the tests to minor model misspecifications. Three goodness-of-fit indices commonly cited in the literature are given in Table 1. An index of the relative amount of the sample covariance explained by the implied covariance created by a fitted model is the GFI, an acronym for goodness-of-fit index. The CFI, an acronym for comparative fit index, yields a comparison between the fitted model and the model of no relationships (the independence model). Last, the CK, a single sample cross-validation index, is an approximation of expected fit if the fitted model is applied to new sample data.

The greatest strength of SEM is the ability to provide an excellent theoretical explanation by employing latent variables. However, since latent variables have an arbitrary metric, the standardized metric is used in computational programs for purposes of mathematical identification and convenience. This requires that an effect analysis, which is designed to measure relative influences of direct and indirect causes, be expressed in unit standard deviations. This is a major disadvantage for a managerial interpretation of output. A related disadvantage is how to apply the statistical results since prediction is limited to latent variables. This limitation is due to the fundamental causal structure of the measured variables, that is, latent variables are the causes of measured variables. Thus, to predict a measured variable we must assign a value to a latent common cause. However, as noted above, the metric of a latent variable is not defined and the relative metric of unit standard deviations is used. As with the discussion of the Kamakura et al. (2002) study, we assess the managerial explanatory abilities of SEM as poor.

A methodology that can simultaneously portray a cause-effect service mapping in terms of actionable variables is a Bayesian network (Pearl, 1988). In theory, a Bayesian network can be applied to both discrete and continuous variables. However, most practical applications are limited to discrete variables due to problems of efficient estimation with a mixture of discrete and continuous variables. Further, the majority of available computational programs deal only with discrete variables. However, the data assumptions of a discrete Bayesian network are an excellent match to the nominal and ordinal level measurements usually collected by consumer surveys. In addition, it can accommodate data collected in branching sequences in a single model and express the nonlinear aspects of variable relationships.

A Bayesian network can utilize significance tests of independence to confirm or reject a hypothesized structure. In addition to significant tests, an estimate of goodness-of-fit as the probability of a correctly specified model given the sample data, p(Model|Data), is available. Probabilistic inference capabilities allow predictions and diagnostics, which are reported in the probability metric.

Theoretical explanations conveyed by a Bayesian network are a reasonable approximation of underlying processes, but not as strong as SEMs. The managerial explanation is strong due to the use of actionable operational variables. A computational program can provide interactive query capabilities for diagnostics and predictions that permit great scenario evaluation flexibility for a nontechnical manager. In total, we consider a discrete Bayesian network the superior modeling technique for investigating both strategic and operational aspects of SPC models.

## DISCRETE BAYESIAN NETWORKS

A Bayesian network has a qualitative component in the form of a graph, called a DAG (directed acyclic graph), which portrays the assumed structure of cause-effect relationships. The quantitative component is a set of conditional probability distributions that provide the statistical interpretation of the cause-effect relationships represented by the graphical structure. The conditional probability distributions can be estimated subjectively by domain experts or can be based on sample data using relative frequency ratios.

A directed acyclic graph, a DAG, portrays the cause-effect structure with the nodes representing the variables and arrows connecting nodes showing the assumed relationships. The graph is characterized as directed because two-headed arrows depicting noncausal association (e.g., X [Lefright arrow] Y) are not allowed, and as acyclic because feedback loops (e.g., X [arrow right] Y [arrow right] X) are not permitted. The directed restriction in the graph is based on the common cause principle (Reichenbach, 1956) that states population associations always arise from an underlying causal structure. Therefore, the graphical structure underlying an observed association between X and Y, when X precedes Y, is X [arrow right] Y or X [arrow left] C [arrow right] Y, where C is a common cause. An independence relationship is portrayed by the absence of a connecting arrow, X Y.

Two approaches have been employed to evaluate potential Bayesian networks: an independence-testing method (Cooper, 1997; Spirtes, Glymour, & Scheines, 1993) and a Bayesian scoring method (Cooper & Herskovits, 1992; Heckerman, Geiger, & Chickering, 1995). The independence-testing approach is consistent with traditional statistical methods, utilizing likelihood ratio significance tests and maximum likelihood estimation. The Bayesian approach employs a prior-to-posterior probability analysis to assign a score to each possible Bayesian network. Both approaches allow incorporation of prior knowledge such as the assumed temporal order of the variables and possible relationship restrictions.

## Independence Testing

The independence-testing method adopts a traditional frequentist viewpoint, where it is assumed that there is a fixed but unknown population Bayesian network that is to be estimated from sample data. Causal assertions are evaluated by sample tests of independence employing the maximum likelihood chi-square test statistic under the assumption of multinomial sampling. The logic of the independencetesting approach is the validity of the claim that X causes Y implies probabilistic dependency, p(Y | X) ≠ p(Y). Therefore, rejection of the null hypothesis of marginal independence of X and Y, H^sub 0^: p(Y | X) = p(Y) or p(X, Y) = p(X)p(Y), provides evidence that p(Y | X) ≠ p(Y) in support of the structure X [arrow right] Y. Conversely, a decision in support of independence is taken as evidence for rejection of the causal claim. An assertion that the association between X and Y is due to a common cause C, X [arrow left] C [arrow right] Y, can be evaluated in a similar manner. That is, if the hypothesis of conditional independence, H^sub 0^: p(X, Y, C) = p(X | C)p(Y | C), is rejected, the common cause claim is supported.

