Exploring the Role of Interaction Effects in Visual Conjoint Analysis

Brian Sylcott
Department of Engineering,
East Carolina University,
Greenville, NC 27858
e-mail: sylcottb15@ecu.edu

Jeremy J. Michalek
Department of Mechanical Engineering,
Engineering and Public Policy,
Carnegie Mellon University,
Pittsburgh, PA 15213
e-mail: jmichalek@cmu.edu

Jonathan Cagan
Department of Mechanical Engineering,
Carnegie Mellon University,
Pittsburgh, PA 15213
e-mail: cagan@cmu.edu

In conjoint analysis, interaction effects characterize how preference for the level of one product attribute is dependent on the level of another attribute. When interaction effects are negligible, a main effects fractional factorial experimental design can be used to reduce data requirements and survey cost. This is particularly important when the presence of many parameters or levels makes full factorial designs intractable. However, if interaction effects are relevant, main effects design can create biased estimates and lead to erroneous conclusions. This work investigates consumer preference interactions in the nontraditional context of visual choice-based conjoint analysis, where the conjoint attributes are parameters that define a product's shape. Although many conjoint studies assume interaction effects to be negligible, they may play a larger role for shape parameters. The role of interaction effects is explored in two visual conjoint case studies. The results suggest that interactions can be either negligible or dominant in visual conjoint, depending on consumer preferences. Generally, we suggest using randomized designs to avoid any bias resulting from the presence of interaction effects. [DOI: 10.1115/1.4031054]

Introduction

When developing a design solution, engineers often find information about the preferences of their target users to be extremely useful. Conjoint analysis [1] is among the most popular preference elicitation methods used to gather data and construct quantitative models of consumer preference for product attributes. In conjoint analysis, data are collected from consumers in the form of product surveys. These surveys are controlled experiments that present participants with multiple hypothetical product profiles chosen to span the space of product attributes without conflating them [2]. This allows for the estimation of both main effects (describing preference for the levels of a product's attributes) and interaction effects (describing how preference for the level of one product attribute is dependent on the level of another attribute). Researchers can create surveys using a subset of all possible profiles to estimate specific effects when other effects are believed to be negligible. This results in a smaller survey size; however, if the effects that were assumed to be negligible are in fact significant, estimates will be biased [3].

In this work, case studies are used to investigate the validity of ignoring interaction effects for conjoint analysis cases that involve shape attributes. The goal of this work is to provide improved clarity about how the inclusion or exclusion of interaction effects can impact the performance of consumer preference models of product shape constructed through visual conjoint analysis.

Contributed by the Design Automation Committee of ASME for publication in the Journal of Mechanical Design. Manuscript received August 20, 2013; final manuscript received July 6, 2015; published online July 30, 2015. Assoc. Editor: Bernard Yannou.

Previous Work

There are several examples of conjoint analysis, as well as other methods, being used to model consumer preference for shape attributes [4–9]. Green and Srinivasan suggested that interaction effects are important in situations that involve surveying sensory phenomena, styling, and esthetic features [10]. The literature on multi-attribute utility theory has also emphasized the importance of nonlinear effects [11–14]. In some situations, missing interaction effects can seriously bias parameter estimates and result in faulty conclusions, while in other situations the inclusion of interaction effects offers little value with added cost and complexity. Additionally, the inclusion of interaction effects can add variance in parameter estimation which can negatively impact model performance [15].

In the following case studies, several different preference models with and without interaction effects are explored. Multiple goodness of fit measures are provided for each model in order to get a better understanding of how interaction effects influence the performance of visual conjoint based models. Contrary to previous findings, this work shows that interaction effects can have a major impact on model performance.

Approach

Here, the random utility discrete choice modeling approach, which assumes that each consumer selects among the set of product alternatives available the one that provides the greatest utility, is followed. A standard maximum likelihood estimation (MLE) [16] is applied to identify the part-worth values that best match observed choice data collected via choice-based conjoint questions. To evaluate the performance of alternative utility function forms that either include or ignore interactions, the log-likelihood (LL), hit rate (HR), and mean absolute share error (MASE) for hold out samples are reported. HR is calculated by comparing the observed choices with the alternative of highest observable utility in each choice set [17]. Each time the observed selection matches the alternative with the highest observable utility counts as a hit—otherwise it is a miss. MASE is used to evaluate how well the observed share of choices matches the predicted share of choices. It is calculated by taking the average of the absolute difference between the observed and predicted share for each design alternative. Together these metrics are used to evaluate the overall performance of the utility models in each of the following case studies.

