ENHANCING DISCRETE CHOICE DEMAND MODELING FOR DECISION-BASED DESIGN

In Press of ASME Journal of Mechanical Design

Henk Jan Wassenaar
Integrated Design Automation Laboratory
Department of Mechanical Engineering
Northwestern University
Evanston, Illinois 60208-3111
wassenaer@yahoo.com

Wei Chen*
Integrated Design Automation Laboratory
Department of Mechanical Engineering
Northwestern University
Evanston, Illinois 60208-3111
weichen@northwestern.edu

Jie Cheng
Power Information Network
J.D. Power & Associates
Troy, Michigan 48098

Agus Sudjianto
Engine Engineering Analytical Powertrain
Ford Motor Company
Dearborn, Michigan 48121

* Corresponding Author, 2145 Sheridan Road, Tech B224, Evanston, IL 60208-3111;
Phone: (847) 491-7019, Fax: (847) 491-3915.

Abstract

Our research is motivated by the need for developing an approach to demand modeling that is critical for assessing the profit a product can bring under the Decision-Based Design framework. Even though demand modeling techniques exist in market research, little work exists on demand modeling that addresses the specific needs of engineering design, in particular that facilitates engineering decision-making. In this work we enhance the use of discrete choice analysis (DCA) to demand modeling in the context of Decision-Based Design. The consideration of a hierarchy of product attributes is introduced to map customer desires to engineering design attributes related to engineering analyses. To improve the predictive capability of demand models, the Kano method is employed to provide the econometric justification when selecting the shape of the customer utility function. A (passenger) vehicle engine case study, developed in collaboration with the market research firm J.D. Power & Associates and the Ford Motor Company, is used to demonstrate the proposed approaches.

Key words: decision-based design, demand modeling, discrete choice analysis, vehicle design, engine design, Kano method, hierarchy of attributes, customer utility

1 INTRODUCTION

Decision-Based Design (DBD) is emerging as a rigorous approach to engineering design that recognizes the substantial role that decisions play in design and in other engineering activities, which are largely characterized by uncertainty and risk [1-6]. It should be noted that there exist differing views about the implementation of DBD. Not all existing approaches are considered correct and thus should be taken with caution. The DBD optimization seeks to maximize the *expected utility* (value[1]) of a designed artifact while considering the interests of both the producer and the end-users [1,2]. Although there is general consensus that for a profit-driven company, the value of a product should be a measure of the profit[2] it brings, there exist concerns on using profit as the single criterion in DBD because of the belief that profit seems too difficult to model. Related to the notion of enterprise-driven product design, one difficulty is the construction of a reliable product demand model that is critical for assessing the revenue, the total product cost, and eventually the profit. Even though demand modeling techniques exist in market research, little work addresses the specific needs of engineering design, in particular that facilitates engineering decision-making.

In market research, two major demand analysis techniques, Discrete Choice Analysis (DCA) and Conjoint Analysis are being used to capture customer choice behavior. DCA is based on probabilistic choice models, which were originated in mathematical psychology [7-10] and developed in parallel by economists and cognitive psychologists. DCA identifies patterns in choices customers make between competing products and generates the probability that an option is chosen. Conjoint Analysis (CA) [11-13] is an approach that determines the preference structure (a multiattribute utility function) of (classes of) consumers, which can be used to predict the demand as function of the product features (attributes). Though conjoint analysis remains a popular approach, DCA is seen by some as an evolutionary improvement (Riedesel, http://www.action-research.com/compare.htm; [14]). Among the advantages of DCA are: more freedom when formulating the survey questions; fewer problems with the degrees of freedom; a

[1] We use the notion "expected utility" in this paper as it stands for the selection criterion in the presence of uncertainty, while the "value" is often interpreted as a selection criterion without uncertainty.

[2] Profit is a result of accounting practices such as depreciation, which need not be related to engineering design. Therefore, with profit is meant net revenue, i.e., the difference between revenue and expenditure. The net revenue can be discounted to present value.

more natural task for the survey respondent; and the ability to handle more customer-oriented design attributes.

Demand analysis in market research is limited, focused on feature upgrades, packaging, product placement, etc., and regularly not intended to guide engineering design decisions. Applications of DCA for demand estimation has been seen in transportation engineering [15], however, very little research on demand analysis exists in the field of engineering design. One is the S-Model Approach proposed by Cook [17, 18], of which the key concept is the Taylor expansion about a reference point where the value and price of all products are identical, resulting in equal demand for each product. Li and Azarm proposed the Comparing Multi-attribute Utility Values Approach [16] and later a Customer Expected Utility Approach [19]. The former approach estimates the demand by comparing multi-attribute utility values obtained through conjoint analysis, while the latter approach assumes that the customer purchases a product when the product attributes satisfy certain limits (thresholds). It should be noted that some of the approaches above have limitations [20] when the assumptions made are not reasonable. In addition, it may not be possible to predict a customer's choice with certainty, therefore a more sophisticated approach is required such as the probabilistic choice theory [8]. In our previous work [2], we enhanced the DBD framework originally proposed by Hazelrigg [1] and proposed a discrete choice analysis (DCA) based approach to demand modeling. We found that the disaggregate demand models built from DCA use data of individuals instead of group averages, which enables a more accurate capturing of the variation of characteristics of individuals and avoids paradox associated with group decision-making.

Even though we have demonstrated the usefulness of DCA for demand analysis in engineering design [2], there are still many important research issues to be addressed before the method can be fully applied and used with confidence. First, more guidelines need to be provided to assist designers when applying the DCA approach to demand modeling that facilitates engineering decision-making, especially in the design of complex engineering systems. The mapping of customer desires to engineering design attributes related to engineering analyses needs to be addressed and the procedure should enable designers to focus the demand survey on specific features of the product without harming the consistency of the demand analysis at the system level of a product. Secondly, a method for selecting the form of the customer utility function should be provided so that the construction of a demand model is based on sound econometric reasoning instead of simple (mathematical) model fitting. A mathematical model may fit the data well however may not capture the true underlying purchase behavior and therefore its value in predicting demand may be poor.

In this work, we provide detailed guidelines for demand modeling that facilitates engineering decision-making. To bridge the gap between business and engineering, a hierarchy of product attributes is considered to map product feature related customer desires down to engineering design attributes that can be represented using engineering languages. We also present an approach for selecting the form of a customer (choice) utility function of demand analysis to enhance the predictive accuracy. The proposed approaches are demonstrated using a real (passenger) vehicle engine case study in collaboration with the market research firm J.D. Power & Associates and the Ford Motor Company. Before presenting our proposed methods, the necessary background of DBD, DCA, and the Kano method are first introduced.