Independence testing is a two-step procedure implemented by the PC algorithm of TETRAD II (Scheines, Spirtes, Glymour, & Meek, 1994). In the first step, independent and conditional independent relationships are tested at a specified significance level, guided by user-specified variable order and constraints. A rejected independence relationship implies a dependency linkage, and these linkages are assembled to provide the graphical structure. In the second step, maximum likelihood estimates of the conditional probability distributions conforming to the graphical structure are computed from sample data.

## Bayesian Scoring

The statistical perspective of the Bayesian scoring approach is that the population Bayesian network is a stochastic variable, denoted by BN. The statistical task is to combine prior knowledge of BN with sample data to estimate a posterior probability distribution for BN. The incorporation of prior knowledge of BN into an analysis requires a prior probability distribution. However, prior knowledge of BN is very often minimal or nonexistent, and a uniform prior probability distribution is assumed (Geiger & Heckerman, 1997).

A prior-to-posterior probability analysis is greatly simplified when the prior and the posterior distributions have the same form, which is termed as being conjugate (Bernardo & Smith, 2000). The family of Dirichlet distributions is conjugate for multinomial sampling, and a prior Dirichlet distribution when combined with sample relative frequency ratios will yield a posterior Dirichlet distribution (Cowell, Dawid, Lauritzen, & Spiegelhalter, 1999, Appendix A). A Dirichlet distribution is specified by values known as hyperparameters. The term hyperparameter is used in the literature rather than the term parameter because a Dirichlet distribution is often used to describe variation in a set of parameters.

Hyperparameters play the role of frequencies in specifying a Dirichlet distribution as a uniform prior distribution. A quantity called the prior precision establishes the values of the prior hyperparameters, and determines the magnitude of the uniform prior probabilities. The value of the prior precision must be specified by the researcher. The value of a prior precision can be conceptually viewed as the equivalent sample size needed to represent past experience (Winkler, 1967). Small values for the prior precision, such as an equivalent sample size of 1, indicate a low level of relevance for the prior distribution in reference to sample data in the estimation of a posterior distribution. If sample size is relatively large compared to a small prior precision, then the posterior probability distribution is primarily a function of sample information.

Bayes' theorem is used to revise the Dirichlet prior distribution of BN, denoted by p(BN), given sample data, denoted by Data, to a Dirichlet posterior distribution of BN, denoted by p(BN | Data). The quantity p(Data | BN) is known as likelihood of the data given BN. The computations necessary to perform the prior-to-posterior analysis is p(BN | Data) - p(BN, Data)/p(Data) = p(BN)p(Data | BN)/p(Data). In words, the posterior probability of BN, p(BN | Data), is equal to the product of prior probability of BN times likelihood of the data given BN, p(BN)p(Data | BN), divided by the probability of the data, p(Data). The p(Data) does not change over the possible BNs, and can be viewed as a scaling constant to assure that the posterior probabilities sum to unity. Therefore, the posterior distribution of BN, p(BN | Data), is proportional to p(BN)p(Data BN). This proportionality property allows the quantity p(BN)p(Data BN) to be viewed as the relative posterior probability, which is often expressed as a natural logarithm, Inp(BN) + Inp(Data | BN). However, under the assumption of a uniform prior probability distribution, In p(BN) is constant, allowing the quantity In p(Data | BN) to be used as the Bayesian network score.

A Bayesian network score reflects the probability of data being produced by a causal process as specified by a given BN, and is interpreted as the ability of that BN to predict the data. A Bayesian network score will always be negative since it is the logarithm of a probability, thus the smallest negative score among a group of BNs indicates the most probable BN.

Combining Independence-Testing and Bayesian Scoring Methods

Although the philosophical positions underlying the independence-testing and Bayesian scoring methods are very different, there has been an effort to combine the two approaches in developing Bayesian networks. Simulation studies show that the PC algorithm, with very large samples, will recover all causal relationships from observational data given valid statistical testing and certain philosophical assumptions common to all causal modeling methods (Cooper, 1999). However, with typical sample sizes, independence testing may indicate too few linkages due to type II errors, or too many linkages due to type I errors (Scheines et al., 1994). Spirtes and Meek (1995) adopted a hybrid approach by using independence testing to establish a baseline Bayesian network, and then compared network alternatives on Bayesian scores. Scheines (1999) also adopted a combination approach in using independence testing to establish a network structure, and then employed a prior-to-posterior analysis to estimate conditional probability distributions.

We will adopt a hybrid approach of using independence testing to confirm or reject the hypothesized causal relationships in the candidate Bayesian networks, and Bayesian scores for selecting the most appropriate network. The conditional probability distributions associated with the selected Bayesian network will be estimated by the prior-to-posterior analysis.

Probabilistic Inference

A Bayesian network can conduct probabilistic inference that involves an intervention analogous to physical manipulations in a classical experiment (Cowell et al., 1999; Pearl, 2000; Jensen, 2001). The process is started by an intervention on selected evidence variables, denoted by set[evidence], which fixes a value of each evidence variable to certainty, that is, p(set[evidence]) = 1. In the Bayesian network literature, an intervention is often referred to as an instantiation of the evidence variables. The probability distributions of a selected set of nonevidence variables, known as the query, are then revised. These postintervention probabilities, p(query | set[evidence]), provide the quantitative assessment of the impact of an intervention.

