Eliciting Preference for Complex Products:
A Web-Based Upgrading Method

Young-Hoon Park, Min Ding, and Vithala R. Rao*

September 2005
Revised August 2006
Revised January 2007
Revised April 2007
Revised July 2007

---

* Young-Hoon Park is Assistant Professor of Marketing at the Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853-6201; phone: 607 255-3217; fax: 607 254-4590; email: yp34@cornell.edu. Min Ding is Associate Professor of Marketing at the Smeal College of Business, Pennsylvania State University, University Park, PA 16802-3007; phone: 814 865-0622; fax: 814 865-3015; email: minding@psu.edu . Vithala R. Rao is the Deane W. Malott Professor of Management and Professor of Marketing and Quantitative Methods at the Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853-6201; phone: 607-255-3987; fax: 607-254-4590; email: vrr2@cornell.edu. All authors have contributed equally and are listed in random order. The authors acknowledge financial support from the Johnson Graduate School of Management, Cornell University and the Smeal College of Business, Pennsylvania State University.

Eliciting Preference for Complex Products:
A Web-Based Upgrading Method

Abstract

We propose a new incentive-aligned approach, upgrading method, for eliciting attribute preferences about complex products, which combines the merits of self-explicated approach and conjoint analysis. The upgrading method first endows a subject with a product profile, and allows her to upgrade it, one attribute at a time, to a more desirable product configuration. During the process, she states her willingness to pay (WTP) for each potential upgrade she is interested in, and BDM procedure (Becker, DeGroot, and Marschak 1964) is used to ensure it is in the best interest of a subject to state truthfully her WTP. Each subject will receive her upgraded product at the end of the study. We implemented this procedure on the Web in an empirical implementation with digital cameras. This procedure is shown to significantly improve predictive performance over the benchmark (self-explicated) approach.

Keywords: Conjoint analysis, Complex products, Incentive alignment, Self-explicated method, Preference measurement

Many new products in the technological era are quite complex, each comes with a massive number of attributes and/or levels within each attribute. Digital cameras, for example, have more than 10 major attributes, with 2 to 10+ levels for each attribute. This poses a tremendous challenge to marketers who are interested in understanding consumer preferences for such complex products. The conventional approach, conjoint analysis, is not suited for such task. Conjoint analysis is a decompositional approach where a researcher tries to infer an individual's preferences for each attribute level based on her stated preferences/choices for selected versions of the product. In order to obtain accurate estimate of individuals' preferences, a minimum number of such versions (called profiles) is usually needed and this number increases quickly as the number of attributes (and/or levels) increases. Conjoint analysis has been widely used in marketing with undeniable success (e.g., Carroll and Green 1995, Cattin and Wittink 1982, Green and Kreiger 1991, 1992, Kohli and Mahajan 1991, Mahajan, Green, and Goldberg 1982). The massive number of attributes and levels associated with complex products, however, would have required a conjoint analysis to employ an unrealistically large number of product profiles and, if such tremendous cognitive burden were forced upon conjoint subjects, would have degraded the quality of data and led to spurious preference structure and managerial insights (Hauser and Rao 2004). As a result, Green and Srinivasan (1990) in their influential review paper listed "methods for coping with large numbers of attributes and levels within attribute" as a key future direction in conjoint analysis. Recently, Bradlow (2005) emphasized the importance of this issue in a wish list for conjoint analysis. A recent major advancement in this domain is Bradlow, Hu, and Ho (2004), which develops a behavioral model for describing and predicting participant's preferences using partial conjoint profiles.

The alternative to decompositional approach, such as conjoint analysis, is compositional approach. In compositional approach, researchers directly ask each individual about their preferences for each attribute (and/or level), and their corresponding preferences for a given product are then obtained by combining their preferences for the included attribute levels. The most widely used compositional approach is the self-explicated approach (Green and Srinivasan 1990, Srinivasan and Park 1997). This approach contains two steps. First, it asks directly each subject to state their desirability of various levels for a given attribute. Next, it asks each subject to allocate a constant sum (e.g., 100 points) across all attributes whereas the numbers correspond to the importance of each attribute. Clearly, the self-explicated approach is much easier to implement compared to conjoint analysis, and imposes less cognitive burden on subjects. The magnitude of such cognitive "savings" increase quickly as the number of attributes and/or levels increases. Thus, self-explicated approach has been widely used to elicit consumer preferences for complex products. To the delight of researchers and practitioners, it is proven to have surprisingly robust predictive performance (Srinivasan and Park 1997). As a result, self-explicated approach has become the standard method in understanding consumer preferences for complex products and is thus used as the benchmark in this paper.[1]

There are, however, arguments that the self-explicated approach could be potentially improved. First, it suffers several limitations inherent to a compositional approach (see Sattler and Hensel-Borner 2000 for a review). The task in compositional approach is not similar to real situation, and it may, among other things, increase the likelihood of double counting when different attribute levels convey similar benefits (Green and Srinivasan 1990), decrease the chance to detect potential nonlinearity in part-worth function (Green and Srinivasan 1990), or

[1] Alternative ways to estimate the part-worths for complex products could be considered. In particular, a conjoint analysis can be examined in which each subject completes a subset of the questions (Lenk et al. 1996). This approach, however, would require a much larger sample in the context examined here.

true weights (Srinivasan 1988), and lead subjects to appear less sensitive to range of levels (Nitzsch and Weber 1993) and more conforming to socially acceptable answers (Green and Srinivasan 1990). The second argument is related to incentive alignment. Subjects are not incentive-aligned during the self-explicated approach (or for that matter, conjoint analysis) and, as recently reported in the literature (Ding, Grewal, and Liechty 2005, Ding 2007), individuals' behavior in non-incentive-aligned environment sometimes systematically deviates from their true preferences. For example, individuals may be more risk seeking and may appear less price sensitive.

As a result, an ideal improvement over self-explicated approach should satisfy several criteria. First, it must preserve the benefits of self-explicated approach while mitigating its limitations. Given that our objective is to understand preferences for a complex product, we could not simply resort to a decompositional approach. Second, it must incentive-align subjects to state truthfully their preferences. Third, it must place minimal cognitive burden on subjects to understand the procedure, while efficiently collecting useful information about preferences. It is important that an improved approach can be easily understood and willingly carried out by subjects.

This paper proposes a web-based upgrading method that satisfies all three criteria described above. It first endows each subject with a bare bone version of the product under study, and then allows her to upgrade the product, one attribute at a time (defined as a *round* in this paper). During each round of upgrading, a subject could seek an improvement that will upgrade one particular attribute to a more desirable level from whatever she owns at that time. The basic experience of upgrading is similar to that purchasing a computer from build-your-own-goods websites (e.g., Dell), where a consumer starts with a generic version and determines which attribute he/she wants to upgrade, and if yes, to which level. In an empirical implementation using digital cameras (with 11 attributes and 60 levels in all), the web-based upgrading method leads to significantly better predictive performance as compared to the self-explicated approach. Furthermore, the upgrading method data revealed a quantitatively different preference structure from the same set of subjects.

The rest of the paper is organized as follows. We first specify the web-based upgrading method design and discuss its properties. We then describe an experiment designed to validate the upgrading method empirically using digital camera, followed by the estimation and results. We conclude with general discussions and directions for future research.

Upgrading Method Design

In this section, we first describe the design in detail, and then discuss its properties with regard to incentive alignment. Finally, we discuss the realism and efficiency of the design.

General Design

First, we define a complete run of upgrading as a *set*. At the beginning of a set, each subject will be endowed with a particular version of the product. She will then attempt to upgrade it to a more desirable product configuration. The upgrading procedure is organized such that a subject can only upgrade one attribute at a time (round), and only have one chance to upgrade any specific attribute. The method is implemented at the individual level, over a web-interface, which allows for dynamic customization of the study based on each subject's responses and outcomes as they evolve. We describe below one possible implementation of upgrading method, which we used in the empirical study reported here. Alternative implementations are possible and are discussed in the last section of this paper.