Case Studies

Case 1: Vase Shape Preference Illustration

Methodology. This simplified example is used to illustrate the potential importance of interactions and their impact on model performance. In order to make this illustration as clear as possible, a simple model was developed using simulated response data. Visual conjoint studies often use outlines and silhouettes to represent product shape. The subject of this illustration is the outline of a flower vase. The vase consists of the four Bézier curves depicted in Fig. 1. The control points of the Bézier curves are parameterized so that three attributes of the vase vary. These attributes are the top to bottom width ratio, the height to average width ratio,

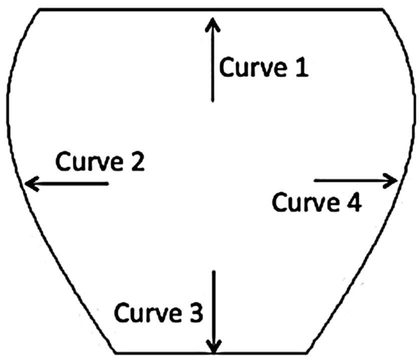

Fig. 1 Four Bézier curves that make up the vase model

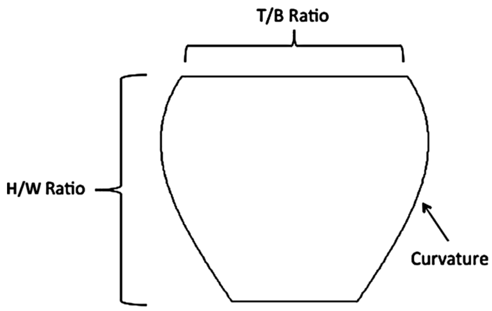

Fig. 2 Vase attributes

Table 1 Vase attribute scoring

Level	Top–bottom ratio	Height–width ratio	Curvature
0	6	4	4
1	3	2	2
2	9	6	6
Level		Height–width ratio × curvature	
1	1	+12	
1	2	−12	
2	1	−12	
2	2	+12	

and the curvature of the sides. Symmetry is enforced between curves 2 and 4. Each attribute has three levels. The attributes are depicted in Fig. 2.

The three attributes, each with three levels, yield 27 possible profiles. Each of the 27 profiles was scored based on the main attribute levels and the interaction between the height to average width ratio and side curvature attributes. Table 1 shows the scores associated with each level of the attributes and the significant interactions. These scores were used to simulate response data from 102 subjects. For each choice set, the simulated respondents selected the alternative with the highest score. In the event of a tie, respondents distributed their selections evenly over the highest rated alternatives. Two of the respondents provided some variance by ensuring each alternative was selected at least once. There are three possible outcomes for the selections {100,1,1}, {50,50,2}, or {34,34,34}.

In order to estimate main and interaction effects, a full factorial of 27 choice sets was developed. Kuhfeld's SAS macros [18] were used to organize the 81 profiles (three full factorial groups) into 27 choice sets (three different 100% D-efficient, nine question surveys) with three alternatives each resulting in an 100% D-efficient 27-question survey. SAS was used to design 27 additional questions as a hold out sample for the 27-question survey. An example trial is depicted in Fig. 3. Two alternative models are proposed in this section, one with interaction terms and one without.

Results—Negligible Interactions. A solution to each of the MLE models was obtained using the modified quasi-Newton method implemented in the *fminunc* function of MATLAB. In this scenario, interaction effects were not included in the ratings. The part-worth model results are presented in Table 2. Higher HR values indicate a higher degree of fit with the data. As MASE approaches 0% so does the difference between the observed and predicted choice shares.

The HR and MASE for the main effects model are comparable to the first-order interaction effects model, suggesting these effects have little influence on these fit metrics.

Results—Significant Interactions. In this scenario, ratings were based on the attribute values as well as the interaction between the height to average width ratio and side curvature attributes. Here, MATLAB was again used to find solutions to each of the MLE models. The part-worth model results are presented in Table 3. Due to the inclusion of interactions in the ratings, the main effects models were not sufficient to accurately estimate preference. As a result, there is a sizable difference in performance between the models that included interactions and the ones that did not. The interaction effects model performed better on all metrics than the main effects model.