An intervention can create forward network inference from cause to effect by selecting a set of causes as the evidence, set[evidence = {causes}]), and selecting a set of effects as the query. Forward inference will yield predictions in the form of p(query = {effects} | set[evidence = {causes}]). Diagnostics can be obtained by backward network inference from effect to cause with an intervention on a set of effects as the evidence, set[evidence = {effects}]), and selecting a set of causes as the query to yield p(query = {causes} | set[evidence = {effects}]).

Computational Programs

User-friendly Bayesian scoring software that requires no programming, special training, or data manipulation is readily available (e.g., www.Bayesware.com offers a demonstration program called Discoverer). Bayesian scoring programs usually adopt the assumption of Dirichlet prior and posterior probability distributions for the prior-to-posterior analysis. Most programs assume that prior knowledge is lacking and a uniform Dirichlet prior distribution for BN is established by specifying a prior precision value. A prior precision equal to 1 is often the default value for prior precision, which gives a minimum weight to prior knowledge in reference to sample data. Software, such as Discoverer, provides on-screen displays of sample frequencies, conditional probability distributions, variable and network scores, wizards for batch prediction, cross-validation and network comparisons, and interactive tools for probabilistic inference (Ramoni & Sebastiani, 1999). Later, we demonstrate some of the capabilities of this type of software to estimate and evaluate alternative models, to predict and explain managerially relevant outcomes, and to perform "what-if" analyses. First, however, we present the underlying conceptual SPC model of transport service satisfaction, and describe the methods and variables used to develop the Bayesian Network.

## SPC MODEL OF TRANSPORT SERVICE SATISFACTION

The model of transport service satisfaction, shown in Figure 1, is based on a service inputs, customer satisfaction, and customer loyalty structure derived from the SPC concept (Heskett et al., 1997, Chapter 8).

Service inputs in Figure 1 are the attributes of timeliness, courtesy, and damage level, plus a claim-filing decision. Customer satisfaction components in the model are attribute satisfactions, performance evaluation, and customer expectation. Customer loyalty variables are retention and endorsement. Figure 1 is built in modular fashion from three overlapping processes that are described below.

### Attribute Level Satisfaction

In the commercial transportation industry, "reliability" (defined as consistency of arrival and delivery times) and "responsiveness to inquiry" (defined as the speed and accuracy of providing information in response to an inquiry) are the two most important factors in evaluating transportation partners (LaLonde & Cooper, 1989). We anticipate these same two factors, along with damage and damage-related service recovery, drive the evaluation of retail transportation services as well. Such attribute-level satisfaction ratings have previously been modeled as antecedent to global performance evaluation (Spreng, Mackenzie, & Olshavsky, 1996; Bolton & Drew, 1994) and to overall satisfaction (Anderson & Mittal, 2000; Mittal et al., 1998).

The damage-claim-satisfaction subprocess involves an obvious branching component. That is, if a customer has no damage, then no claim is filed. Further, a customer may incur damage, but decide not to file a claim. A consumer's ultimate satisfaction with damage level, whether or not the consumer utilized the claims (service recovery) process, is thought to be a driver of perceived performance.

### Satisfaction, Performance, and Expectation

Attribute level satisfactions are assumed to be direct causes of perceived performance, and performance the direct cause of disconfirmation of expectations (Bolton & Drew, 1994). Disconfirmation of expectations is conceptualized as the degree to which experienced performance falls below, meets, or exceeds the consumer's prior expectations regarding performance. Empirical evidence consistently demonstrates a significant, positive relationship between disconfirmation and satisfaction (e.g., see Bolton & Drew, 1991 ; Oliver, 1981 ; Tse & Wilton, 1988). While others have also conceptualized and found evidence supporting a link between expectations and disconfirmation (Anderson & Sullivan, 1993; Churchill & Surprenant, 1982; Spreng & Olshavsky, 1993), the use of expectations measured postexperience has been criticized (Droge & Halstead, 1991; Halstead, 1993; Yi, 1990). Since the data in this study were collected following service delivery, no separate expectations-disconfirmation link appears in Figure 1. Others modeling the

SPC chain, likewise, have focused on disconfirmation and performance, but not expectations, as key antecedents of relevant outcome variables (Bolton & Drew, 1991).

### Performance, Expectation, and Loyalty

Customer loyalty, represented in part by retention and endorsement (Loveman, 1998), is viewed as a function of repeat purchase behavior and word of mouth intentions (Rust et al., 1995; Olivia, Oliver, & MacMillan, 1992). Given practical data collection considerations, repeat purchase intentions and likelihood of recommending the service provider are used as indicators of these behaviors. Previous theoretical and empirical work specifies these behaviors as consequences of overall satisfaction (Bolton, 1998; Mittal & Kamakura, 2001; Anderson, 1994), which is conceptualized as resulting from disconfirmation. Given the consistent, positive, and generally strong correlation between disconfirmation and satisfaction reported in the literature, we propose that, as a practical matter, the two outcome variables can be modeled as consequences of disconfirmation directly.