Specifically, the steps involved in a set of upgrading, as shown in Figure 1, are: (1) A subject accesses the web-based upgrading study through a web browser (e.g., Internet Explorer); (2) She is endowed with a (bare bone) configuration of the product; (3) She is shown all attributes that are available for upgrading (she can only upgrade once for each attribute), and is asked to select the attribute to upgrade next; (4) She is shown all levels in that attribute, and is asked to state her willingness to pay (WTP) to upgrade from her current level to each of the levels she is interested in for that attribute; (5) The computer randomly generates a cutoff price for each level, and determines whether a level is upgradeable (defined as the stated WTP for this level is larger or equal to the randomly drawn cutoff price for the same level); (6) Her product will remain the same if no level is upgradeable, otherwise it will be upgraded to one of the upgradeable levels (randomly chosen by the computer), but only pays the randomly chosen cutoff price for the upgraded level; (7) Steps (3) to (6) are repeated, until she has upgraded all attributes of interest to her, or she decides not to upgrade on remaining attributes.

Insert Figure 1 about here

At the end of the upgrading, a subject will receive her final upgraded configuration of the product, and pay the cumulative cost of the upgrades she has made. Following the standard practice of experimental economics, we recommend endowing each subject with an amount of cash at the start of the study, and the cost of upgrading is then subtracted from this cash endowment (thus a subject does not need to take money out of her own pocket). If the product under study is expensive and endowing every subject is not financially feasible, a lottery mechanism may be used to determine which participant will end up receiving the final product (which is what we employed in the empirical study in this paper). We highlight the difference between self-explicated method and the proposed upgrading method in Table 1. We note that the self-explicated approach directly provides unbiased individual-level estimates because each subject is expected to state the desirability of level for a given attribute and the importance of each attribute. Under the upgrading approach, on the other hand, WTP data are not collected for some attributes/levels for some subjects. Thus, the hierarchical Bayesian estimation is desirable to share the data across subjects in the sample.

Insert Table 1 about here

<u>Incentive Alignment</u>

The proposed upgrading method ensures incentive-alignment through two design features. First, it will give a subject the final upgraded product and require a subject to pay the total cost of upgrades (or equivalently, give the subject an amount of cash that equals to the difference of an initial endowment and total cost of upgrades). This feature thus ensures a participant will treat her decision exactly as real life decision. Second, it utilizes the Becker-DeGroot-Marschak (BDM) procedure (1964), which ensures that it is in the best interest of a subject to state her true WTP (see Wertenbroch and Skiera 2002 for a recent application in marketing). It involves the following steps: (1) A subject states her WTP for an item (in this case, a specific upgrade); (2) A number (cutoff price) is randomly drawn from a (typically, uniform) distribution, and the range of the distribution should include all reasonable WTPs for this item but is otherwise unknown to subjects; and (3) The outcome is determined as follows: If the number drawn is higher than the stated WTP, the subject will not be able to purchase the item, but if the number drawn is lower than or equal to the stated WTP, the subject will be able to purchase the item but pay only the randomly drawn number (price).

Although a participant will state one WTP for each of the levels of an attribute that he/she is interested in upgrading to, it is unrealistic to allow upgrading to more than one level for a given attribute at the same time. To ensure it is incentive aligned for a participant to state simultaneously his/her true WTPs for all levels of an attribute that he/she is interested in upgrading to, we will draw a random cutoff price for *each level* for each participant to determine whether a participant will be allowed to upgrade to this level. Since a participant will receive only one upgrade from each attribute (as in real life), the next step is to determine which level this participant will be given. We discuss below two approaches that can be used to achieve this objective, their strength and weakness, and finally which one we use in this paper.

The first approach is to randomly select a level from all levels in an attribute. If a participant did not bid on this level, or bid on this level but bid was below the cutoff price for this level, this participant will not receive any upgrade. If the level is upgradeable, the participant will then receive this upgrade. While this implementation is solid in theory and will achieve complete incentive alignment, it will lead to substantially a smaller number of successful upgrades during the study. This causes a serious problem in practice as it becomes much less realistic and participants will be disappointed and quickly lose interest in the study.

The second approach is to randomly draw a level from those levels that are upgradeable. This approach ensures a participant will get an upgrade if he/she has at least one upgradeable level. It can also be shown that, as long as the distribution of the cutoff price is designed such that the density is high on the upper end (which is in the control of the researchers/practitioners), it is sub-optimal for a participant to either overstate or understate her WTP during upgrading, and a participant's optimal strategy is to simply state her true WTP. Nevertheless, some participants might still believe shaving their bids (i.e., state a number that is smaller than their WTP) is a better strategy, if they do not carefully consider the consequence of such strategy and its dependency on the distribution of the cutoff price.[2] As a result, it is possible that WTP collected (from some participants on some attributes) are somewhat smaller than their WTP. On balance, we believe the second approach is a better option in practice and have thus adopted it in the empirical test in this paper.

Realism and Efficiency

Realism is important because subjects may not be able to accurately state their preferences in an unfamiliar task even though they are incentive-aligned to do so. The upgrading method in general (and the proposed design in particular) mirrors the real task individuals engage in choosing a product in the marketplace; the individuals are familiar with this approach. For example, an individual visiting a major computer vendor's website will first be shown some default configuration, and she is then allowed to customize this computer one attribute a time, with additional money for each upgrade, until the most desirable configuration (given the cost of each upgrade) is identified.

Another critical design issue is efficiency in terms of extracting as much useful information as possible without imposing tremendous cognitive burden on subjects. The upgrading method described contains three features that specifically address this efficiency issue. First, we will endow each subject with the bare bone (on average, least desirable configuration) version. As a result, subjects are interested in participating in several upgrading rounds, and within each round, most levels will be more desirable than the one they have at that time. As a result, researchers and practitioners are able to extract abundant WTP information from each subject. Second, a subject would state her WTPs for all levels she is interested in upgrading during each round, and the computer randomly picks one upgradeable level for each attribute. This allows researchers/practitioners to obtain, simultaneously, a subject's preferences on all

[2] We thank guest editor for pointing this out.

attribute levels she is interested in, instead of only the one she will be upgraded to. Finally, the upgrading method enables each subject to complete more than one set of upgrading. This is possible because a specific upgrade a subject receives is determined by two random events (randomly drawn cutoff price and the random selection if there is more than one upgradeable level). It is also possible because we can allow them to start with a different initial configuration in a set.

Experiment

In order to validate the upgrading method empirically, we conducted a within-subject contrast experiment between the upgrading method and the benchmark self-explicated approach. We used the following criteria to select the product for experiment: (1) The potential study subjects (university students) must be a key target segment of the product; (2) The potential study subjects must be reasonably familiar with the product category and be interested in purchasing such product at the right price; and (3) The product must be complex, characterized by a massive number of attributes and/or levels, and reasonably expensive.

The first two criteria are essential for any realistic study, while the third criterion is specific for the particular problem we are addressing (complex product). We interviewed a small sample of potential subjects (none of them later participated in the study) about what types of products, in the price range of a few hundred dollars, that they would really like to have. We choose this price range as such goods considered by university students tend to be complex products with a large number of attributes/levels. The top categories revealed by the subjects included digital camera, MP3 player, and cellular phone. To ensure subjects will actually be interested in purchasing such products if the price is right, we then polled how many of them own products in each of the categories and their purchase intention. Digital camera was the product category with the lowest ownership but highest purchase intention (Everyone wants a chance to purchase a nice digital camera; even those who currently own one indicated they could use a better digital camera). We then conducted research using such sources as the Bestbuy.com and other similar websites and found that these sites used more than 10 major attributes in describing digital cameras; thus this product also satisfied our third criterion (complex product). As a result, digital camera is chosen for the experiment.