Together these examples illustrate how interactions can be either crucial or negligible when developing an accurate model of consumer preference. Because the attributes that make up a shape are so interconnected it is likely, perhaps more so than with functional attributes, that interactions will be an important part of expressed preference. The significance of interactions in a visual conjoint model can also depend on how a shape is parameterized. Interactions have been shown to be critical when a shape is modeled with one representation and negligible when the same shape

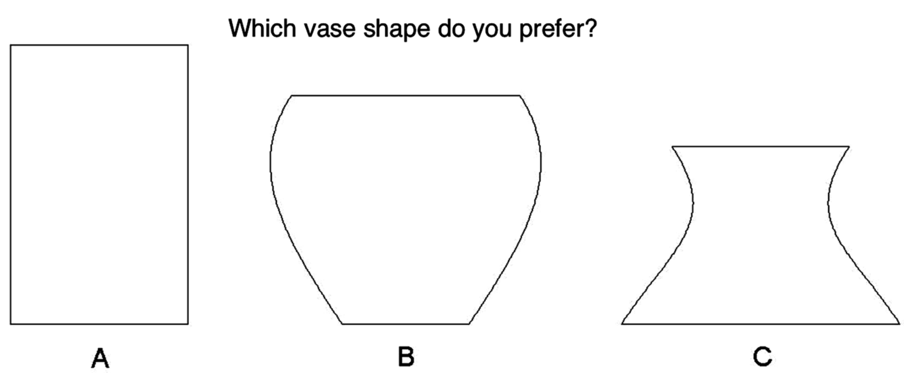

Fig. 3 Example vase preference trial

Table 2 Vase part-worth results for negligible interactions

		Main effects	Interaction effects
Hold out	LL	−1045	−1355
	HR	85%	82%
	MASE	10%	13%

Table 3 Vase part-worth results for significant interactions

		Main effects	Interaction effects
Hold out	LL	−3034	−3839
	HR	60%	84%
	MASE	35%	13%

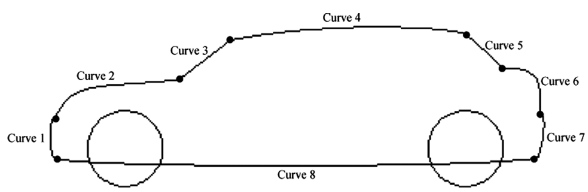

Fig. 4 Example vehicle

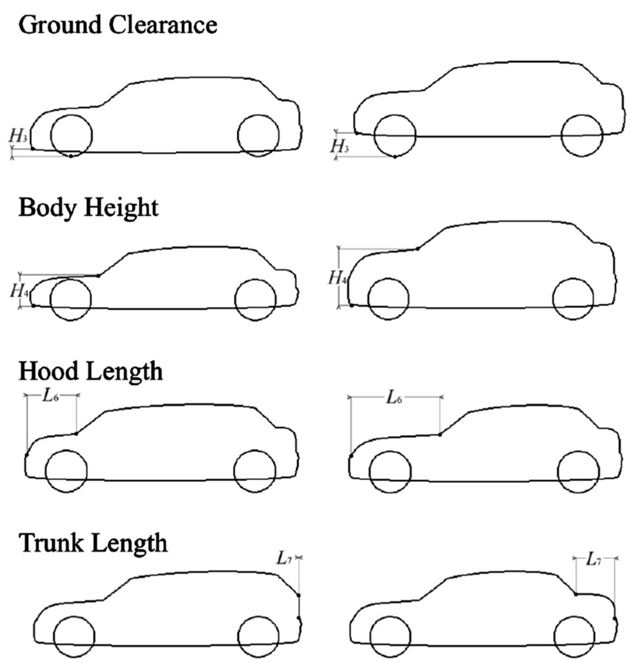

Fig. 5 Vehicle shape attributes

is modeled with a different representation [19]. As such, the significance of interactions terms in visual conjoint models depends not only on the shape and preferences for that shape but also on the way the shape is parameterized. Shape parameterization is an important consideration when designing representations in visual conjoint studies.

Case 2: Vehicle Shape Preference

Methodology. The first case was a simple example designed to illustrate how interaction effects can be either critical or negligible and how modeling decisions can impact performance. It is often difficult to know the extent that interaction effects will matter in a given situation beforehand, particularly in visual conjoint applications. Single surveys that are capable of estimating all main and interaction effects for every individual are often too large to be practical. An alternative approach is to use a randomized survey design for each participant. With a large enough sample, randomized surveys are capable of estimating both main and interaction effects.