## METHOD

### Data

A major retail transport firm commissioned a survey to collect data from customers on ordinal measures of service value, customer satisfaction, and customer loyalty variables. Transport firm managers, consultants, and university researchers collaborated on question and response scale wording. Managers insisted that scales have ordinal data characteristics that matched response scales used in previous inhouse research. The questionnaire was pretested with a small sample (n = 20), revised, and finalized. Questionnaires were mailed to 5,520 randomly selected customers who had completed a move in the previous month.

The mail-out mail-back methodology yielded a response rate of 26%. Given the potential for nonresponse bias, telephone interviews were conducted with a small sample (n = 72) of nonrespondents to the original survey. As no evidence of significant nonresponse bias was detected, analysis proceeded using the mailback questionnaire responses. A subset of 1,101 cases of completed responses is available for the modeling effort. The data are randomly divided into 1,000 respondents for the analysis sample and 101 respondents for the holdout sample to be used in external validation of the selected Bayesian network.

### Variable Measurement

The service input variables, displayed in Figure 1, include the arrival time to pick up goods, labeled Arrival, and the time of delivery of goods, labeled Delivery. The variable Arrival has four response categories (on-time, 2 hours late, 1⁄2 day late, 1 or more days late), and the variable Delivery has four categories (on-time, 2 hours late, 1⁄2 to 1 day late, 2 or more days late). The perceived level of employee courtesy, labeled Courtesy, has three categories (low, average, high). The customer satisfaction variables of Timeliness Satisfaction and Information Satisfaction have three categories (low, average, high).

The level of damage experienced, labeled Damage Level, has four response categories (none, minor, moderate, major). The indication as to whether a claim is filed is coded in the Claim Satisfaction variable, which has five categories (no damage, no claim, low, average, high). The first two categories of Claim Satisfaction are the events of no damage experience and the decision not to file a claim. The last three categories are the satisfaction ratings for the claim experience. The damage satisfaction variable, labeled Damage Satisfaction, has five categories (no damage, no claim, low, average, high).

The coding of the Damage Level and Claim Satisfaction variables permit modeling the decision points shown in Figure 1 in a manner consistent with the branching sequences employed in the questionnaire. That is, the combination of responses on Damage Level and Claim Satisfaction determine the response category of Damage Satisfaction. When no damage is reported, the coded response set is {Damage Level - none, Claim Satisfaction = no damage, Damage Satisfaction = no damage}. When damage is reported but no claim, the coded response set is {Damage Level = minor or moderate or major, Claim Satisfaction = no claim, Damage Satisfaction - low or average or high}. When damage is reported with a claim, the coded response set is {Damage Level - minor or moderate or major, Claim Satisfaction = low or average or high, Damage Satisfaction = low or average or high}.

A measure of perceived overall performance, labeled as Perform, has three categories (low, average, high). Disconfirmation of expected performance, labeled as Overall Disconfirmation, has three categories (below, equal, above). The customer loyalty variables of Figure 1 are repeat purchase intention, labeled Retention, with two categories (no, yes), and the likelihood of recommendation, labeled Endorsement, with three categories (no, maybe, likely).

Alternative Bayesian Networks

The conceptual model displayed in Figure 1 is translated into a Bayesian network. The causal implication of the SPC framework applied to this study is symbolized as service inputs [arrow right] customer satisfaction [arrow right] customer loyalty. This SPC structure expands to attributes [arrow right] satisfaction [arrow right] performance [arrow right] disconfirmation [arrow right] loyalty, which when specialized to the questionnaire variables becomes the structure defined in Figure 2. The structure in Figure 2 is the qualitative component of a Bayesian network labeled BN1.

The structure of BN1 assumes the attribute satisfactions are independent, and performance is only an indirect cause of the customer loyalty variables. These assumptions can be evaluated by comparing BN1 to the alternatives BN2 and BN3. Network BN2 assumes that the attribute satisfactions are not independent as portrayed in Figure 2. Specifically, information satisfaction is postulated a cause of timeliness and damage satisfaction. The structure of BN2 is created by adding the linkages Information Satisfaction [arrow right] Timeliness Satisfaction and Information Satisfaction [arrow right] Damage Satisfaction to the structure of BN1. Network BN3 allows evaluation of the BN1 assumption that performance and the loyalty variables are independent given disconfirmation of expectation. The BN3 structure is created by adding Performance [arrow right] Retention and Performance [arrow right] Endorsement to the BN1 structure. Among the networks to be evaluated, BN1 is the most parsimonious and BN3 is the most complex.

## RESULTS

### Model Evaluations

Independence testing is applied to BN3, subject to the independence constraints of the structure. The testing is conducted with the probability of a type I error set at α^sub nominal^ = -001 for each test of independence. This rather strict nominal level of significance is adopted since there are 18 linkages in BN3 and each linkage requires a test of marginal or conditional independence. Multiple testing will inflate the overall probability of a type I error, but setting the nominal significance level to .001 will result in approximately an overall type I error rate of α^sub overall^ = 1 (1 - .001)18 = .018. The testing results show rejection of independence for each of the assumed relationships, and thus supporting each linkage in BN3. Since the networks of BN1 and BN2 are subsets of BN3, the linkages in all networks are supported. The probability of type II errors, β^sub overall^, is not controlled in this application of Neyman-Pearson null hypothesis testing. Thus, the overall power of testing, rejecting the hypothesis of independence when it should be rejected, computed by Power^sub overall^ = 1 - β^sub overall^ is also unknown. However, it is well known that in a single null hypothesis test decreasing α^sub nominal^ will increase β^sub nominal^ and decrease Power^sub nominal^, given constant sample size and effect size. Since α^sub overall^ is relatively small, but sample size is fairly large, it is difficult to even speculate about the Power^sub overall^ of the multiple testing. This problem of approximating error rates in the traditional null hypothesis framework is the major motivation for applying the Bayesian scoring approach.