2 TECHNICAL BACKGROUND

2.1 The Decision-Based Design Framework

The flowchart of the DBD framework that we proposed [2] as an enhancement to the Hazelrigg's DBD framework is shown in Fig. 1. The arrows in the flowchart indicate the existence of relationships between the different entities (parameters) in DBD, instead of showing the sequence of implementing DBD. One of our major contributions lies in introducing Discrete Choice Analysis as a systematic approach to establish the relationship between the customer-oriented design attributes $\boldsymbol{A}$, the socioeconomic and demographic background $\boldsymbol{S}$ of the market population, time t, and the demand Q. In this flow chart, the engineering design attributes $\boldsymbol{E}$ are any *quantifiable* product properties that are used in the product design process by a design engineer, represented as functions of design options $\boldsymbol{X}$ through engineering analysis. The customer-oriented design attributes $\boldsymbol{A}$ are product features and financial attributes (such as service and warranty) that a customer typically considers when purchasing the product. From the demand modeling perspective, the input ***A*** of a demand model could be either quantifiable or non-quantifiable (e.g., level of comfort). However, to assist engineering decision-making, the product feature related customer-oriented design attributes ***A*** need to be converted to quantifiable engineering design attributes ***E*** represented using engineering languages (see more details in Section 3.1). The engineering design attributes ***E*** may also include non customer-oriented design attributes that are only of interest to design engineers, e.g., stress level of a structure. The flow chart in Figure 1 coincides with an optimization loop that identifies the best design option ***X*** to maximize the expected utility. For enterprise-wide product design, Design Options ***X*** may include non-engineering related design options, e.g., warranty options, which have no relationship with the engineering design attribute ***E***, but will have a direct impact on some customer-oriented design attributes ***A*** (in this case the warranty feature).

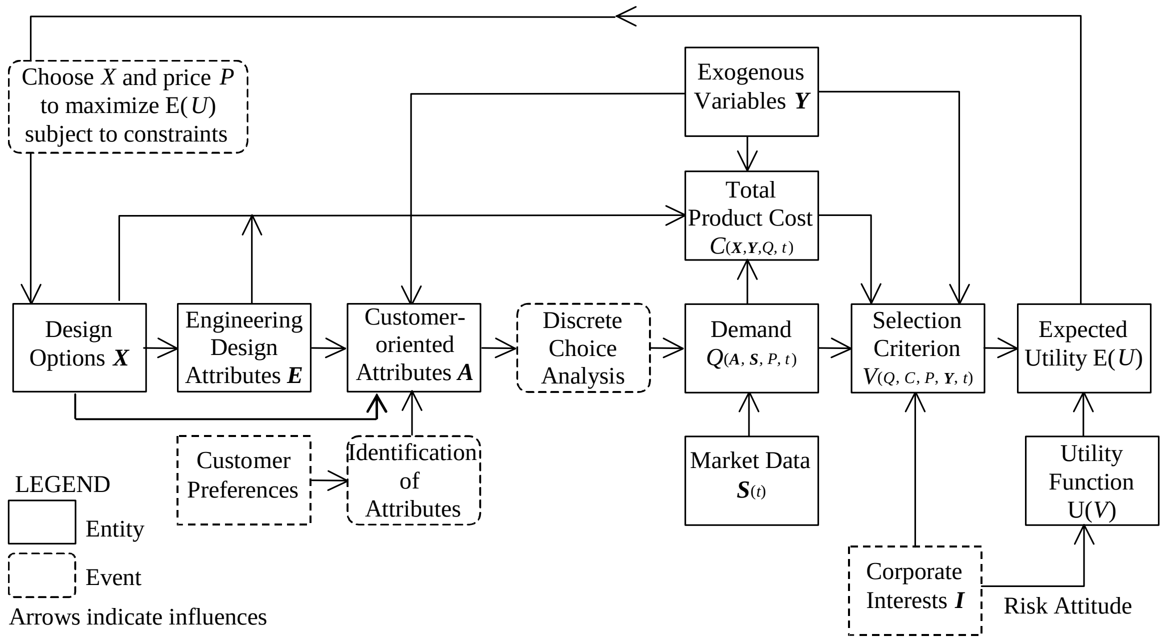

Figure 1. Decision-Based Design flowchart [2]

2.2 Background of Discrete Choice Analysis

DCA identifies patterns in choices customers make between competing products and generates the probability that an option is chosen. A key concept of DCA is to estimate the customer's (choice) utility at the individual level, employing the probabilistic choice theory to address unobserved taste variations, unobserved attributes, and model deficiencies. A quantitative process based on multinomial analysis is used to generate the demand model. The probabilistic choice theory entails the assumption that the *individual*'s true utility U can be *estimated* using a deterministic utility W supplemented with a random disturbance ε (see Eq. 1). The deterministic part can be parameterized as a function of observable independent variables (customer-oriented design attributes ***A***, socioeconomic and demographic attributes ***S***, and price P) and unknown coefficients β, which can be estimated by observing the choices respondents make (real or stated) and thus represent the respondent's taste, see Eq. 2. The β-coefficients and utility functions are indicated with the subscript n, representing the n^{th} respondent, the index i refers to the i-th choice alternative.

$$U_{in} = W_{in} + \varepsilon_{in} \tag{1}$$

$$W_{in} = f(A_i, P_i, S_n; \beta_n) \tag{2}$$

The probability that alternative 1 is chosen from a choice set containing two alternatives (binary choice) is then defined as the probability that the utility of alternative 1 exceeds the utility of alternative 2 or alternatively, on the probability that the difference between the disturbances does not exceed the difference of the deterministic parts of the utility, i.e.

$$\begin{aligned} \Pr(1)\big|[1,2] &= \Pr(W_{1n} + \varepsilon_{1n} \geq W_{2n} + \varepsilon_{2n}) \\ &= \Pr(\varepsilon_{2n} - \varepsilon_{1n} \leq W_{1n} - W_{2n}) \end{aligned} \tag{3}$$

Depending on the assumed distribution of the random disturbance ε different choice models can be formed. Assuming a normal distribution of the disturbance ε leads to the probit choice model [21, 22] and an extreme value (Gumbel) distribution results in the logit choice model [2, 15, 22] that predicts the choice probabilities. The choice probability of the multinomial logit model is shown in Eq. 4, where Prn(i) denotes the *estimated* probability that respondent n chooses alternative i.

$$\Pr_n(i) = \frac{e^{W_{in}}}{\sum_{l=1}^{J} e^{W_{ln}}} \tag{4}$$

Estimation techniques such as the maximum likelihood method can be used to determine the β-coefficients such that the predictions of the model matches the observed choices as closely as possible. The total demand for a particular design *i* is the summation of the predicted choice probabilities across the choice alternatives for the entire market population [15].

As stated in the introduction, DBD is developed to perform in design environments characterized by uncertainty and risk. Multiple sources of uncertainty can be accounted for in DCA. Probabilistic choice theory [8] addresses the uncertainty in predicting a customer's choice through the assumption of the random disturbance (ε) in the model of the customer's utility. Confidence interval analysis can be used to quantify the uncertainty associated with demand estimation. DCA models can also consider uncertainty in design options ***X*** through simulation in which the distribution of the uncertain design option is assessed. Uncertainty in exogenous variables ***Y*** and market data ***S*** can be considered in a similar way. The distribution of demand estimates that results from all uncertainties is included in the determination of the optimal design in the DBD framework through the (normative) utility function of the decision-maker *U*(*V*), see Fig. 1.