In this case, randomized surveys are issued to real participants in order to capture their preference for vehicle shape. The subject of this study is a vehicle design depicted by line drawing silhouettes built using the scheme developed by Tseng et al. [8]. An example of the vehicle representation is shown in Fig. 4.

As shown in Fig. 4, these representations are the composition of eight Bézier curves. The control points of the curves are parameterized in a manner that allows several major features of the design to be varied continuously. In this study, four attributes were varied, the ground clearance, body height, hood length, and trunk length. These attributes are depicted in Fig. 5. Each of the four attributes consisted of three levels. The lower and upper bounds on each of the attributes are depicted in Fig. 5. The remaining level is midway between the bounds.

There were a total of 81 possible product profiles. The SAWTOOTH software package was used to organize the profiles in to random designs for each participant. Each survey question presented three design alternatives for the participants to choose from. A sample trial is shown in Fig. 6.

Each survey consisted of 18 questions to build the preference model, six hold out questions, and a repeat question to check for consistency. The SAWTOOTH SSI web tool was used to administer the survey. Participants were recruited using Amazon.com's Mechanical Turk site. There were a total of 109 survey participants. The average participant age was 38.5-yr old; there were 45 male and 64 female participants. Each participant was over 18-yr old.

In this case, no information about which, if any, interaction effects were important was available beforehand. Consequently, there was no guidance on whether main effects model could sufficiently represent preference. As shown previously, the impact interaction effects can have on model performance ranges from negligible to critical. A full factorial of choice sets would require 81 questions. This would be a fairly long survey, especially with the inclusion of hold out questions. As an alternative, each participant was given a random survey. The random approach has been shown to be robust in situations where the importance of interactions is not known beforehand [20].

Results. Two approaches were taken to model the preference exhibited in this case. First, the script developed by Train [21] was used to implement MLE and solve for $\boldsymbol{\beta}$. The results presented in Table 4 show the main, first, and first-, second-, and third-order interaction effects models performed better than the null model. Due to the large number, the regression coefficients are not listed. The models that included interaction effects outperformed the main effects model on all metrics.

Many of the first- and second-order interactions in these models were found to be statistically significant, and the presence of higher order interactions improved predictive capability measured via HR and MASE. For example, in the model with first-order effects there is a significant interaction that indicates respondents preferred low ground clearance on vehicles that had a long hood length. But significance of specific effects is not strongly robust over the different model specifications, and interpretation becomes challenging when many higher order interactions are present as intuitions may not hold.

Next, the latent class analysis routine within the SAWTOOTH software package was used to develop main effects model as well as one that included first-order interactions (higher order interactions were not available in SAWTOOTH) based on the survey responses. There were five segments in each model sized [13.50%, 45.20%, 17.90%, 9.00%, and 14.40%] and [10.00%, 43.30%, 10.00%, 27.70%, and 9.00%], respectively. Table 5 summarizes the performance of these models.

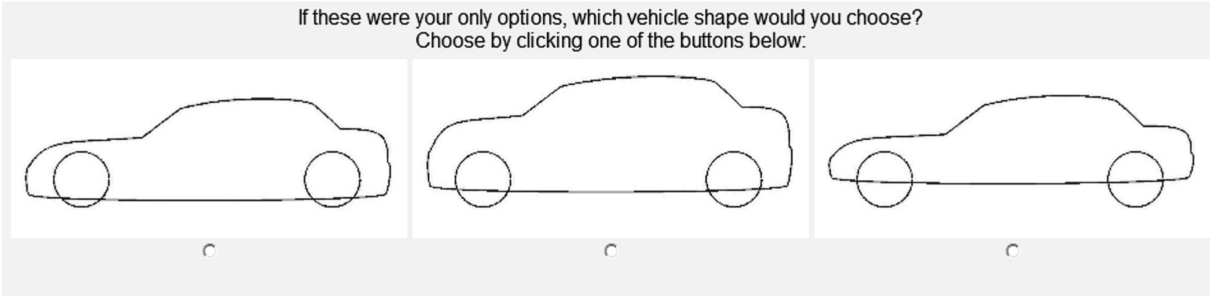

Fig. 6 Screenshot from vehicle preference survey

Table 4 Vehicle design preference model performance results (one segment)

		Null model	Main effects	First-order interactions	First-, second-, and third-order interactions
Hold out	LL	−718	−706	−680.34	−612
	HR	19%	47%	45%	54%
	MASE	16%	16%	14%	5%

Table 5 Vehicle design preference model performance results (five segments)

		Null model	Main effects	First-order interactions
Hold out	LL	−718	−708	−682
	HR	19%	23%	47%
	MASE	54%	24%	15%

As with the single-segment models, results from the multisegment models show better performance than the null model and improvement in all metrics when interactions are included in the model. A large percentage of the respondents remained in a single segment while the other segments were much closer in size to each other suggesting there was some heterogeneity in the dataset. However, the multisegmented model that included first-order interaction performed comparably to the single-segment model.