The Bayesian scores of the three networks, with a prior precision equal to 1, are -9173.2 for BN1, -9427.7 for BN2, and -9204.8 for BN3. Bayesian scores decompose into additive variable scores for each network structure, so the assumption of attribute satisfaction independency between attributes in BN1 can be evaluated relative to the hypothesized satisfaction dependency in BN3. A Bayes factor (Kass & Raftery, 1995) is used to compare the independence assumed in the attribute satisfaction substructure of BN1 with the dependency substructure of BN2. A Bayes factor is calculated by exp(Score^sup Structure i^ - Score^sup Structure j^) and interpreted as the ratio of the likelihood of the observed sample data given Structure i versus the likelihood of the observed sample data given Structure j, p(Data Structure i)/p(Data | Structure j). The score for Timeliness Satisfaction for BN1 is -706.0, whereas the score for Timeliness Satisfaction for BN3 is -815.4. The odds of Timeliness Satisfaction independency versus dependency of Information Satisfaction, p(Data | Information Satisfaction no arrow Timeliness Satisfaction)/p(Data | Information Satisfaction [arrow right] Timeliness Satisfaction), is exp[-707.0 - (-815.4)] = 3.2 × 10^sup 47^ to 1 in favor of the independence structure. Similarly, the score for Damage Satisfaction for BN1 is -688.2 and for BN2 is -833.3. The odds of Damage Satisfaction independency versus dependency of Information Satisfaction is exp[-688.2 - (-833.3)] = 1.0 × 10^sup 63^ to 1 in favor of the independence structure. Thus, the independency of the attribute satisfaction variables in BN1 is supported.

The substructure of Performance as a cause of loyalty variables, Retention and Endorsement, in BN1 versus BN3 can also be evaluated by Bayes factors. The scores for Retention and Endorsement are -185.3 and -417.4 in BN1, - 191.0 and -443.1 in BN3. The odds of Performance being only an indirect cause of Retention versus being both a direct and indirect cause of Retention is exp[-185.3 - (-191.0)] = 298.8 to 1 in favor of the indirect cause structure. The odds of Performance being only an indirect cause of Retention versus being both a direct and indirect cause of Endorsement is exp[-417.4 - (-443.1)] = 1.5 × 10^sup 11^ to 1 in favor of the indirect cause structure. Thus, the independence of each loyalty variable and Performance given Disconfirmation portrayed in BN1 is supported.

Although the above analysis indicates BN1 is the most probable network, Bayes factors can be computed to compare BN1 versus BN2 and BN1 versus BN3 for the complete structures. The Bayes factor p(Data | BN1 )/p(Data BN2) is exp[-9173.4 - (-9427.9)] = 3.4 × 10^sup 110^ to 1 in favor of BN1. The Bayes factor p(Data | BN1)/p(Data | BN3) is exp[-9173.4 - (-9204.8)] = 4.3 × 10^sup 13^ to 1 in favor of BN1. Thus, BN1 is not only more parsimonious than BN2 and BN3, it is far more probable. BN1 is selected as the Bayesian network for the transport service satisfaction model of Figure 1. The conditional probability distributions consistent with the structure of BN1, Figure 2, are estimated by a multinomialDirichlet conjugate analysis with a prior precision of 1. The conditional probability distributions for BN1 are not displayed due to their size.

### Network Validation

Internal and external validation studies are conducted on the adopted Bayesian network, BN1. A cross-validation procedure is used to assess the internal consistency of the adopted Bayesian network (Stone, 1977). The procedure selects 100 cases as the response set and uses the remaining 900 cases to predict the response. The procedure is repeated ten times, with each case being included in a response set over the course of the cross-validation. The internal validation results for the customer satisfaction and loyalty variables are displayed in Table 2. The coding for data branching required that Damage Satisfaction and Claim Satisfaction be combined for the validation.

The overall predictive accuracy for the seven customer satisfaction and loyalty variables is 79%. The accuracy ranges from 69% for Information Satisfaction to 95% for Retention.

Prediction of the response variables for the 101 cases in the holdout sample provides a measure of external validity. The overall accuracy for the holdout sample is 86%. The accuracy ranges from 80% for Damage Satisfaction to 93% for Performance. The results of the external validation are also summarized in Table 2.

### Probabilistic Inference

The query capabilities of a Bayesian network can be employed to address managerial inquires that have direct implications for action. We demonstrate the query capabilities of probabilistic inference to (1) predict the probabilities of the customer loyalty variables given levels of customer satisfaction, (2) compute diagnostic probabilities for service input variables given levels of overall performance and claim satisfaction, and (3) investigate how service recovery impacts on satisfaction with damage.

Prediction of Customer Loyalty

Table 3 displays the predictions of the probability distributions of the customer loyalty variables, Retention and Endorsement, as the query resulting from interventions on the satisfaction variables, Timeliness Satisfaction, Information Satisfaction, and Damage Satisfaction, as the evidence.