Experimental Design

Using the product descriptions and comparison features at Bestbuy.com as the benchmark, we selected a total of 11 most important attributes that have been commonly used to describe and compare different types of digital cameras. We also identified the levels that are most common within each attribute. This has led to a total of 60 attribute levels across 11 attributes: 10 levels for optical zoom, 9 levels for brand, 7 levels for Weight, 5 levels for Resolution, Warranty-Parts, Warranty-Labor, Focus Range, Viewfinder Size, 3 levels for Text Overlay, Video, and Flash Range (see Table 2 for the detailed description of the attributes and levels).

Insert Table 2 about here

Each subject in the experiment completed four tasks as shown in Figure 2; these are self-explicated task, upgrading task, external validity task, and a brief survey on oneself and the experiment. The general instructions contain directions for the experiment, a description of the attributes and levels to be used in the experiment and a glossary explaining the terms used to describe the attributes and levels. Each subject received $7 for participating in the study. Further, we randomly selected one subject out of every 40-50 subjects to receive a digital camera (the specific configuration of camera was determined by her choice/outcome in the upgrading method and external validity task) and some cash (the difference between $400[3] and the price of the digital camera received).

Insert Figure 2 about here

The self-explicated task was designed following the standard format in the literature (Green and Srinivasan 1990) and contained two sections. In the first section a subject was asked to evaluate, one attribute at a time, how desirable each level within this attribute was by assigning a number between 1 to 10 to each level, with the most preferred level as 10 and the least preferred level as 1. In the second section, a subject evaluated the importance of attributes by allocating a total of 100 points to all attributes based on his preference.

The instruction for upgrading task closely followed the theoretical design (previous section) with three variations. First, we randomly picked one out of 40-50 subjects to reward a digital camera she selected/upgraded, plus the cash balance (the difference between $400 and camera's cost computed by the upgrading procedure). Second, we charged $100 (out of $400 endowed to a subject) for the base model (starting product). Third, we asked each participant to complete two mandatory sets of upgrading, with the option to do as many additional sets of upgrading as they wanted to. Information on the additional sets completed also gave us a sense of how involved a participant is during upgrading task.

The external validity task had two choice questions. In each choice question, a participant evaluated 17 different digital cameras, and decided which camera she would like to buy. To make the choice task more natural, we also included the option of not buying any of the cameras. A subject who selects a no purchase option will be excluded for the validation purpose, as self-explicated approach does not predict this outcome. The profiles of digital cameras in the first

[3] Based on informal interviews of potential participants, it appeared that nobody at the university where our research was conducted was willing to pay $400 or more for any digital camera. $400 is chosen also because most digital cameras that fit our product space are below this threshold.

question were generated using the SAS experiment design macros to ensure objectivity, and after eliminating clearly dominant profiles. The profiles of digital cameras in the second question were generated from real digital cameras from Bestbuy.com.[4] Finally, subjects completed a brief survey on their general experience with cameras as well as feedbacks on the study.

The web-implementation of the upgrading method was written using PHP (a programming language mainly used for server side software applications) and MySQL database, on a server hosted by Yahoo. With this implementation, the upgrading method (study) can be conducted on any computer that has internet connections. Both PHP and MySQL are free and the specific codes are available from the authors upon request.

Experimental Procedure

A total of 88 subjects at a major U.S. university participated in the study in a campus computer lab. Subjects were randomly assigned to 4 conditions. Subjects in conditions 1 and 2 completed self-explicated task first, followed by upgrading task; subjects in conditions 3 and 4 completed upgrading task first, followed by self-explicated task. All subjects then completed external validity task, and finally responded to the brief survey. We constructed two configurations as the starting products during the upgrading task, each contains either the worst or second worst levels in each attribute (HP and Kodak for brands). Subjects in conditions 1 and 3 started with the first initial configuration, and subjects in conditions 2 and 4 started with the second initial configuration. The actual study took about 30 minutes to complete. Two subjects were randomly selected at the conclusion of the study and were compensated as described in the study (camera plus cash balance).

[4] Based on informal interviews, students normally evaluate about 20 cameras initially (although quickly narrow down) when they buy digital cameras. We thus used SAS macro to generate 20 holdout profiles. Out of the 20, 3 are either dominated or dominant and are thus eliminated, leaving us with 17 digital cameras. For the second holdout task (Bestbuy cameras), we decided to use 17 cameras again to be consistent with the first holdout task.

Analysis

From the self-explicated approach, we obtained attribute importance evaluations and attribute-level desirability from each subject. We multiplied these two quantities to obtain self-stated importance of all attribute levels and computed the utility of items in the validation tasks for each of the 88 subjects for the self-explicated approach.

From the upgrading method, we obtained the attribute levels a subject had submitted an offer to, and the actual amount of the WTPs. We considered two types of interpretations of the upgrading data and estimated the data separately. The first interpretation is based on the indifference judgment that the utilities are equal between the starting configuration (plus the amount of offers made) and the targeted configuration.[5] The second interpretation assumes some participants (erroneously) adopted "shaving bids" strategy and the stated amount of offers are smaller than their true WTPs. In this case, a paired inequality comparison is suited.

A subject completed an average of 2.67 sets of upgrading (std. dev. = 0.77, max = 5.00); this indicates some interest (2.67 is significantly higher (p-value < 0.01 on the mean comparison t-test) than the two that were mandatory) on the part of subjects in the upgrading approach. On average, there are 29 and 25 offers submitted per respondent in sets 1 and 2 of the upgrading method, respectively; the average amount of offers made across all 11 attributes is $33.09 (std. dev. = $30.25) and 3.02 offers for upgrading in a given attribute (std. dev. = 2.07).[6] Table 3 shows the number of offers made per attribute across respondents. The average time to complete

[5] We would like to thank an anonymous reviewer for this interpretation of the data from the upgrading method.
[6] We take the only attributes/levels for which the respondent stated offers in both sets and compute the correlation coefficient for each respondent. The average correlation coefficient across respondent is 0.64 with the standard deviation of 0.33, indicating a relatively high reliability of the procedure. Note we are not expecting perfect correlation as the amount of offer made for a given level also depends on the current profile they have (at the time of each attribute upgrade), and these profiles are different between the two sets.

one set of the upgrading method was 332 seconds, compared to 426 seconds for completing the self-explicated method.

Insert Table 3 about here

Estimation Procedure

We implemented two models to test the predictive power of our proposed approach. As a benchmark model, Model 1 uses the data from the self-explicated method. Model 2 uses the data from the upgrading method. To provide the best possible comparison between the self-explicated method and the upgrading method, we assess both individual subjects' preferences and out-of-sample predictions. Individual subjects' part-worths under the self-explicated approach is the product of attribute importance and level importance, which are in turn used to construct utility for each of the 17 cameras in the external validity tasks for the out-of-sample predictions.

To analyze the data from the upgrading method, we used a random-effects hierarchical Bayesian logit model, similar to the model specified by Allenby, Arora, and Ginter (1998). The utility of price is assumed to be linear. Under the second interpretation noted above, the probability that the *i*-th subject chooses the *k*-th alternative (the profile after potential upgrade) from the *j*-th pair (including both profiles before and after potential upgrade, plus the amount of offer made for the profile before upgrading) is given by[7]

$$p_{ij}^{k} = \frac{\exp\left(\beta_i^T x_{ij}^k\right)}{\sum_l \exp\left(\beta_i^T x_{ij}^l\right)},$$

[7] If an attribute has 4 levels (e.g., A, B, C, and D), a participant started with level A, and is interested in upgrading to C or D (but not B), she will state one offer (say, $10) for upgrading to level C and one offer (say, $15) for upgrading to level D. In this example, we infer two paired comparisons. That is, her utility for the product profile before upgrading (with level A on the attribute) plus $10 is less than her utility for the upgraded product (with level C on the attribute). Similarly, her utility for the product profile before upgrading (with level A on the attribute) plus $15 is less than her utility for the upgraded product (with level D on the attribute).

where x_{ij}^{k} (including the stated amount of offer if applicable) describes the k-th camera by the i-th subject from the j-th pair, and β_i is a vector of the part-worths for the i-th subject. We assume a hierarchical shrinkage specification for the individual part-worths, where a priori,

$$\beta_i \sim N\left(\bar{\beta}, \Lambda\right).$$

This specification allows for estimation of part-worths β_i at the individual-level, the aggregate or average part-worth $\bar{\beta}$, as well as of the amount of heterogeneity (Λ).[8] We assume average amount of offers made for each attribute level with diffuse priors to ensure proper posteriors but also allow the data to primarily govern the inferences. We assess the convergence properties of the Markov Chain Monte Carlo analysis to ensure that the algorithm converged properly.