2.3 Kano Method

Traditional market analysis often assumes that customer satisfaction is proportional to product performance, i.e., linear. The Kano method [23], introduced in the late 1970s by Dr. Noriaki Kano of Tokyo Rika University, provides an approach to determine the generalized shape of the relation between product performance and customer satisfaction
y to 9.stime 1970s15.161 Tf0.55 0 TD0.0007 Tc0.1cannaand cusorm
by classifying the customer-oriented attributes into three distinctive categebe conscustomtegebo8c-0.ming-

reason that the customer does not expect them. For instance, the oil-change interval, a (unexpected) long interval may be expected to significantly increase satisfaction. Attributes are thought to move over time from excitive to basic to must-be. For example, cupholders were once excitive when first introduced but are now expected and their absence leads to great dissatisfaction.

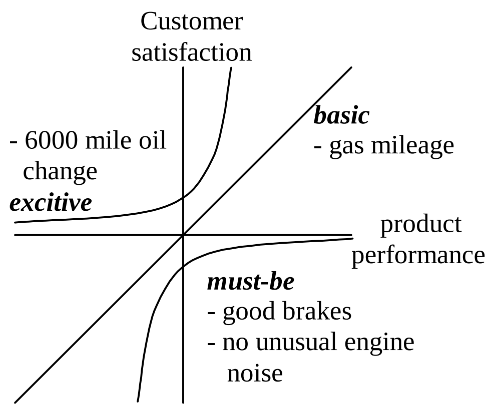

Figure 2. Example of Kano Diagram

The Kano method is frequently used in the literature to rank order the attributes according to their importance to determine where to focus the engineering effort to maximize customer satisfaction. In our view, it is wrong to attempt to rank order attributes based on the responses from a group of consumers due to the paradox associated with aggregating the preference of a group of individuals [24]. We believe that Kano's Method only allows a qualitative assessment of product attributes, i.e., the shape of the curves. In this work, we propose to enhance the predictive capability of DCA by using the Kano method to assess the generalized shape of the customer utility functions, W_{in} in Eq. 2. Doing so provides guidance to the demand modeling specialist in capturing the true customer behavior, improving explanatory power and predictive accuracy of demand models. In the next section we detail how an approach for demand modeling that facilitates engineering decision-making can be implemented.

3 IMPLEMENTING DISCRETE CHOICE ANALYSIS FOR DEMAND MODELING

From the demand modeling perspective, the input $\boldsymbol{A}$ (customer-oriented design attributes) of a demand model could be either quantifiable or non-quantifiable (e.g., level of comfort). However, to facilitate engineering decision-making, a demand model is expected to relate the market demand to engineering measures of product attributes that can be used to guide product design decision-making. The mathematical construct of the DCA approach to demand modeling was published in our earlier work [20]. In this paper, we focus on the procedure for implementing DCA for product demand modeling and discuss the potential issues involved in each phase of demand modeling. More detailed descriptions can be found in [19]. Our discussion follows the sequence of the four major phases for implementing DCA:

Phase I Identify customer-oriented design attributes $\boldsymbol{A}$, the range of price P and survey choice set; (**Attributes and Choice Set Identification**)

Phase II Collect choice data of proposed designs versus alternative choice options and record customers' socioeconomic and demographic background $\boldsymbol{S}$; (**Data Collection**)

Phase III Create a model for demand estimation based on the probability of choice. (**Modeling**)

Phase IV Use the demand model for market share and demand estimation (Demand Estimation)

3.1 Phase I - Attributes and Choice Set Identification

A useful demand model requires that the selection of customer-oriented design attributes $\boldsymbol{A}$ (explanatory variables) is based on econometric reasoning, that is, there should exist a causal relationship between the attributes and the customer's purchase decision. There are several methods available to assess what customers desire, what product attributes customers consider [25], and what competing alternatives should be considered in a discrete choice survey. Focus groups [26] can be used for both existing products and products that are new (e.g., innovative design).

Through surveys, the identified customer desires can be clustered together into groups that share similar characteristics such as, cost, performance, safety, operability, comfort, style, convenience, etc. These groups can be considered as "top-level *customer desires*". The next step is to identify for each customer desire what are the corresponding customer-oriented *design attributes* $\boldsymbol{A}$ that can be used to describe the product either in a product survey (stated preference) or in a product sales database (revealed preference).

To facilitate engineering decision-making, the customer language of the customer desires (e.g., good engine sound quality) should be translated into quantifiable engineering language, i.e., to identify suitable units of measurement for each customer desire. This task consists of cooperation between market researchers and engineering specialists, and perhaps consultations with customers to verify the correct understanding of the customer desires. It implies that a design engineer must develop a preliminary understanding of the design and how the design can fulfill the customer desires. Identification of the customer-oriented design attributes for some customer desires is straight forward, e.g., miles per gallon for fuel economy in vehicle design. For other customer desires this can be quite complicated, e.g., vehicle style or engine sound quality. It is possible that multiple attributes need to be used to capture the properties of a customer desire, while one attribute can impact multiple customer desires. Figure 3 demonstrates how the top-level customer desires are mapped to specific customer desires (in customer language), and then to customer-oriented design attributes $\boldsymbol{A}$, which are further related to engineering design attributes $\boldsymbol{E}$. The quantifiable customer-oriented design attributes $\boldsymbol{A}$ become a subset of the engineering design attributes $\boldsymbol{E}$; the later also also include attributes that are of interest only to designers but not customers. Establishing such a mapping relationship is especially important in the design of complex engineering systems.

Demand modeling deals with the functional relationship between the market demand (Q) and the customer-oriented design attributes $\boldsymbol{A}$, along with the socioeconomic background of customers $\boldsymbol{S}$, price P, and the time factor t. To integrate the demand model into the Decision-Based Design framework (see Figure 1), engineering analysis (modeling) needs to be carried further to establish the relationship between design options $\boldsymbol{X}$ and the customer-oriented design attributes $\boldsymbol{A}$. During the engineering analysis process, additional design attributes might be introduced as intermediate variables or variables that impose physical restrictions on the design or impact the product cost. As an example of mapping customer desires to specific design, we show at the right side of Figure 3 that (low) engine sound while idling could be considered as a specific customer desire that belongs to the group of (engine) performance, which is a product benefit. Radiated sound and engine mount vibration can be considered as customer-oriented design attributes for measuring the engine sound while idling. At the next level, engineering models will consider other engineering design attributes such as main bearing clearance, and crankshaft stiffness. Quantifiable customer-oriented design attributes ***A*** like radiated sound, engine mount vibration, and additional design attributes like bearing clearance and crankshaft stiffness will all be considered as engineering design attributes ***E***, which are modeled as functions of design options such as crankshaft material, pin diameter, and cheek thickness, through engineering analysis.