The first three columns show the interventions on Timeliness Satisfaction of set[Timeliness Satisfaction = high], set[Timeliness Satisfaction = average], and set[Timeliness Satisfaction = low]. The resulting postintervention probabilities for Retention and Endorsement are listed in the lower query section of Table 3. Since Timeliness, Information, and Damage Satisfactions are mutually independent indirect causes of the customer loyalty variables, their distributions change only with an intervention. Thus, the probability distributions for Information Satisfaction and Damage Satisfaction are not revised in the Timeliness interventions. Columns 4, 5, and 6 show the postintervention probabilities for the customer loyalty variables given instantiated levels of Information Satisfaction. Columns 7, 8, 9, and 10 provide the predicted customer loyalty probability distributions given interventions on Damage Satisfaction.

Each of the attribute satisfaction variables has a substantial influence on the customer loyalty variables. The high level of Information Satisfaction yields the largest probabilities of positive retention (.79) and likely endorsement (.66). Information Satisfaction also exhibits the greatest variation in the predicted customer loyalty responses over the range of interventions. In this sense, Information Satisfaction has a greater influence on customer loyalty than does Timeliness and Damage Satisfaction. The last column in Table 3, intervention 11, shows the predicted probabilities of the customer loyalty variables when each satisfaction variable is fixed at its most favorable level, set[Timeliness Satisfaction = high], set[Information Satisfaction = high] and set[Damage Satisfaction = none]. The resulting probabilities for positive retention and likely endorsement are .92 and .80. These predictions, at maximum satisfaction levels, far exceed the current sample estimates of .66 for positive retention and .53 for likely endorsement. Thus, improvement in the satisfaction variables would yield very positive gains in customer loyalty.

Service Input Diagnostics

The diagnostic inference process proceeds backward from customer satisfaction (effects) to service inputs (causes), and reports p(query = {service inputs} | set[evidence = {customer satisfaction}]). These probabilities are estimated for Arrival, Delivery, Courtesy, and Damage given interventions on Performance and Claim Satisfaction. Table 4 displays the postintervention probabilities for service inputs, p(query = {service inputs} | set[Performance], set[Claim Satisfaction]).

The first three rows of Table 4 display the postintervention probabilities for service inputs given instantiated evidence of no damage and the varying levels of performance. The Courtesy variable exhibits the greatest variation over this set of interventions, p(Courtesy = high set[Claim Satisfaction = no damage], set[Performance = high]) = .81 to p(Courtesy = low | set[Claim Satisfaction = no damage], set[Performance = high]) = .48, indicating that employee courtesy has a greater influence than arrival and delivery times. The postintervention probability magnitudes for Arrival = on time, Delivery = on time, and Courtesy = high are very similar at a given level of Performance over the no claim, high, average, and low levels of Claim Satisfaction. As with the no damage interventions, the Courtesy variable exhibits the greatest variation.

The probabilities for the minor damage level show a pattern of decreasing magnitudes between levels of Claim Satisfaction for given levels of Performance. That is, p(Damage - minor | set[Claim Satisfaction], set[Performance = high]) are .84, .68, .57, and .33 for the no claim, high, average, and low levels of Claim Satisfaction. Interventions that fix Performance to the average level, p(Damage = minor | set[Claim Satisfaction], set[Performance = average]) are .79, .59, .51, and .26 for the no claim, high, average, and low levels of Claim Satisfaction. When Performance is set to the low level, p(Damage = minor | set[Claim Satisfaction], set[Performance = low]) are .75, .54, .46, and .21. The patterns within each of the above sets of Damage probabilities sharply decrease monotonically in interventions from set[Claim Satisfaction = no claim] to set[Claim Satisfaction = low]. These estimates indicate that Damage is a very influential cause of performance and claim satisfaction.

Service Recovery: Damage-Claim Interactions

The impact of interventions on combined Damage (minor, moderate, or major) and Claim Satisfaction (no claim, high, average, or low) provide some insight into service recovery efforts. Figure 3 displays the postintervention probabilities for high damage satisfaction given varying levels of Damage and Claim Satisfaction, p(Damage Satisfaction = high | set[Damage], set[Claim Satisfaction]).

When damage is minor, the differences in damage satisfaction across levels of claim satisfaction range from p(Damage Satisfaction = high | set[Damage = minor, Claim Satisfaction - low]) = .24 to p(Damage Satisfaction = high | set[Damage = minor, Claim Satisfaction = high]) = .74. The probability of high damage satisfaction is greater given minor damage with a high level of claim satisfaction (.74) than with no claim (.54).

The postintervention probabilities of high damage satisfaction, when damage is moderate, across levels of claim satisfaction range from p(Damage Satisfaction = high | set[Damage = moderate, Claim Satisfaction - low]) = .01 to p(Damage Satisfaction = high | set[Damage = moderate, Claim Satisfaction = high]) = .28. The probability of high damage satisfaction, given moderate damage with a high claim satisfaction, is double that of no claim (.28 vs.. 14). When damage is major, the probability of high damage satisfaction is zero across all levels of claim satisfaction. Thus, damage satisfaction varies in a nonlinear manner across the interaction of damage level and claim satisfaction as expected.