Results

We estimated the model using the two different interpretations of the upgrading data (equality and inequality). In order to estimate the data with indifference judgment from the upgrading method, we generated two paired comparisons from one indifference judgment (equality): the utility of the first alternative is greater or less than (and equal to) that of the other alternative. The predictive performance based on the inequality interpretation is much better than that based on the equality interpretation (see later in this subsection for details), we thus only present the results from the inequality interpretation in this paper. There are no differences across the four experimental conditions and the data were thus pooled and results reported here are across all conditions. For upgrading data, we estimated the part-worths using data from the first set of upgrading and all upgrading data. As results obtained from using all upgrading data do not show noticeable improvement in out-of-sample prediction, we report the results from the first set of

[8] In line with the literature (Ding et al. 2005), we did not find substantial difference in predictive performance between the two models (with a diagonal and non-diagonal matrix). We report the results with a diagonal matrix, which leads to slightly better predictive performance.

upgrading in this paper. We include the mean (and standard deviation) of self-stated part-worths from the self-explicated method and part-worth estimates from the upgrading method in Table 4.[9]

Insert Table 4 about here

It should be noted that no direct comparison between the two sets of part-worths can be made. While the upgrading method provides a measure of preferences that can be interpreted in dollar value (as the choice conjoint analysis), the self-explicated approach only provides a relative measurement. Nevertheless, we could compare the two sets of part-worths on the preference order of levels within an attribute.

For most attributes with uniform directional benefit (either more is better or less is better), the preference order of the part-worths across the levels within an attribute is remarkably similar between the two methods. Upgrading method, however, appears to recover nonlinear preference, compared to the self-explicated method. This is consistent with Green and Srinivasan (1990)'s observation that participants in self-explicated approach tend to assign linear preference to different levels in an attribute, if these levels are linear.

We briefly discuss here two attributes with uniform directional benefits, warranty on parts and weight (Figure 3 and Figure 4). While the preference order of the levels for warranty-parts is identical between the two methods, a noticeable difference is observed between 12 months and 18 months. Under the self-explicated approach, the self-stated importance continuously (almost linearly) increases as the warranty for the parts increases. Under the upgrading method, however, the posterior mean remains almost constant between 12 and 18 months. Similarly, the preference is relatively constant between 24 and 36 months. For the

[9] Note that the self-explicated approach allows nonlinearity in price. We thus investigated both linear and quadratic specification for price to capture nonlinearity in price of the upgrading data, and found the quadratic term to be insignificant. We only present the results with the linear specification of price.

attribute of weight, the nonlinearity is observed for the levels from 5-6oz to 7-8oz. Under the self-explicated method, the self-stated importance decreases almost linearly as the weight increases. Under the upgrading method, however, the preference remains almost constant between 5-6oz and 7-8oz for the upgrading method.[10]

For the attribute with no *a priori* uniform preference order across levels, i.e., brand (Figure 5), it is worth noting that for the attribute of brand, the rank order of the levels is remarkably similar between the self-explicated method and the upgrading method and almost identical for the more preferred levels.

Insert Figures 3, 4, and 5 about here

Finally, we examine the predictive performance for two holdout tasks (one with computer-generated profiles and the other with profiles from Bestbuy.com) from both approaches. As Green and Srinivasan (1990) noted, out-of-sample prediction provides true validation for conjoint methodology and therefore, should serve as the best yardstick to judge whether the proposed upgrading method adds value to the self-explicated method. Since there are normally a large number of competing alternatives for complex products in real life (e.g., there were close to 100 different digital cameras on Bestbuy.com website), from which a consumer has to make her purchasing decision, we implemented holdout tasks with 17 different digital cameras. Compared to the more traditional holdout task (where an individual selects from three or four different product profiles), our holdout task will provide a more realistic test of the validity of the upgrading method. We provide the out-of-sample predictions in Figure 6 for both

[10] In the previous versions of the manuscript, we included non-bid information in the analysis. In particular, for attribute levels where a subject submitted no offer (level B in footnote 7), a pairwise comparison was utilized if the focal attribute's benefit is directional (e.g., higher resolution is always better than the lower resolution). For attribute levels with no uniform preference orderings (e.g., brand), no relative preference was inferred when there are no offers submitted. In our empirical application, the predictive performance with bid-only data was better than that with data from non-bid as well as bid. This might be due to the fact that all attributes except brand have uniform directional benefits. Furthermore, the above interpretation for non-bid might require unnecessarily strong structural assumptions on consumer preferences.

holdout tasks (Four subjects selected no purchase option in the first validation task, and are thus not included in calculating predictive performance regarding the first validation task.). The upgrading method leads to significantly better predictive performance: the percent of matches between the actual choice and the top predicted option are 42% for the first holdout task (computer-generated profiles) and 26% for the second holdout task (Bestbuy.com profiles) under the upgrading method, versus 27% and 19% under the self-explicated method, respectively.[11] Note the baseline prediction in a naïve model is 6% (i.e., random select 1 of 17 choices). This result provides strong empirical evidence for the validity and managerial usefulness of the proposed upgrading method in understanding preferences for complex products.

Insert Figure 6 about here

We found that the subjects were able to understand the upgrading method quickly, and responded well to the incentives built into the approach. For example, when asked at the end of the experiment whether "they understood it is in their best interest to state exactly the price they were willing to pay during the upgrading method", the average responses from the subjects is 4.06 (on a scale from 1 to 5, with 1 being not clear at all and 5 being extremely clear). Subjects also indicated that the upgrading method is more stimulating than the self-explicated method (p-value = 0.00 on the paired-t test) and that subjects would be happier to do the upgrading task again in the future (p-value < 0.01 on the paired-t test).

Discussion and Future Research

Motivated by the need to better understand consumer preferences for complex products, this paper specifies a new upgrading method, which builds upon the benchmark self-explicated method while ensuring the merits of both decompositional and compositional approaches are

[11] Under the first interpretation of the upgrading data (equality for indifference), the predictive performance is 36% and 19% for the first and second holdout tasks, respectively.

incorporated, and subjects are incentive-aligned to reveal their true preferences. In addition to these desirable features, upgrading method is also built upon a realistic task that most people are familiar with (e.g., upgrading from a default computer, one attribute at a time, to a configuration that she likes most, given the cost of each upgrade). As a result, subjects do not need to spend major cognitive effort to understand and participate in upgrading method. Furthermore, the upgrading approach is very efficient in obtaining useful information and is comparable to the benchmark self-explicated approach. In the empirical study we conducted, the upgrading method took less time to complete (per set) than the self-explicated approach. Moreover, subjects appear to enjoy the upgrading method more than the self-explicated approach and are much more involved. It should be noted that any improvements in performance of the upgrading method could come from a combination of three sources: upgrading procedure, incentive alignment, and information sharing across individuals in hierarchical Bayes estimation; but we can not ascertain their relative contributions. It will be helpful if future research can parse out the unique contribution of each source, possibly by designing an incentive aligned mechanism for self-explicated condition, which has never been done before and yet to be developed by future research.[12]

The empirical test conducted using digital camera, a complex product with 11 attributes (with a total of 60 levels), shows a quantitatively different preference structure under upgrading method than that revealed by the self-explicated task. Most importantly, the preference structure uncovered from upgrading method data has superior external validity than that uncovered from the self-explicated method. The empirical results also reveal that the inequality interpretation of

[12] Since a participant does two or more sets of upgrading, and only one set of self-explicated task, those participants (one-half) who first performed the upgrading method may be biased towards the upgrading method (compared to those who does self-explicated first) that could lead to potentially higher validation results. We thank an anonymous reviewer for pointing this out.

the data has better external validity than the equality (indifference) interpretation, indicating that at least some of the participants stated offers that were below their true WTP and had not behaved in a way that is consistent with that under incentive compatible mechanism. As discussed previously, this could be due to the fact that these participants erroneously believed that shaving bids is a better strategy than truth-telling.