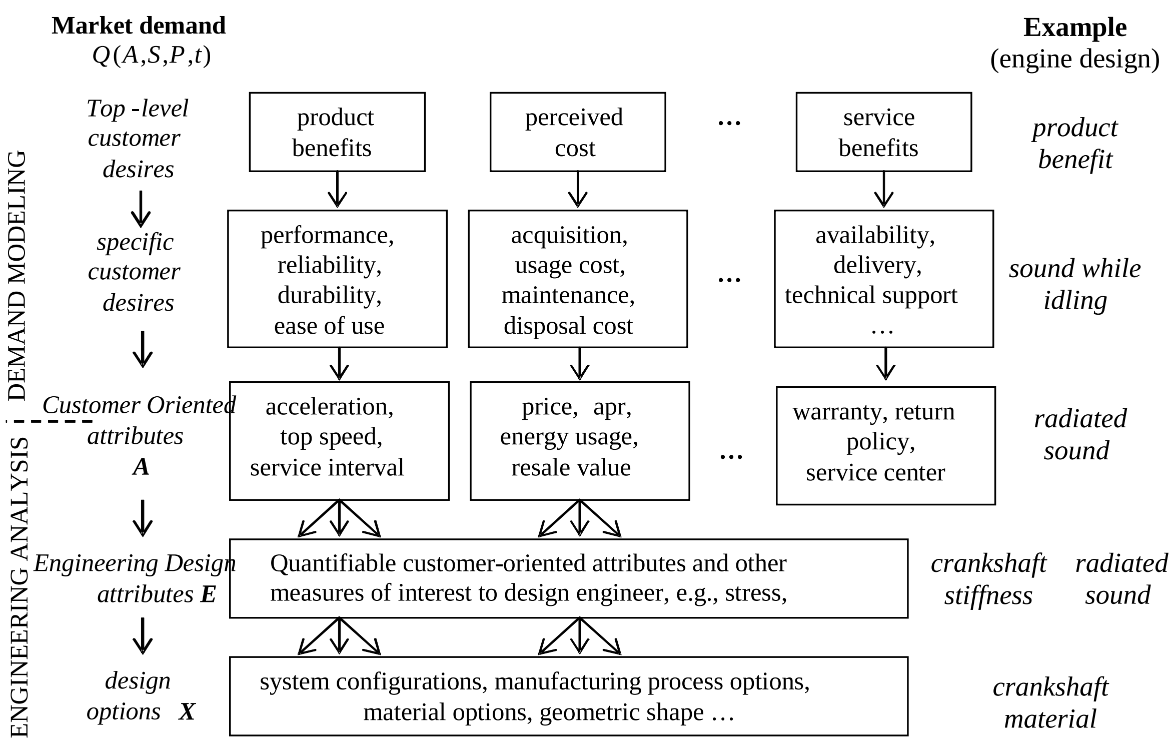

Figure 3. Mapping top-level customer desires to design options

3.2 Phase II – Data Collection

Stated Choice Vs. Revealed Choice

There are two ways of collecting choice data: stated choice and revealed choice. Revealed choice concerns actual (purchase) behavior that can be observed in real choice situations. Stated choice concerns controlled choice experiments that ask the respondents to state their intent without needing to commit to the consequence of their decision (e.g., pay the purchase price). Revealed choice can be used when similar products or services exist (e.g., when redesigning a powertool), while stated choice can be used for innovative designs, product features, or services that do not yet exist, which have new attributes or whose levels of the customer-oriented attributes lie outside the attribute level range of existing products. One should be aware that either stated choice or revealed choice has advantages and disadvantages [19], [27].

Survey Respondent Sampling

Several techniques can be used to sample a population [15], however any sampling design starts with defining the target market population, i.e., what consumers (customer background distribution) buy what (competing) products, the target market population size, etc. and a definition of the sampling unit, i.e., a customer or a product purchase. Random sampling entails randomly drawing an observation (customer or product purchase) such that the probability of that observation to be drawn equals 1/N, where N represents the market population. Random sampling cannot adequately capture the choice behavior of a very small population subgroup. This issue can be addressed using stratified random sampling [15], which divides the market population into mutually exclusive and exhaustive segments. Random samples are then drawn from each market segment. A demand model for each market segment can be constructed to predict each market segment's demand, which can then be properly weighted to arrive at an unbiased estimate for the total market demand.

3.3 Phase III - Modeling

Phase III is a quantitative process to generate the demand model. Based on the collected discrete choice data, whether they are revealed or stated choice, modeling techniques such as logit [15], [20] or probit [21, 22] can be used to create a choice model that can predict the choices individual customers make and forecasts the market demand for a designed artifact. Details of logit have been introduced in Section 2.

The predictive capability of a demand model depends among other things on the customer-oriented design attributes considered and the form of the utility function used in the demand model. An important step in Phase III is to pre-determine the functional form of the utility function *W*, shown in Eqn. 2. It is common to initially assume a linear shape of the customer utility function and then to test different functional shapes (e.g., quadratic, exponential) for improvement of the model fit. However, a model obtained using such a try-and-improve approach lacks econometric reasoning (causality). The model may fit the (survey) data well but its predictions may lack accuracy. In this work, we propose to use the Kano method, introduced in Section 2, to facilitate the identification of the appropriate functional relationship between customer satisfaction and product performance. Using the Kano method, the various customer-oriented attributes ***A*** are first classified into different categories, such as those shown in Fig. 2. Designers can then choose different function forms for different attributes. For example, quadratic or exponential function shapes can be used to fit the customer-oriented attributes that are classified as must-be and excitive, respectively. This approach is expected to better capture the underlying behavior of consumers as opposed to randomly trying different functional shapes without proper econometric reasoning.

Representing customer-oriented attributes (***A***) using a sufficient number of quantifiable engineering design attributes in a choice model is often desirable to facilitate engineering decision-making. For instance, to capture the sound quality as experienced by the car occupants (may depend on the occupant's position in the car), the customer-oriented attributes could include: noise level, harmonics, and frequency. When these attributes are included, the demand model can be used to guide engineering decision-making related to air intake design, engine configuration, firing order, exhaust design, engine mount design, noise insulation, etc. However, while including more explanatory variables (attributes) may improve the model fit as with each additional variable more data variance can be explained, using too many explanatory variables may lead to the model fitting aspects of the data that are not due to underlying parametric features but due to for instance sampling variability, which is not desirable. Two criteria can be used for comparing alternative model fits and for determining whether including additional explanatory variables is useful. They are the Akaike's Information Criterion (AIC), Eq. 5, and the Bayesian Information Criterion (BIC) [28], Eq. 6. Both criteria penalize models for having too many explanatory variables.

$$AIC = -2L + 2p \tag{5}$$

$$BIC = -2L + p\ln(n), \tag{6}$$

where L is the log-likelihood, p the number of explanatory variables and n the number of observations (sample size). For both criteria, the best-fitting model is the model with the lowest score. A difference of 6 points on the BIC scale indicates strong evidence that the model with the lower value is preferred [30].

Another issue that may arise when using large numbers of explanatory variables is collinearity, that is, some explanatory variables may be explained by combinations of other explanatory variables. If collinearity occurs then these explanatory variables cannot be used in the choice model simultaneously. Factor analysis [28] or latent variable modeling [29] could be used to combine customer-oriented attributes (explanatory variables) that are correlated to each other into a fewer number of factors. Another solution approach is to constrain the (beta) coefficients of collinear variables to be equal.