Overall, these results suggest that while it is possible to model preference for shapes without them, in many cases first order and even higher order interaction effects are important. Models that assume without evidence that higher order interactions are negligible in visual conjoint are at risk of introducing substantial bias and error. These results have important implications for engineering design. In some cases, there is a direct link between esthetic form and functional performance. For example, in the previously cited work by Tseng et al. [8] efforts are made to optimize esthetics, the vehicle's silhouette, and performance, aerodynamics, concurrently. The ability to accurately model preference for product form can be an important part of the design process that leads to more desirable products.

Discussion

In this work, the role of interaction effects in visual conjoint analysis was explored through two case studies. In case 1, main effects models are shown to be sufficient when interactions are known to be negligible. The interaction effects model slightly underperforms compared to main effects model. This is most likely caused by over fitting resulting from including the interaction terms. This example demonstrates how erroneously including interaction terms can negatively impact performance. One additional area of interest that will be explored further in future work is the relationship between the part-worths and the predefined scores. These case results illustrate how large the difference in performance can be between main effects and first-order interaction effects models when interactions are important to respondents. The interaction effects model outperforms the main effects model by 24% and 22% on HR and MASE, respectively. Inaccuracies in these types of models can negatively impact the design process as design decisions resulting from such models can lead to the development of undesirable products.

In case 1, the survey design that included only a subset of all possible choice sets (e.g., a full factorial of profiles organized randomly into choice sets) produced similar results on average as using all possible choice sets. However, some 100% D-efficient survey designs result in substantially lower performance than others, indicating that random assignment of profiles to choice sets can introduce error [19]. This error is mitigated if every respondent receives different randomized choice sets.

In case 2, most of the model performance improved with the inclusion of first-order interaction effects. In practice, firms use prior experience with products to select the subset of interaction effects which are expected to be significant. This set is likely to include some first-order effects and few, if any, higher order effects. The improvement in model performance gained from including second- and third-order interaction effects suggests that these effects should be investigated before assuming them negligible. This is particularly true when prior product data are unavailable. The single-segment model saw additional improvement with the inclusion of second- and third-order interaction effects suggesting that prior literature on vehicle shape preference that has ignored interactions may have bias and error in the interpretation of coefficients (importance of attributes) and in choice predictions. Although the multisegmented model indicated clear differences in preferences between the segments, the inclusion of first-order interactions proved to have a positive impact on overall model performance.

One limitation in this work is that the MLE approach is reliant on the independence of irrelevant alternatives (IIA) assumption. IIA was a reasonable assumption to make in case 1 as all of the simulated respondents selected the design with the highest score regardless of the other alternatives. In case 2, it is unclear whether this assumption holds. As an alternative to MLE, a hierarchical Bayesian approach can be employed. As such methods are more robust to IIA, they will be pursued in the future [22].

Conclusion

Since individual products vary so greatly it is difficult to make generalizations about the significance of interaction effects that will apply in all situations. However, in light of the results of these case studies the authors make two recommendations. First, because shape parameters are generally parts of a whole it is likely that the parameters that make up a shape and preference for those parameters will depend on one another in some fashion. It is recommended that models of shape preference include at least first-order interaction effects. Exploration into this possibility is preferred before opting for a fractional factorial design. Second, in the absence of prior knowledge of whether or not interactions are important for a given model parameterization, randomized designs should be used. This approach requires more data for a given level of statistical significance but is capable of estimating all interaction effects and averaging out bias associated with how the design profiles are organized into choice sets. Conjoint studies that assume without evidence that interactions are negligible and inappropriately use theory designed for cases where interactions are negligible (such as fractional factorial designs and main effects model specifications) can introduce substantial bias in interpretation and prediction. While this has always been an issue with conjoint studies, interaction effects appear to be particularly pronounced in visual conjoint applications, and we urge a correspondingly heightened degree of awareness and caution.