While the upgrading method is conceptually sound and the empirical evidence is convincing, there remain several promising areas for future research. First, it is not easy to determine the base level for an attribute without a prior preference order (e.g., brand, color). It will be helpful to practitioners if rules of thumbs can be identified. Second, the proposed upgrading method is designed to closely resemble (thus preserve the benefits of) the benchmark self-explicated method. It is possible, however, to use other types of upgrading formats. Instead of upgrading one attribute at a time, a subject can upgrade a combination of multiple attributes (e.g., upgrade to 6 mega pixel with 5× optical lens) by developing a super attribute, where each level in this super attribute is a combination of levels of closely inter-related attributes. The *n*-th price auction, instead of individual-based BDM procedure, can also be used. It may, however, impose design and implementation restrictions as it requires multiple subjects upgrading the same attribute levels at the same time, and need to address potential collusion among subjects.

Third, while we have kept the model parsimonious to highlight the key aspects of the upgrading method, the sequence of attributes for upgrading could be considered in the model estimation because subjects tend to upgrade more important attributes first. Fourth, there is an (unstated) assumption in the empirical analysis that the paired comparison judgments are independent. This assumption might not hold, especially for multiple levels of the same attribute for which the participant provided the offers. It will be interesting to investigate in the future how important this assumption is for upgrading method.

Fifth, following the convention in experimental economics, the upgrading method endows subjects with a sum of money to be used for upgrading. The effects of this endowment need further examination because such endowment might artificially change price sensitivity among the participants. The empirical implementation with digital camera also used the lottery incentive structure (with 1 in 44 chance of winning); it would be interesting to see the relative effectiveness of the lottery incentive compared to an incentive structure that rewards every participant.

Sixth, while self-explicated method is the appropriate benchmark model for the proposed upgrading method, alternative approaches exist both in academic (e.g., Lenk et al. (1996) gave a fraction of conjoint questions to each participant) and practice (e.g., SIMALTO; www.simalto.com). Although the literature has not reported whether these alternative approaches lead to better predictive performance than self-explicated approach, it will still be interesting to contrast these approaches with the upgrading method in the future.

Seventh, we have proposed and tested a design that a participant starts with a barebone version and gradually improve it (upgrade on various attributes). Under our proposed framework, it is also possible to trade up and/or down from the product which may lie in the middle level or to trade down from the top product profile that a participant is endowed with. We have restricted ourselves in this implementation to trade up because it is a natural task that an ordinary participant does regularly. But, it will certainly be interesting to investigate trading up and/or down in future research.

Eighth, there is a literature on build-your-own products that we did not discuss early. This literature does not involve incentive-aligned bidding process and could not be used for our intended purposes. Nevertheless, it will be interesting to explore the relationship between that literature the upgrading method in the future.

Finally, it is possible that our upgrading method results reported here for digital cameras could be reflective of the product category. We hope our proposed method provides a framework for further empirical exploration in other product categories.

In conclusion, the proposed upgrading method has appealing features, and is easy to implement. The empirical study conducted demonstrates its superior predictive performance compared to the benchmark (self-explicated) model. We hope this method will provide another important tool in the quest to better understand consumer preferences.

References

Allenby, Greg M., Neeraj Arora, and James L. Ginter (1998), "On the Heterogeneity of Demand," *Journal of Marketing Research*, 35 (3), 384-389.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak (1964), "Measuring Utility by a Single-Response Sequential Method," *Behavioral Science*, 9 (July), 226-232.

Bradlow, Eric T. (2005), "Current Issues and a Wish-List for Conjoint Analysis," *Applied Stochastic Models in Business and Industry*, 21 (4-5) 319-323.

Bradlow, Eric T, Ye Hu, and Teck-Hua Ho (2005), A Bayesian Learning Model for Imputing Missing Attributes in Partial Conjoint Profiles. *Journal of Marketing Research*, 41 (4), 369-381.

Carroll, J. Douglas and Paul E. Green (1995), "Psychometric Methods in Marketing Research: Part I, Conjoint Analysis," *Journal of Marketing Research*, 32 (4), 385-391.

Cattin, Philippe and Dick R. Wittink (1982), "Commercial Use of Conjoint Analysis: A Survey," *Journal of Marketing*, 46 (3), 44-53.

Lenk, Peter J., Wayne S. DeSarbo, Paul E. Green, and Martin R. Young (1996), "Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs," *Marketing Science*, 15 (2), 173-191.

Ding, Min, Rajdeep Grewal, and John Liechty (2005), "Incentive Aligned Conjoint Analysis," *Journal of Marketing Research*, 42 (1), 67-82.

Ding, Min (2007), "An Incentive Aligned Mechanism for Conjoint Analysis," *Journal of Marketing Research*, 44(2), 214-223.

Green, Paul E. and Abba M. Krieger (1991), "Segmenting Markets with Conjoint Analysis," *Journal of Marketing*, 55 (4), 20-31.

Green, Paul E. and Abba M. Krieger (1992), "An Application of Product Positioning Model to Pharmaceutical Products," *Marketing Science*, 11 (2), 117-132.

Green, Paul E. and V. Srinivasan (1990), "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing*, 54 (4), 3-19.

Hauser, John R. and Vithala R. Rao (2004), "Conjoint Analysis, Related Modeling, and Applications" in *Advances in Marketing Research: Progress and Prospects*, Jerry Wind, ed., Kluwer Academic Publishers.

Kohli, Rajiv and Vijay Mahajan (1991), "A Reservation-Price Model for Optimal Pricing of Multiattribute Products in Conjoint Analysis," *Journal of Marketing Research*, 28 (3), 347-354.

Mahajan, Vijay, Paul E. Green, and Stephen M. Goldberg (1982), "A Conjoint Model for Measuring Self- and Cross-Price/Demand Relationships," *Journal of Marketing Research*, 19 (3), 334-342.

Nitzsch, Rudiger von and Martin Weber (1993), "The Effect of Attribute Ranges on Weights in Multiattribute Utility Measurements," *Management Science*, 39 (8), 937-943.

Sattler, Henrik and Susanne Hensel-Borner (2000), "A Comparison of Conjoint Measurement with Self-Explicated Approaches," in *Conjoint Measurement: Methods and Application*, Anders Gustafsson, Andreas Herrmann and Frank Huber ed., Springer Verlag.

Srinivasan, V. (1988), "A Conjunctive-Compensatory Approach to the Self-Explication of Multiattributed Preferences," *Decision Sciences*, 19 (2), 295-305.

Srinivasan, V. and Chan Su Park (1997), "Surprising Robustness of Self-Explicated Approach to Customer Preference Structure Measurement," *Journal of Marketing Research*, 34 (2), 286-291.

Wertenbroch, Klaus and Bernd Skiera (2002), "Measuring Consumers' Willingness to Pay at the Point of Purchase," *Journal of Marketing Research,* 39 (2), 228-241.