3.4 Phase IV- Demand Estimation

The choice model obtained through Phases I to III can be used to predict the choice probabilities for each alternative in the choice set given a customer's background (***S***) and descriptions of the choice alternatives. Sample enumeration is the most advanced and most accurate approach to determine a product's market demand for the entire market population. A logit demand model is shown in Eq. 7, that estimates demand based on sample enumeration using random samples of the market population *N*, where *i* denotes the choice alternative and *n* being the sampled individual.

$$Q(i) = \sum_{n}^{N} \Pr_n(i) = \sum_{n}^{N} \frac{e^{W_{in}}}{\sum_{l=1}^{J} e^{W_{ln}}} \tag{7}$$

The accuracy of demand prediction can be improved by estimating a choice model per market segment to account for systematic variations of taste parameters (β coefficients) among population subgroups. Ultimately one can assume that taste parameters are log normal distributed across the market population [15]. Including customer specific data in the customer background ***S*** that relate to the customer's (potential) use of the product can improve the accuracy of the demand predictions, e.g., in case of a car one can think of annual mileage driven, type of usage (commuting/recreational), etc. This product usage data can be recorded for each respondent when collecting the customer data and incorporated in the demand model.

A different approach for estimating the market demand is to use the choice model to predict the average choice probabilities (i.e., market shares) of the market population, (e.g., by using sample enumeration). A separate, specialized, model can be formed to estimate the total market sales volume. An advantage of this approach is that a separate model for predicting the market sales volume may be more accurate by accounting for economic growth, seasonal effects, market trends, etc., potentially leading to more accurate demand predictions.

4 IMPLEMENTATION EXAMPLE

In this section we show an implementation of the proposed demand modeling approach to constructing a vehicle demand model with emphasis on evaluating engine design changes in a DBD model. The demand model developed in this case study can be used to assess the impact of engine design changes on vehicle demand, facilitating the evaluation of engine design and making proper tradeoffs between performance and cost. Twelve vehicles (7 models, 12 trims) are considered in the demand model representing the midsize car segment, which includes vehicles like the Ford Taurus, Toyota Camry, and Honda Accord. All data illustrated are normalized to protect proprietary rights of providers. From the historical data collected by JD Power, it is found that compared to other car segments, the customers of midsize cars often pay more attention to vehicle engine performance when choosing vehicles, thus facilitating the test of our approach. Our implementation is subject to the assumption that customers only consider the 12 vehicle trims when purchasing a vehicle. The demand model developed is a static model, as such demand changes over time are not considered. In “what if” studies and DBD optimization, we assume that only the design of one vehicle changes at a time while the other designs are kept the same.

4.1 Vehicle Demand Modeling - Attributes and Choice Set Identification

Based on J.D. Power’s Vehicle Quality Survey (VQS), we identify five groups of top-level customer desires related to vehicle choice at the vehicle system level. These are: engine/transmission performance, comfort and convenience, ride and handling performance, product image, and price/cost (see Table 1). For reasons of simplicity we do not consider customer desires related to sound system, seats, and style. In Section 3.1 we detailed the process of translating customer desires into quantifiable customer-oriented design attributes that are meaningful to both the demand-modeling specialist and to a design engineer. Specific customer desires can be identified for each top-level, vehicle system, customer desire. The customer-oriented design attributes considered in our model are presented in Table 1, which shows a representative mapping of top-level customer desires via the specific customer desires to the customer-oriented design attributes. Take engine and transmission performance as an example. The specific customer desires include performance during rapid acceleration, passing power at highway speeds, and also a pleasant sound while idling and at full throttle acceleration and low vibration levels. Interaction between engineering experts at the Ford Motor Company and market research specialists from J.D. Power helped identify the customer-oriented design attributes corresponding to their specific customer desires. Linking the customer-oriented design attributes, further down the hierarchy, to design options, is also an important activity of designers. However, it is not covered in this case study. An example of this linking can be found in our paper on DBD [2] for a universal motor case study. The design options in this vehicle engine design case study are represented by the different settings of attribute levels.

Table 1. Customer-oriented design attributes structure for vehicle engine design example

Top-level customer desires	**Specific customer desires**	**Customer-oriented attributes A**
Engine and Transmission Performance	Performance	Horsepower Torque Low-end torque Displacement Type (I4/V6)
	Noise	Noise @ 2000 rpm (highway) Noise @ 4000 rpm (accelerating) Noise @ rpm of max Hp (db)
	Vibration	Overall Vibration Level Vibration @ 2000 rpm (highway) Vibration @ 4000 rpm (accelerating)
Comfort and Convenience	Comfort	Front legroom Front headroom Rear legroom Rear headroom
	Convenience	Trunk space Range between fuel stops
Ride and Handling	Handling	Roll Gradient (deg/g)
	Steering	SWA@0.5 g (deg) Window Rolling Parking Efforts Static Parking Efforts

	Ride	Choppiness (M/sec^2/minute)
Product Image	Brand	Vehicle Make
	Origin	USA/import
	Reliability	IQS (Initial Quality Index)
	Durability	VDI (Vehicle Dependability Index)
	Vehicle Size	Vehicle mass Vehicle width Vehicle length
Product Cost	Acquisition Cost	MSRP_price Rebate
	Usage Cost	APR Fuel Economy Resale index

4.2 Vehicle Demand Modeling – Data Collection

The demand model is created using revealed choice data at the respondent level provided by J.D. Power. The data consist of 2,552 observed individual vehicle purchases (of the seven vehicles – 12 trims considered in this case study) in the USA of the year 2000 vehicle market, including respondents' background. The values of customer-oriented attributes related to the general vehicle descriptions of the 12 discrete choices, such as weight, fuel economy, and legroom are obtained from Ward's Automotive. The values of some other customer-oriented attributes such as, ride, handling, noise, and vibration are provided by Ford. A representative part of the choice set input data table for one customer is presented in Table 2.

Table 2. Partial demand model input data table (normalized)

Cus.id.	vehicle id.	observed choice	Customer background ***S***			Customer-oriented attributes ***A***			
			gender	age	income	Msrp price	Horse power	torque	Fuel economy
1	1	0	0	27	5	1.07	1.13	1.09	0.96
1	2	1	0	27	5	0.87	0.89	0.85	1.15
1	3	0	0	27	5	1.15	1.09	1.02	0.98
1	4	0	0	27	5	1.02	1.06	1.02	0.90
1	5	0	0	27	5	1.05	1.08	1.12	0.98
1	6	0	0	27	5	0.89	0.77	0.82	1.12
1	7	0	0	27	5	0.96	1.04	0.94	1.00
1	8	0	0	27	5	0.89	0.93	0.97	1.00
1	9	0	0	27	5	1.07	1.02	1.10	1.00
1	10	0	0	27	5	1.03	0.92	0.98	1.02
1	11	0	0	27	5	1.11	1.23	1.16	0.98
1	12	0	0	27	5	0.89	0.83	0.94	0.94

For each respondent there are 12 rows of data in the database, one for each choice alternative, each row containing the customer background, the customer-oriented attributes that describe the vehicle, and the respondent's observed choice (real purchase). The customer choice is treated as a binary variable, and in this particular case the customer selected vehicle 2. In total the database contains 30624 observations (2552 respondents * 12 vehicles). The correlation of a number of attributes of the collected data is presented in Table 3.