SUMMARY

Although our modeling effort is specialized to transport service satisfaction, we contend a similar effort would produce an operational implementation of a conceptual model in any service sector. Our modeling approach is to expand the modules of an established conceptual causal chain to form an operational Bayesian network structure. We selected the SPC as the conceptual causal framework since it is widely known and has great face validity. Further, the SPC has some empirical validation by the Kamakura et al. (2002) study discussed earlier.

The general SPC prescription of service inputs [arrow right] customer satisfaction [arrow right] customer loyalty is expanded to a generic causal chain portraying attributes [arrow right] attribute satisfactions [arrow right] overall performance [arrow right] disconfirmation of expectations [arrow right] loyalty outcomes. The generic structure is specialized to the conceptual transport service satisfaction model given in Figure 1. The final step in developing an operational structure is to represent each of the components in Figure 1 with variables that have a track record of successful implementation in data collection. Figure 2 presents the resulting Bayesian network structure. We see no great barriers to generalizing this approach to the full SPC structure of employee relations [arrow right] service inputs [arrow right] customer satisfaction [arrow right] customer loyalty [arrow right] profit for any service sector.

We are strong proponents of the Bayesian network methodology as the implementation mechanism for causal modeling. The strengths and weaknesses of the method are reviewed above in some detail, and comparisons are presented with two more widely known alternatives. If valid data inputs are available from survey data collection or internal records, the procedures described in this article should provide a substantive contribution to service management.

## REFERENCES


Anderson, E. W. (1994). Cross-category variation in customer satisfaction and retention. Marketing Letters, 72(Winter), 19-30.

Anderson, E. W., & Mittal, V. (2000). Strengthening the satisfaction-profit chain. Journal of Service Research, 3(2), 107-120.

Anderson, E. W., & Sullivan, M. W. (1993). The antecedents and consequences of customer satisfaction for firms. Marketing Science, 12(2), 125-143.

Bernardo, J. M., & Smith, A. F. M. (2000). Bayesian theory. New York: Wiley.

Bolton, R. N. (1998). A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. Marketing Science, 17(1), 45-65.

Bolton, R. N., & Drew, J. H. (1991). A multistage model of customers' assessments of service quality and value. Journal of Consumer Research, 17(March), 375-384.

Bolton, R. N., & Drew, J. H. ( 1994). Linking customer satisfaction to service operations and outcomes. In R. T. Rust & R. L. Oliver (Eds.), Service quality: New directions in theory and practice. Thousand Oaks, CA: Sage Publications, 173-200.

Churchill, G. A., & Surprenant, C. (1982). An investigation into the determinants of consumer satisfaction. Journal of Marketing Research, /9(4), 491-504.

Cooper, G. F. (1997). A simple constraint-based algorithm for efficiently mining observational databases for causal relationships. Journal, of Data Mining and Knowledge Discovery, 1, 203-224.

Cooper, G. F. (1999). An overview of the representation and discovery of causal relationships using Bayesian networks. In C. Glymour & G. Cooper (Eds.), Computation, causation & discovery. Menlo Park, CA: MIT Press, 3-62.

Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic network from data. Machine Learning, 9, 309-347.

Cowell, R., Dawid, A. P., Laurit/,en, S. L., & Spiegelhalter, D. J. (1999). Probabilistic networks and expert systems. New York: Springer-Verlag.

Droge, C., & Halstead, D. (1991). Postpurchase hierarchies of effects: The antecedents and consequences of satisfaction for complainers versus noncomplainers. international Journal of Research in Marketing, 8(4), 315-328.

Geiger, D., & Heckerman, D. (1997). A characterization of Dirichlet distributions through local and global independence. Annals of Statistics, 25, 1344-1368.

Halstead, D. (1993). The expectation-satisfaction relationship revisited: An empirical test and directions for future research. Proceedings of the 1992 AMA Winter Educators' Conference. Chicago: American Marketing Association, 371-372.

Heckerman, D., Geiger, D., & Chickering, D. M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. Machine Learning, 20, 197-243.

Heskett, J. L., Sasser, W. E., & Schlesinger, L. A. (1997). The service-profit chain. New York: The Free Press.

Jensen, F. (2001). An introduction to Bayesian networks. London: UCL Press.

Kamakura, W., Mittal, V., de Rosa, F., & Mazzon, J. (2002). Assessing the service profit chain. Marketing Science, 27(3), 294-317.

Kass, R., & Raftery, A. E. (1995). Bayes factors and model uncertainty. Journal of the American Statistical Association, 90, 773-795.

LaLonde, B. J., & Cooper, M. C. (1989). Partnerships in providing customer service: A third-party perspective. Oak Brook, IL: Council of Logistics Management.

Loveman, G. W. (1998). Employee satisfaction, customer loyalty, and financial performance: An empirical examination of the service profit chain in retail banking. Journal of Service Research, 7(1), 18-31.

Mittal, V., Ross, W. T., Jr., & Baldasare, P. M. (1998). The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions. Journal of Marketing, 62, 33-47.

Mittal, V., & Kamakura, W. (2001). Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics. Journal of Marketing Research, J8(February), 131-142.

Olivia, T. A., Oliver, R. L., & MacMillan, I. C. (1992). A catastrophe model for developing service satisfaction strategies. Journal of Marketing, 56(3), 8395.