Table 1: Self-Explicated Method versus Upgrading Method

	Self-Explicated	Upgrading
Incentive-aligned	No	Yes
Efficiency (compared to Conjoint)	Yes, subjects evaluate all attributes/levels	Yes, subjects only evaluate attributes/levels they are interested in
Task realism	Less natural	More natural (people upgrade in real life)
Measurement of WTP	No*	Yes
Estimation	Unbiased direct estimates	Indirect estimates using hierarchical Bayes

* While not used in academia nor in practice, WTP could technically be estimated from self-explicated data if the price is treated as linear and only the difference between the stated desirability is considered.

Table 2: Attributes and Levels

	Attribute	Total levels	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8	Level 9	Level 10
1	Brand	9	Canon	Fuji	HP	Kodak	Minolta	Nikon	Olympus	Samsung	Sony	--
2	Resolution	5	4 mp	5 mp	6 mp	7 mp	8 mp	--	--	--	--	--
3	Optical Zoom	10	1×	2×	3×	4×	5×	6×	7×	8×	9×	10×
4	Warranty–Parts	5	6 months	12 months	18 months	24 months	36 months	--	--	--	--	--
5	Warranty–Labor	5	6 months	12 months	18 months	24 months	36 months	--	--	--	--	--
6	Focus Range	5	1-6” to inf	6-12” to inf	13-18” to inf	19-24” to inf	24-36” to inf	--	--	--	--	--
7	Viewfinder Size	5	1.0”	1.5”	2.0”	2.5”	3.0”	--	--	--	--	--
8	Text Overlay	3	No	Date	Date & time	--	--	--	--	--	--	--
9	Video	3	No	Video only	Video & audio	--	--	--	--	--	--	--
10	Weight	7	3-4 oz	4-5 oz	5-6 oz	6-7 oz	7-8 oz	8-9 oz	9-10 oz	--	--	--
11	Flash Range	3	< 8’	8’–12’	12’–18’	--	--	--	--	--	--	--

Table 3: Number of Offers Made per Attribute in Upgrading Method

	Total Levels	Mean (Std. Dev.)	% of participants who bid on at least one level of the attribute
Brand	9	4.14 (2.38)	92
Resolution	5	2.87 (1.06)	95
Optical Zoom	10	6.03 (2.94)	98
Warranty – Parts	5	2.90 (1.15)	68
Warranty – Labor	5	2.83 (1.16)	65
Focus Range	5	2.73 (1.14)	95
Viewfinder Size	5	2.76 (1.08)	91
Text Overlay	3	1.40 (0.49)	77
Video	3	1.34 (0.47)	89
Weight	7	3.93 (1.89)	89
Flash Range	3	1.44 (0.50)	92

Table 4: Model Parameter Estimates

Variables	Model 1: Self-Explicated Method		Model 2: Upgrading Method		
	Mean	Std. Dev.	Posterior Mean	95% Posterior Interval	Heterogeneity
<u>Brand</u>					
Canon	83.71	63.81	46.35	[45.42, 49.53]	1.00
Fuji	52.87	46.53	28.79	[25.79, 31.33]	0.84
HP: Base	45.01	41.73	0.00	--	--
Kodak	60.13	50.94	25.14	[18.41, 31.82]	2.68
Minolta	35.73	40.00	25.87	[23.65, 28.11]	0.75
Nikon	69.92	52.50	34.95	[32.68, 37.21]	1.12
Olympus	67.02	53.15	38.89	[34.64, 43.35]	1.11
Samsung	58.13	45.52	29.77	[27.29, 32.25]	1.09
Sony	88.72	61.12	54.23	[49.30, 59.26]	1.21
<u>Resolution</u>					
4mp: Base	18.14	25.59	0.00	--	--
5mp	60.49	49.19	14.75	[8.27, 16.35]	4.79
6mp	95.13	70.42	22.47	[19.99, 27.14]	1.40
7mp	117.52	87.20	38.69	[35.82, 41.72]	0.96
8mp	143.19	102.70	51.53	[48.79, 52.81]	0.71
<u>Lens (Optical)</u>					
1×: Base	12.94	6.83	0.00	--	--
2×	22.95	13.18	1.09	[0.20, 1.93]	0.79
3×	35.14	23.11	16.88	[14.00, 18.89]	0.99
4×	52.33	33.46	26.77	[24.35, 29.08]	1.17
5×	64.61	41.08	28.45	[25.66, 31.46]	1.86
6×	79.67	50.58	35.40	[32.45, 38.55]	1.29
7×	90.77	52.36	40.66	[38.32, 41.57]	0.81
8×	104.24	66.57	43.28	[39.82, 47.05]	0.93
9×	113.24	68.93	56.11	[53.31, 58.98]	1.63
10×	127.23	70.11	61.11	[55.15, 67.48]	0.99
<u>Warranty – Parts</u>					
6 Months: Base	5.64	5.16	0.00	--	--
12 Months	20.82	22.12	13.51	[11.58, 19.37]	1.29
18 Months	29.55	28.43	16.02	[13.59, 20.73]	0.97
24 Months	38.07	31.10	37.10	[33.70, 40.07]	1.43
36 Months	45.33	32.91	39.02	[36.41, 41.68]	1.10
<u>Warranty – Labor</u>					
6 Months: Base	5.71	6.67	0.00	--	--
12 Months	19.90	22.05	14.37	[10.59, 19.00]	3.06
18 Months	27.27	26.71	17.03	[15.00, 19.54]	0.90
24 Months	35.76	30.59	28.67	[22.43, 34.65]	1.16
36 Months	44.17	32.93	33.45	[31.02, 35.54]	0.79
<u>Focus Range</u>					
1-6” to inf	81.75	64.73	46.77	[40.67, 50.63]	1.31
6-12” to inf	64.93	61.52	33.14	[31.46, 35.13]	1.15
13-18” to inf	47.14	38.81	23.69	[16.64, 25.88]	1.04
19-24” to inf	31.73	25.27	10.85	[9.15, 11.68]	0.88
24-36 to inf: Base	17.75	27.27	0.00	--	--
<u>Viewfinder</u>					
1.0”: Base	8.81	10.13	0.00	--	--
1.5”	22.69	19.46	13.70	[13.45, 14.08]	0.52

2.0”	46.20	54.30	19.76	[13.41, 22.96]	0.96
2.5”	69.31	81.94	33.18	[28.81, 37.05]	1.14
3.0”	83.75	99.86	42.48	[40.27, 47.11]	1.03
Text Overlay					
No: Base	5.41	9.59	0.00	--	--
Date	18.17	19.94	17.15	[10.88, 22.44]	2.31
Date/Time	29.37	27.43	22.01	[19.11, 24.83]	1.00
Video Mode					
No: Base	7.80	10.97	0.00	--	--
Video Only	35.40	55.87	15.77	[11.90, 23.27]	2.37
Video and Audio	73.63	110.80	43.63	[41.17, 45.13]	0.78
Weight					
3-4oz	75.78	51.03	57.86	[52.14, 60.09]	1.08
4-5oz	66.82	48.49	40.39	[37.50, 42.35]	1.42
5-6oz	51.13	37.78	25.88	[22.18, 28.81]	1.08
6-7oz	40.75	30.78	22.88	[20.22, 24.18]	1.18
7-8oz	26.98	21.75	18.91	[16.40, 20.81]	0.81
8-9oz	17.87	17.99	7.50	[6.38, 9.27]	0.97
9-10oz: Base	9.52	8.98	0.00	--	--
Flash Range					
< 8’: Base	9.37	11.69	0.00	--	--
8’-12’	38.96	33.43	19.77	[13.44, 24.91]	4.26
12’-18’	66.46	61.03	39.82	[35.12, 44.00]	1.04
Price					
$149	121.52	154.57			
$189	90.91	105.85			
$229	76.56	90.84			
$269	60.73	72.24			
$309	37.78	45.83			
$349	23.96	31.05			
$389	13.00	15.36			
Amount of offers made ($)			-1.82	[-3.68, -0.12]	1.90

Figure 1: Upgrading Method (Flowchart for One Set of Upgrading)