Table 3. Partial correlation matrix of vehicle attributes and customer background

	gender	age	income	USA/import
gender	1			
age	-0.192	1		
income	-0.074	-0.176	1	
USA/import	0.150	-0.220	0.087	1
msrp_price	0.006	-0.041	0.141	0.183
rebate	-0.101	0.256	-0.141	-0.869
apr	-0.072	0.173	-0.017	-0.425
resale index	0.178	-0.215	0.031	0.869
vdi (dependability)	-0.117	0.036	0.024	-0.746
iqs (initial quality)	-0.162	0.187	-0.059	-0.928
horsepower/mass	-0.011	-0.104	0.180	0.212
torque/mass	-0.051	-0.005	0.148	0.013
low-end torque/mass	-0.087	0.036	0.120	-0.255
fuel economy	0.127	-0.047	-0.102	0.444
fuel range	0.138	-0.063	-0.045	0.680
wheel base	-0.106	0.076	0.050	-0.667
vehicle width	-0.119	0.157	-0.066	-0.918
vehicle length	-0.149	0.154	-0.038	-0.907
front-headroom	-0.013	-0.103	0.145	0.290
front-legroom	0.072	-0.094	0.116	0.762
rear-headroom	-0.162	0.132	0.053	-0.695
rear-legroom	-0.140	0.157	0.013	-0.731
trunk space	-0.132	0.139	0.004	-0.844

The variables gender and USA/import of Table 3 are binary variables that is, female = 1, and import = 1, otherwise 0. Some conclusions can be deducted from Table 3. For instance, the negative sign of the correlations related to gender for wheel base, vehicle width, and vehicle length indicate that women apparently buy smaller cars, and the negative coefficient (-0.220) for USA/import indicates that older consumers tend to prefer American built cars. The negative coefficient for rebate and USA/Import (-0.869) reveals that imports are generally sold with smaller rebates. The correlation between customer background (gender, age, and income) and customer-oriented attributes appears to be very weak, which is desirable. Highly correlated variables are prone to being collinear. Further, *high correlation between the dependent variable (in this case the vehicle choice) and independent explanatory variables (i.e., customer-oriented design attributes) implies that few variables are sufficient to predict vehicle choice, limiting the use of many variables (customer-oriented attributes) desirable for engineering design decision-making*.

4.3 Vehicle Demand Modeling – Multinomial Logit

In this case study we use STATA (www.stata.com) to estimate the choice model. STATA employs a substitute for multinomial logit: grouped logit. Grouped logit considers multinomial choice (i.e., one vehicle picked from the choice set) as a grouped set of binary choices (i.e., pick/not-pick; 0 or 1). Only one binary choice of a group is allowed positive (i.e., a pick).

A linear customer utility function shape is initially considered for the utility function used in the logit choice model (Eq. 2). All customers share the same utility function coefficients, i.e., market segmentation is not considered. We tested over 200 customer utility functions with different combinations of linear and interaction items. Eventually a model using 38 explanatory variable items (including interactions) is selected based on the Bayesian Information Criterion score (BIC) (see description in Section 3.3). With the selected model, the observed and the estimated market shares for the 12 vehicle models are shown in Table 4. Table 4 shows that the observed choice rate/market shares and the market shares as predicted by the model match quite well. The MS_R2, i.e., R2 error measure of the observed market shares vs. predicted market shares for this model is 0.996.

As proposed earlier in Section 2.3, the Kano method is used to further improve the predictive accuracy by identifying appropriate shapes for the customer utility function of the choice model (Equation 2). According to Kano study results at Ford, all key customer-oriented attributes should be considered as *basic*, except for *fuel economy beyond 27.5 mpg* which can be classified as *excitive*. The econometric reasoning for this is as follows; fuel economy is considered *basic* if the fuel mileage is near what is expected for the vehicle's class, in this case the midsize market segment. But, when the fuel mileage is significantly higher than its competitors then it becomes a distinguishing feature, e.g., "I bought this car because of its remarkable fuel economy." We test a quadratic function shape for the key customer-oriented attributes "fuel economy" and "range between fuel stops" in the customer utility function of the demand model. The demand model using the utility function shape as assessed by the Kano method should be preferred as it provides a better fit of the data given the more than six point difference in the BIC score shown in Table 5.

Table 4. Observed and estimated market shares for vehicle demand model

Vehicle id.	Choice Rate (#)	Market Shares observed	Market Shares estimated
1	251	0.098354	0.098814
2	190	0.074451	0.074544
3	335	0.131270	0.130938
4	220	0.086207	0.086117
5	231	0.090517	0.090972
6	192	0.075235	0.075440
7	199	0.077978	0.077447
8	167	0.065439	0.064866
9	67	0.026254	0.027256
10	435	0.170455	0.170324
11	213	0.083464	0.083507
12	52	0.020376	0.019776

Table 5. Comparison between linear and quadratic customer utility function fits

	Kano (quadratic)	Regular (linear)
MS_R2	0.998293	0.995984
Max. likelihood	-5820.69	-5831.48
BIC	11930.61	11941.85

4.4 Cross Validation of Demand Model

Due to the scope of our study, we cannot use the current market demand data to validate the demand model created using the data of year 2000. The approach we take for validating the obtained vehicle demand model is through the technique of cross-validation [31] which does not require the collection of additional data. The data set consisting of 2,552 individuals is divided into 5 subsets of approximately equal size using random sampling. The model is fitted to the combined data of 4 out of the 5 datasets. The fitted model is then used to predict the choice for the remaining choice set and the r^2

value for the market shares, which is used as error measure, is calculated. This procedure is repeated 5-fold, every time using a different data set from the 5 data sets for prediction and error measure calculation. The r^2 value of the demand model fitted on the full data set is 0.99. The r^2 value decreased to an average 0.92 for the 5 cross validation tests, which is still an acceptable value. The cross validation helps us build more confidence in using the proposed DCA approach to demand modeling and demand prediction. It also shows that the accuracy of the obtained demand model is satisfactory.

4.5 Market Share Prediction and "*What if*" Scenarios

The impact of customer-oriented design attribute changes (which reflect engineering design changes) on the vehicle market shares can be predicted by updating the vehicle descriptions and recalculating the predicted choice probabilities for each individual. To illustrate how the demand model can be used to study the impact of design changes and possible actions of competitors, we consider the following "*what if*" scenarios. Vehicle 11 and vehicle 12 are two trims of one vehicle model from the same manufacturer, a basic version and a more powerful luxury version. We assume that the manufacturer decides to improve the fuel efficiency of the base model (vehicle 11) by 10%, the impact on the market shares is shown in Table 6 under the heading "scenario 1." It appears that increasing the fuel efficiency of vehicle 11 increases its market share from 8.35 to 9.25% but it also shows that vehicle 12's market share is negatively affected. This negative impact of feature upgrades of a product on other members of the same manufacturer is known in marketing literature as "cannibalism." It implies that *the product being designed should not be considered in isolation*. Scenario 2 shows the impact on the market shares if the producer of vehicle 5 decides to introduce a rebate of $500 to boost its market share. Finally, Scenario 3 shows the impact of increasing vehicle 12's engine power by 5%.