Acknowledgment

This work was supported in part by the National Science Foundation CAREER Award Nos. 0747911 and CMMI 1233864 and by Ford Motor Company. The views expressed are those of the authors and not necessarily those of the sponsors.

References

[1] Green, P. E., and Rao, V. R., 1971, "Conjoint Measurement for Quantifying Judgmental Data," J. Mark. Res., **8**(3), pp. 355–363.
[2] Huber, J., and Zwerina, K., 1996, "The Importance of Utility Balance in Efficient Choice Designs," J. Mark. Res., **33**(3), pp. 307–317.
[3] Huber, J., 2004, "Conjoint Analysis: How We Got Here and Where We Are (An Update)," Sawtooth Software Conference Proceedings, Sawtooth Software, Sequim, WA.
[4] Swamy, S., Orsborn, S., Michalek, J., and Cagan, J., 2007, "Measurement of Headlight Form Preference Using Choice-Based Conjoint Analysis," ASME Paper No. DETC2007-35409.
[5] Petiot, J.-F., and Dagher, A., 2010, "Preference-Oriented Form Design: Application to Cars' Headlights," Int. J. Interact. Des. Manuf., **5**(1), pp. 17–27.
[6] Reid, T. N., Gonzalez, R. D., and Papalambros, P. Y., 2010, "Quantification of Perceived Environmental Friendliness for Vehicle Silhouette Design," ASME J. Mech. Des., **132**(10), p. 101010.
[7] Kelly, J. C., Maheut, P., Petiot, J.-F., and Papalambros, P. Y., 2011, "Incorporating User Shape Preference in Engineering Design Optimisation," J. Eng. Des., **22**(9), pp. 627–650.
[8] Tseng, I., Cagan, J., and Kotovsky, K., 2012, "Concurrent Optimization of Computationally Learned Stylistic Form and Functional Goals," ASME J. Mech. Des., **134**(11), p. 111006.
[9] Sylcott, B., Cagan, J., and Tabibnia, G., 2013, "Understanding Consumer Tradeoffs Between Form and Function Through Metaconjoint and Cognitive Neuroscience Analyses," ASME J. Mech. Des., **135**(10), p. 101002.
[10] Green, P. E., and Srinivasan, V., 1990, "Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice," J. Mark., **54**(4), pp. 3–19.
[11] Thurston, D. L., 2001, "Real and Misconceived Limitations to Decision Based Design With Utility Analysis," ASME J. Mech. Des., **123**(2), pp. 176–182.
[12] Abbas, A. E., 2009, "Multiattribute Utility Copulas," Oper. Res., **57**(6), pp. 1367–1383.
[13] Ross, A. M., Hastings, D. E., Warmkessel, J. M., and Diller, N. P., 2004, "Multi-Attribute Tradespace Exploration as Front End for Effective Space System Design," J. Spacecr. Rockets, **41**(1), pp. 20–28.
[14] Kulok, M., and Lewis, K., 2007, "A Method to Ensure Preference Consistency in Multi-Attribute Selection Decisions," ASME J. Mech. Des., **129**(10), pp. 1002–1011.
[15] Hagerty, M., 1986, "The Cost of Simplifying Preference Models," Mark. Sci., **5**(4), pp. 298–319.
[16] Train, K. E., 2003, *Discrete Choice Methods With Simulation*, Cambridge University Press, New York.
[17] Orme, B. K., Alpert, M. I., and Christensen, E., 1997, "Assessing the Validity of Conjoint Analysis-Continued," Sawtooth Software Conference Proceedings, Sawtooth Software, Sequim, WA.
[18] Kuhfeld, W. F., 2010, "Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint, and Graphical Techniques," SAS Institute Inc., Cary, NC, http://support.sas.com/techsup/technote/mr2010title.pdf
[19] Sylcott, B., 2013, "Understanding the Role of Aesthetic Judgment in Consumer Choice and Preference Modeling," Ph.D. thesis, Carnegie Mellon University, Pittsburgh, PA.
[20] Chrzan, K., and Orme, B., 2000, "An Overview and Comparison of Design Strategies for Choice-Based Conjoint Analysis," Sawtooth Software Conference Proceedings, Hilton Head Island, SC, Sawtooth Software, Sequim, WA.
[21] Train, K., 1998, "Recreation Demand Models With Taste Differences Over People," Land Econ., **74**(2), pp. 230–239.
[22] Orme, B., 2000, "Hierarchical Bayes: Why All the Attention?" Sawtooth Software Conference Proceedings.