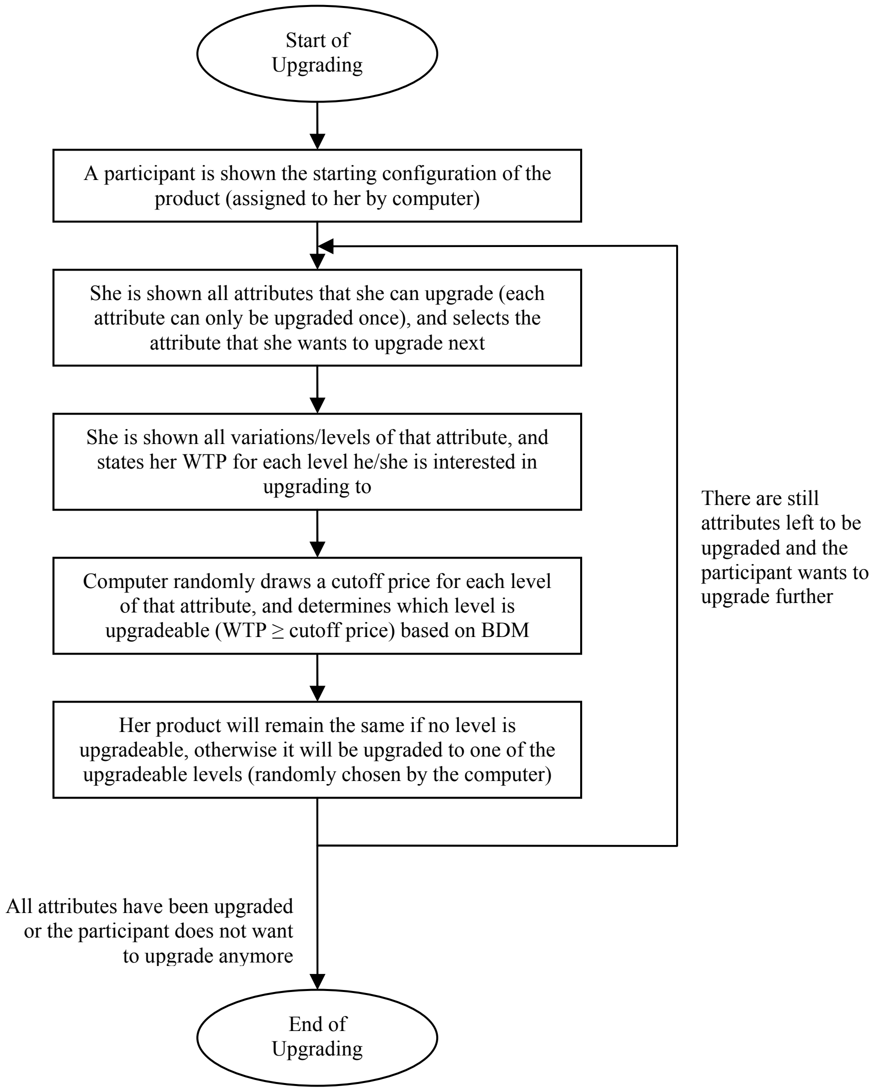

Figure 2: Experimental Design

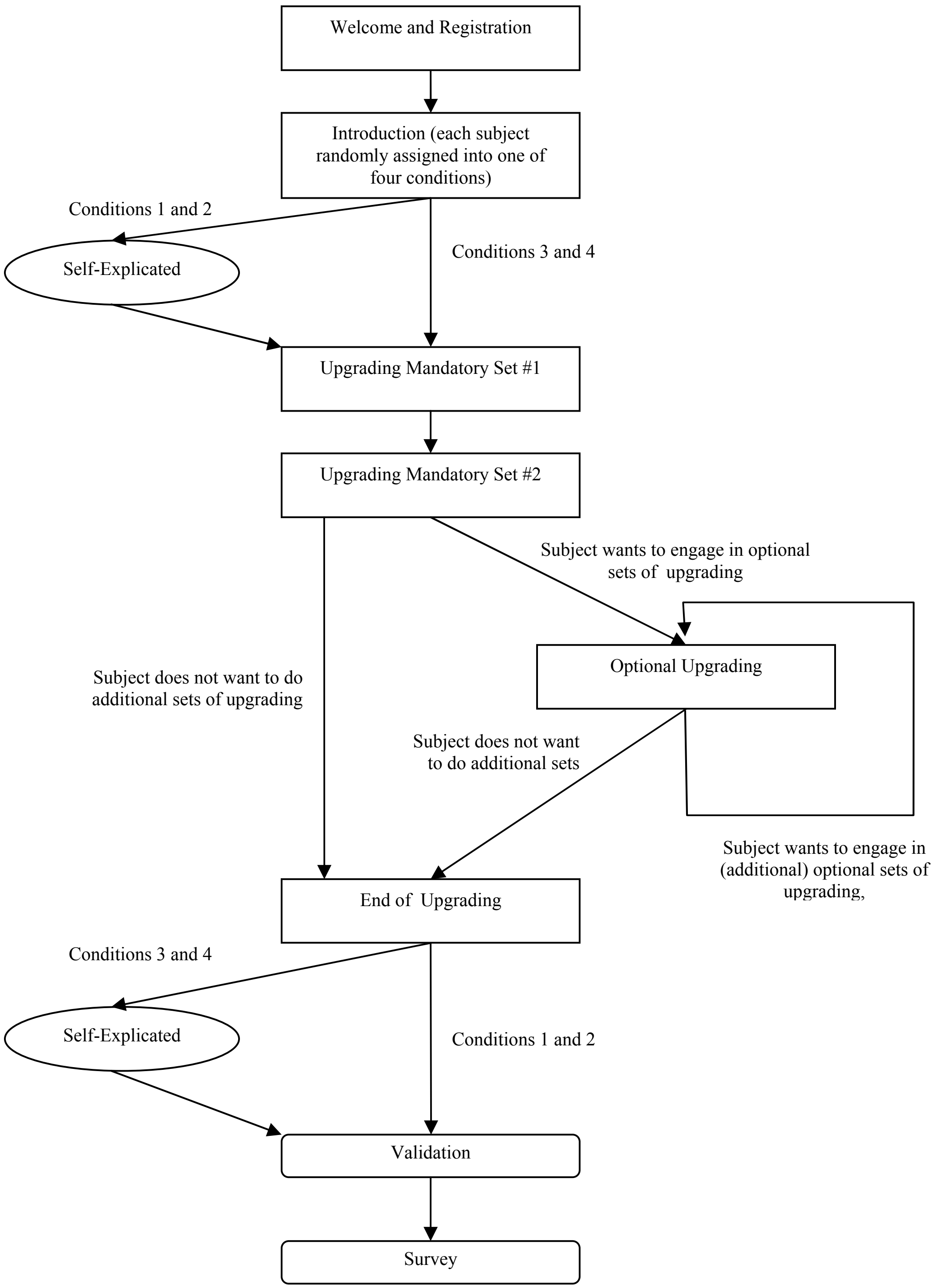

Figure 3: Model Results for Warranty-Parts

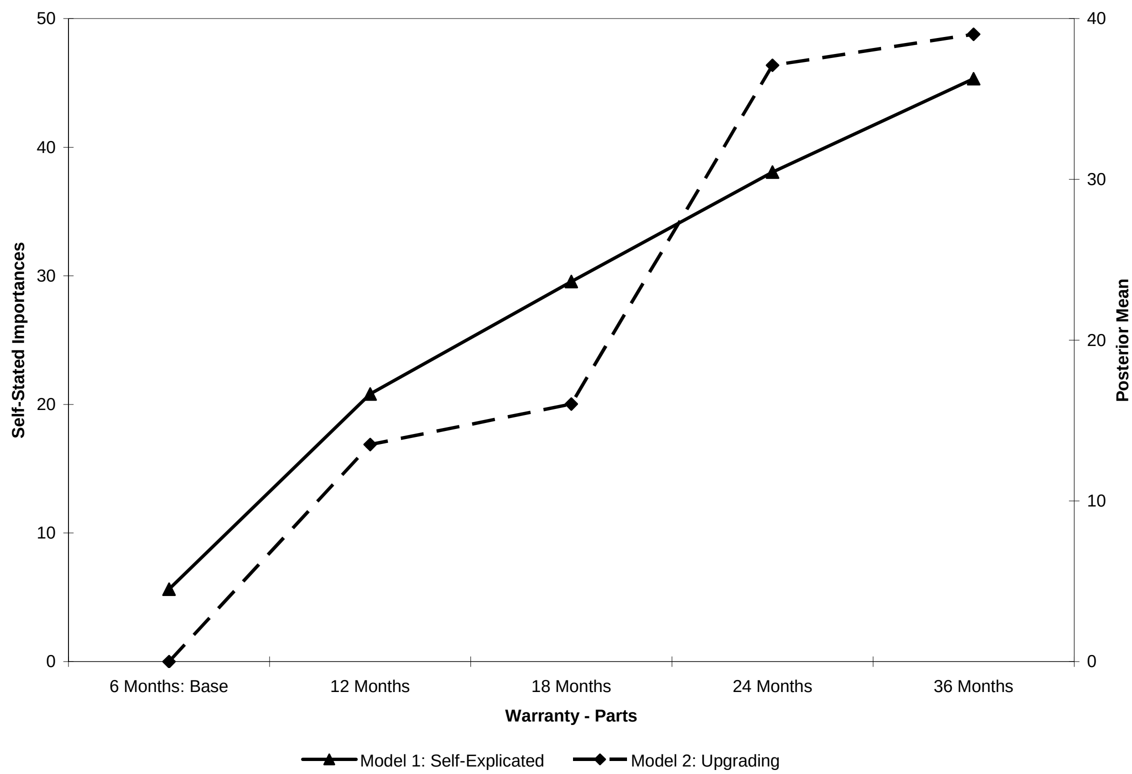