In addition to the market share, the feasibility or the desirability of design changes depends on the impact on profit, which necessitates the consideration of the cost of such changes. This is considered in the DBD design alternative selection example in the next section.

Table 6. Results of "*What if*" scenarios

Veh. Id.	Market shares (%) base	scenario 1.	scenario 2.	scenario 3.
1	9.84	9.81	9.41	9.38
2	7.45	7.47	7.18	7.15
3	13.13	12.91	12.42	12.37
4	8.62	8.53	8.21	8.18
5	**9.05**	**8.81**	**12.15**	**12.08**
6	7.52	7.37	7.12	7.08
7	7.80	7.63	7.38	7.34
8	6.54	6.45	6.20	6.17
9	2.63	2.71	2.62	2.60
10	17.05	17.09	16.49	16.41
11	**8.35**	**9.25**	**8.92**	**8.87**
12	**2.04**	**1.95**	**1.89**	**2.36**

4.6 Decision-Based Design for Vehicle Engine Alternative Selection

We integrate the vehicle demand model with a cost model into a DBD model (see its framework in Figure 1). The DBD model is used to select the best engine design from 5 different engine design configurations, considered for vehicle 11. To simplify matters, the design options are represented by the setting of the customer-oriented design attributes ***A*** rather than the design options ***X***. The cost model considers the impact on cost of performance improvements related to power, torque, and low-end torque. Low-end torque is the maximum torque an engine produces at approximately 2000 rpm and is important for accelerating to pass a vehicle when driving at highway speed.

The 5 alternative engine designs for use in vehicle 11 are presented in Table 7. Engine design 1 offers increased power, torque, and low-end torque with 3% and a price increase of 5% relative to the performance of the existing engine used in vehicle 11. Engine design 2 is similar in performance to Engine design 1 but is sold at the base price. Engine design 3 offers a 3% power and 5% price increase relative to the base model, while the performance of Engine design 4 is the same as Engine design 3 but sold at the base price. A fifth engine design alternative (Engine design 5) is added by considering re-using an existing engine design for vehicle 11 of a different model, which is less powerful but enables a reduction in price of 5% when compared with the base model.

Table 7. Design alternatives for Decision-Based Design case study

	Design Alternative (Vehicle 11) (% change attribute level)				
design #	**1**	**2**	**3**	**4**	**5**
price	5	0	5	0	-5
hp	3	3	3	3	0
torque	3	3	0	0	-10
low-end torque	3	3	0	0	-10

The market size *M* of the 12 midsize vehicles is estimated at 1,000,000 vehicles annually. Uncertainty is introduced by assuming a normal distribution of the market size with standard deviation of 50,000 vehicles. To facilitate the consideration of the impact of engine changes of vehicle 11 on vehicle 12 and on the same manufacturer's profit we assume that vehicle 12 contributes $1,100 per vehicle to the profit. The manufacturer's expected utility is obtained by assuming a risk averse attitude, which is obtained by taking the log of the profit. The DBD optimization problem, shown in Figure 4, is formulated as follows: *given* the vehicle demand model (Section 4.3), and decision-maker's risk attitude (log of profit), *maximize* the expected utility of profit *with respect to* price, horsepower, torque, and low-end torque.

The market share impact (% change) for the 12 vehicles and the impact on the profit (in millions of dollars) of vehicle 11's manufacturer and the expected utility for the five design alternatives (vehicle 11) are presented in Table 8. For example, it is noted that under design alternative 1, increasing the horsepower, torque, and low-end torque leads to a 9.7% gain in the market share of vehicle 11 when increasing the price by 5% and a drop of the market share of vehicle 12, also produced by the same manufacturer, by 3.8%. When considering the (maximum of) expected utility of the five design alternatives then it appears that design alternative 4, consisting of a 3% torque increase while leaving the price unchanged, should be preferred. It should be noted that even though the DBD model is used to select the best design among a set of discrete alternatives in this study, the DBD model can be used to select the best alternative among a range of continuous decision variables via optimization.

GIVEN

Market size M 1000,000 vehicles annually

Standard deviation σ_M 50,000

Customer-oriented attributes *A*

Demand model Q

The demand model is obtained using the multinomial logit technique to fit the discrete choice survey data

Cost model C

Determines the relationship between ***A*** and C

Corporate interests *I*

None other than the single selection criterion, V

Single criterion V

Net revenue $V = Q\,P - C$

Utility function $U(V)$

$U(V) = \log(V)$

Market Data *S* (Socioeconomic and demographic attributes)

Data related to gender, age, and income

FIND

Key customer attributes ***A*** and price P

SUBJECT TO

N/A

MAXIMIZE

Expected utility of the net present value of profit V

Figure 4. Vehicle engine DBD description

Table 8. Market share impact (% change), profit ($ million), and expected utility for case study

Veh. Id.	Design Alternative 1	2	3	4	5
1	-0.4	-0.6	0.1	-0.1	0.2
2	-0.8	-0.9	-0.3	-0.5	-0.1
3	-1.1	-1.3	-0.6	-0.9	-0.5
4	-1.0	-1.1	-0.5	-0.7	-0.3
5	-0.3	-0.5	0.1	-0.1	0.4
6	-0.6	-0.7	-0.1	-0.4	0.1
7	-1.5	-1.7	-1.1	-1.3	-0.8
8	-1.8	-1.9	-1.3	-1.5	-1.0
9	2.9	2.7	3.4	3.1	3.7
10	-1.0	-1.1	-0.5	-0.7	-0.3
11	**9.7**	**11.4**	**4.4**	**7.0**	**2.0**
12	-3.8	-3.9	-3.4	-3.6	-3.0
Exp. Impact on Profit	77.77	77.00	87.60	89.10	31.01
Exp. Utility	90.84	90.78	91.43	91.52	86.24

5 CONCLUSION

Building upon our earlier work on using the discrete choice analysis approach to demand modeling we develop in this paper guidelines for implementing the discrete choice demand modeling approach in the context of Decision-Based Design. The transformation of top-level customer desire groups down to customer desires, and further into customer-oriented design attributes and engineering design attributes in general, is introduced to bridge the gap between market analysis and engineering modeling in the process of demand analysis. As such, the customer-oriented design attributes form the link between the design options and demand, and eventually profit, thus facilitating engineering design decision-making. Kano's Method is adapted to provide econometric justification for selecting the shape of the customer utility function, which better captures the underlying purchase behavior, and enhancing the predictive capability of demand models. The proposed approaches are demonstrated using a real (passenger) vehicle engine design problem as a case study in collaboration with the market research firm J.D. Power and Associates and the Ford Motor Company. The obtained demand model is shown to be satisfactory through cross validation.