Oliver, R. L. (1981). Measurement and evaluation of satisfaction processes in retail settings. Journal of Retailing, 57(3), 25-48.

Pearl, J. (1988). Probabilistic reasoning in intelligent systems. San Mateo, CA: Morgan Kaufmann.

Pearl, J. (2000). Causality: Models, reasoning and inference. New York: Cambridge University Press.

Ramoni, M., & Sebastian!, P. (1999). Bayesian methods for intelligent data analysis. In M. Bethold & D. J. Hand (Eds.), An introduction to intelligent data analysis. New York: Springer-Verlag, 129-166.

Reichenbach, H. (1956). The direction of time. Berkeley, CA: University of California Press.

Rust, R. T., Zahorik, A. J., & Keiningham, T. L. (1995). Return on quality (ROQ): Making service quality financially accountable. Journal of Marketing, 59(April), 58-70.

Scheines, R. (1999). Estimating causal influences: TETRAD II model selection and Bayesian parameter estimation. In C. Glymour & G. Cooper (Eds.), Computation, causation and discovery. Menlo Park, CA: MIT Press, 425437.

Scheines, R., Spirtes, P., Glymour, C., & Meek, C. (1994). TETRAD II: Tools for causal modeling. Mahwah, NJ: Lawrence Erlbaum.

Schumacker, R. E., & Marcoulides, G. A. (1998). Interaction and nonlinear effects in structural equation modeling. Mahwah, NJ: Lawrence Erlbaum.

Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A model of customer satisfaction with service encounters involving failure and recovery. Journal of Marketing Research, 26, 356-372.

Soteriou, A. C., & Zenios, S. A. (1999). Operations, quality, and profitability in the provision of banking services. Management Science, 45(9), 1221-1238.

Spirtes, P., Glymour, C., & Scheines, R. (1993). Causation, prediction, and search. New York: Springer-Verlag.

Spirtes, P., & Meek, C. (1995). Learning Bayesian networks with discrete variables from data. In U. M. Fayyad & R. Uthurusamy (Eds.), Proceedings of the first international conference on knowledge discovery and data mining. San Francisco: Morgan Kaufmann, 294-299.

Spreng, R. A., Mackenzie, S. B., & Olshavsky, R. W. (1996). A reexamination of the determinants of consumer satisfaction. Journal of Marketing, 60, 15-32.

Spreng, R. A., & Olshavsky, R. W. ( 1993). A desires congruency model of consumer satisfaction. Journal of the Academy of Marketing Science, 27(3), 169-178.

Stone, M. (1977). Asymptotics for and against cross-validation. Biometrika, 64, 29-35.

Tse, D. K., & Wilton, P. C. (1988). Models of consumer satisfaction formation: An extension. Journal of Marketing Research, 25(2), 204-212.

Winkler, R. (1967). The assessment of prior distributions in Bayesian analysis. American Statistical Association Journal, 62, 776-800.

Yi, Y. (1990). A critical review of consumer satisfaction. In V. Zeithaml (Ed.), Review of marketing 1990. Chicago: American Marketing Association.

## Author Affiliation

Ronald D. Anderson

Kelley School of Business, Indiana University, Indianapolis, IN 46202-5151, e-mail: randers@iupui.edu

Robert D. Mackoy

College of Business Administration, Butler University, Indianapolis, IN 46208-3443, e-mail: rmackoy@butler.edu

Vincent B. Thompson

Indiana Business Research Center, Indiana University, Bloomington, IN 47405-1703, e-mail: vbthomps@indiana.edu

Gilbert Harrell

Eli Broad Graduate School of Management, Michigan State University, East Lansing, MI 48824-1122, e-mail: harrell@msu.edu

Ronald D. Anderson is the American United Life Professor of Business Administration at the Kelley School of Business. Prior to that appointment, he held the Krannert chair in business administration at Butler University. His publications have appeared in Journal of Consumer Research, European Journal of Operations Research, Journal of Marketing Research, Journal of Retailing, Journal of Service Research, Journal of the Academy of Marketing Science, Multivariate Behavioral Research, Organizational Research Methods, Structural Equation Modeling, Transportation Journal, Journal of Business Logistics, International Journal of Physical Distribution, Logistics and Transportation Review and others. His research interests are in the application of Bayesian networks and structural equation models.

Robert D. Mackoy is associate professor of marketing at Butler University. His research has appeared in the Journal of Retailing, Journal of Services Marketing, Journal of the Academy of Marketing Science, Journal of Macromarketing, and Journal of Social Psychology, as well as in numerous professional conferences. His research interests focus on consumer satisfaction and consumer thought processes, trust formation in a computer-mediated environment, and ecotourism.

Vincent B. Thompson is economic research analyst for the Indiana Business Research Center at Indiana University's Kelley School of Business. His research interests are in causal modeling, simulation modeling, and applications of artificial intelligence in business. He is a member of the American Society of Transportation & Logistics.

Gilbert D. Harrell is professor of marketing in the Eli Broad Graduate School of Management, Michigan State University. His research publications have appeared in the Journal of Marketing, Journal of Marketing Research, Journal of Consumer Research, Journal of Retailing, Journal of the Academy of Marketing Science, Journal of Consumer Affairs, Journal of Industrial Marketing Management, and numerous other journals. His current research and consulting focus is on relationship marketing, strategic planning, and sustainable competitive advantage.