Figure 4: Model Results for Weight

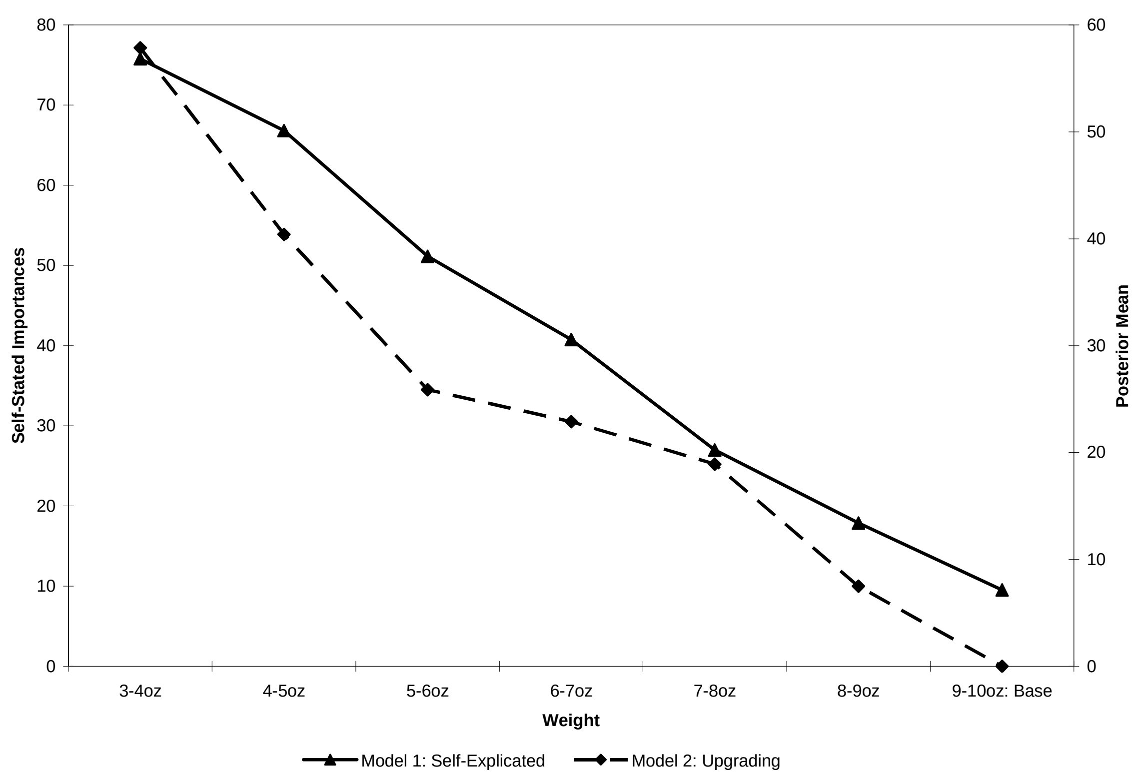

Figure 5: Model Results for Brand

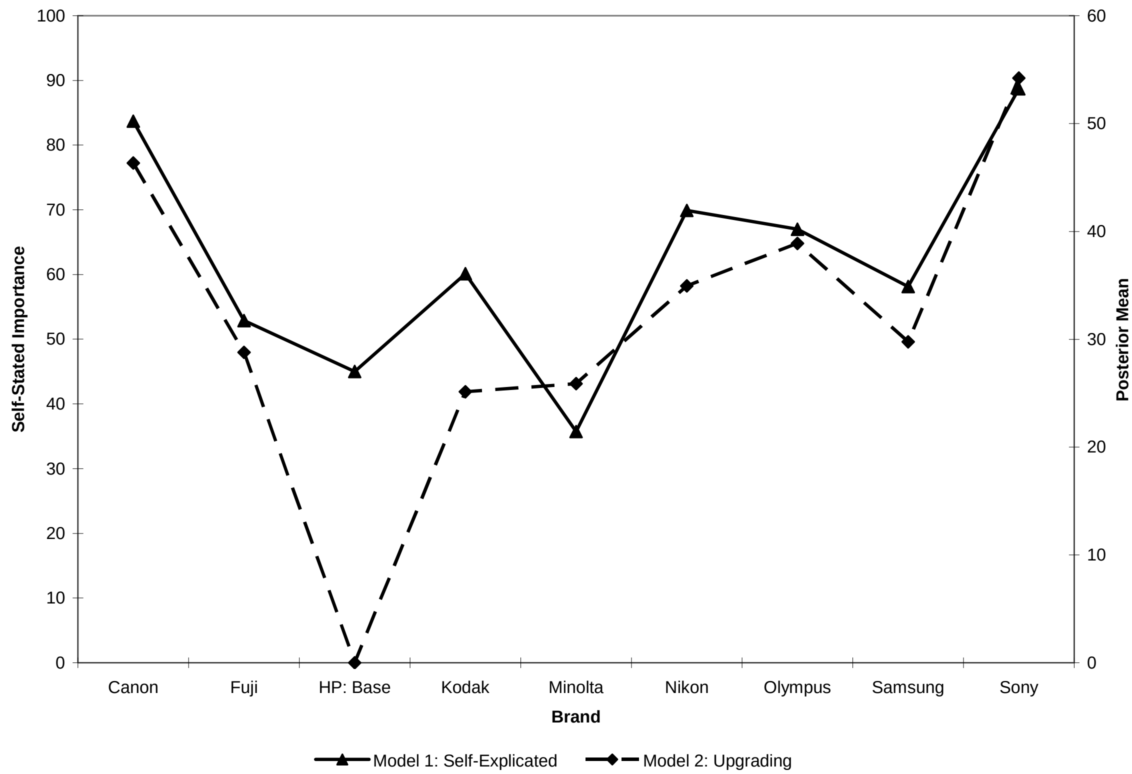

Figure 6: Predictive Performance for the External Validity Tasks