It should be noted that different from some existing design approaches that construct a single utility function for a group of customers, the proposed DBD approach optimizes a single-criterion utility function that is related to the profit of a product. As a part of the profit estimation, the demand modeling based on DCA utilizes separate utility function for each individual and aggregates the customer choices (not preferences) by summing the choice probabilities across individual decision-makers (customers), thus avoiding the paradox associated with aggregating the utility or preference of a group of customers.

The demand modeling approach developed in this work, along with the Decision-Based Design framework are expected to facilitate the communication and collaboration of a company's employees in engineering, marketing, and management towards achieving the enterprise' goal of making profit. Application of the methodologies developed in this work is expected to lead to more competitive products because products will be improved in a systematic way, considering not only the engineering requirements, but also the business interests, customers' preferences, competitors' products, and market conditions.

Our proposal to employ Kano's Method to select and econometrically justify the customer utility function shape is a first step in improving the predictive capabilities of the proposed demand modeling approach. Another approach being examined is to enhance the capturing of the customer's perception of the product attributes through consideration of the unobservable top-level customer desires in the customer utility function using latent variables. In addition, the impact of marketing incentives, distribution, and competition will also be addressed within the DBD framework in our future work.

ACKNOWLEDGMENTS

We like to thank the specialists at J.D. Power & Associates and the Ford Motor Company for their thoughtful contributions and their efforts to gather data for our vehicle demand model. We also thank J.D. Power & Associates for providing the opportunity to work with vehicle demand modeling experts during an internship in Summer 2002. The support from NSF grant DMII 0217702 are acknowledged.

REFERENCES

[1] Hazelrigg, G.A., "A Framework for Decision Based Engineering Design", *ASME Journal of Mechanical Design*, Vol. 120 , 653-658, 1998.

[2] Wassenaar, H.J. and Chen, W., "An Approach to Decision Based Design with Discrete Choice Analysis for Demand Modeling", *ASME Journal of Mechanical Design*, 125(3), 490-497, 2003.

[3] Gu, X., Renaud, J.E., Ashe, L.M., Batill, S.M., Budhiraja, A S., and Krajewski, L.J., "Decision-Based Collaborative Optimization", *ASME Journal of Mechanical Design*, 124(1), 1-13, 2002.

[4] Tappeta, R.V. and Renaud, J.E. "Interactive Multiobjective Optimization Design Strategy for Decision Based Design", *ASME Journal of Mechanical Design*, 123(2), 205-215, 2001.

[5] Wan, J. and Krishnamurty, S., "Learning-Based Preference Modeling in Engineering Design Decision-Making", *ASME Journal of Mechanical Design*, 123(2), 191-198, 2001.

[6] Thurston, D.L., "Real and Misconceived Limitations to Decision Based Design With Utility Analysis", *ASME Journal of Mechanical Design*, 123(2), 176-186, 2001.

[7] Thurstone, L., "A Law of Comparative Judgment", *Psychological Review* 34, p273-286, 1927.

[8] Luce, R., *Individual Choice Behavior: A Theoretical Analysis*, *Wiley*, New York, 1959.

[9] Marschak, J., "Binary Choice Constraints on Random Utility Indicators", *Stanford Symposium on Mathematical Methods in the Social Sciences*, K. Arrow, ed., Stanford University Press, Stanford, California, 1960.

[10] Tversky, A., "Elimination by Aspects: A Theory of Choice", *Psychological Review* 79, p281-299, 1972.

[11] Green, P.E. and Srinivasan, V., "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice", *Journal of Marketing,* 1990.

[12] Green, P.E. and Srinivasan, V., "Conjoint Analysis in Consumer Research: Issues and Outlook", *Journal of Consumer Research*, Vol. 5, 1978.

[13] Green, P.E. and Wind, Y., "New Ways to Measure Consumer Judgments", *Harvard Business Review*, 1975.

[14] Louviere, J.J, "Why Stated Preference Discrete Choice Modelling is NOT Conjoint Analysis (and what SPDCM IS)", *Memetrics white paper*, 2000.

[15] Ben-Akiva, M and Lerman S.R., *Discrete Choice Analysis*, The MIT Press, Cambridge, Massachusetts, 1985.

[16] Li H. and Azarm, S., "Product Design Selection under Uncertainty and with Competitive Advantage", *ASME Design Technical Conference,* DETC2000/DAC-14234, Baltimore MD, 2000.

[17] Cook, H. E., *Product Management: Value, Quality, Cost, Price, Profit, and Organization*, Chapman & Hall, London, 1997.

[18] Donndelinger, J., Cook, H.E. “Methods for Analyzing the Value of Automobiles,” SAE Paper 970762, Warrendale, PA, *Society of Automotive Engineers, Inc.*, Feb. 1997.

[19] Besharati, B., Azarm, S., and Farhang-Mehr, A., "A Customer-Based Expected Utility for Product Design Selection", *Proceedings of the ASME Design Engineering Technical Conference*, Montreal, Canada, 2002.

[20] Wassenaar, H.J., “*An Approach to Decision-Based Design*”, PhD Dissertation, University of Illinois at Chicago, October 2003.

[21] Hensher, D.A. and Johnson, L.W., “*Applied Discrete Choice Modeling*”, Halsted Press, New York, 1981.

[22] Daganzo, C., “*Multinomial Probit, the theory and its application to demand forecasting*”, Academic Press Inc., New York, 1979.

[23] Shiba, S., Graham, A., Walden, D., ***N**ew American TQM: four practical revolutions in management*, Productivity Press, Cambridge, Mass. 1993.

[24] Saari, D.G., “Mathematical structure of voting paradoxes. I; pairwise vote. II; positional voting”, *Economic Theory* 15, p1-103, 2000.

[25] Otto K.N. and Wood, K., *Product Design: Techniques in Reverse Engineering and New Product Development*, Prentice Hall, Upper Saddle River, NJ, 2001.

[26] Krueger, R.A., “*Focus groups: a practical guide for applied research*”, 2nd ed, Thousand Oaks, Sage Publications, California, 1994.

[27] Louviere, J.J., Hensher, D.A., Swait, J.D., “*Stated Choice Methods, Analysis and Application*”, Cambridge University Press, 2000 [24] Hair, J.F., (editor), Anderson, R.E., Tatham, R.L., Black, W.C., “*Multivariate Data Analysis*”, 5th ed, Prentice Hall College Div,., 1998.

[28] Hastie, T., Tibshirani, R., and Friedman J., *The Elements of Statistical Learning*, Springer 2001.

[29] Loehlin, J.C., *Latent Variable Models, an introduction to factor, path, and structural analysis*, 3rd ed., Mahwah, NJ: L. Erlbaum Associates, 1998.

[30] Raftery, A., “Bayesian Model Selection in Social Research”, *Social Methodology*, 1995.

[31] Breiman, L., Spector, P., ‘Submodel selection and evaluation in regression: The X-random case”, *International Statistical Review*, 60, 291-319, 1992.