CHAPTER

# Economic foundations of conjoint analysis

3

**Greg M. Allenby**[a], **Nino Hardt**[a], **Peter E. Rossi**[b,*]

[a]*Fisher College of Business, Ohio State University, Columbus, OH, United States*

[b]*Anderson School of Management, University of California at Los Angeles, Los Angeles, CA, United States*

**Corresponding author: e-mail address: perossichi@gmail.com*

## Contents



**Handbook of the Economics of Marketing, Volume 1, ISSN 2452-2619, https://doi.org/10.1016/bs.hem.2019.04.002**

# 1 Introduction

Conjoint analysis is a survey-based method used in marketing and economics to estimate demand in situations where products can be represented by a collection of features and characteristics. It is estimated that 14,000 conjoint studies are conducted yearly by firms with the goal of valuing product features and predicting the effects of changes in formulation, price, advertising, or method of distribution (Orme, 2010). Conjoint analysis is often the only practical solution to the problem of predicting demand for new products or for new features of products that are not present in the marketplace.

The economic foundations of conjoint analysis can be traced to seminal papers on utility measurement in economics (Becker et al., 1964) and mathematical psychology (Luce and Tukey, 1964). Conjoint analysis was introduced to the marketing literature by Green and Rao (1971) as an approach to take ranked input data and estimate utility part-worths for product attributes and their levels (e.g., amount of computer memory). The use of dummy-variables to represent the utility of attribute-levels provided a flexible approach to measure preferences without imposing unnecessary assumptions on the form of the utility function. Later, nonparametric and semiparametric models of choice were developed to avoid incorrect assumptions about the distribution of the error term (Matzkin, 1991, 1993). This earlier work does not lend itself to models of heterogeneity that employ a random-effect specification across respondents (Allenby and Rossi, 1998) which have become popular in marketing.

Virtually all conjoint studies done today are in the form of a "choice-based" conjoint in which survey respondents are offered the choice between sets of products represented by combinations of well-defined product attributes or features. For many durable good products, the assumption of mutually exclusive unit demand might be appropriate. In these cases, respondents can only choose one of the product offerings (or the "outside" alternative) with unit demand. In other situations, it may be more reasonable to allow respondents to choose a subset of products and to consume continuous quantities. This form of conjoint analysis is called "volumetric" conjoint and is not practiced very extensively due to lack of software for analysis of volumetric conjoint data. As a companion to this chapter, we hope to promote use of volumetric conjoint designs by providing general purpose software in an R package.

Both standard and volumetric conjoint analysis can best be viewed as a simulation of market conditions set up via an experimental design. As such, choice-based conjoint (if executed well) is closer to revealed rather than stated preferences. Virtually all researchers in marketing accept the premise that choice-based conjoint studies offer superior recovery of consumer preferences than a pure stated preference method in which direct elicitation of preferences is attempted. However, there is remarkably little research in the conjoint literature that attempts to compare preferences estimated from choice-based conjoint with preferences estimated from consumer panel observational data (cf. Louviere and Hensher, 1983).

Even with existing products for which marketplace data is observed, there are many situations in which it is not possible to identify consumer preferences. In many markets, there is very little price variation (at least in the short run where preferences might be assumed to be stable) and the set of existing products represent only a very sparse set of points in the product characteristic space. In these situations, conjoint has a natural appeal. In still other situations, only aggregate demand data is available. The considerable influence of the so-called BLP (Berry et al., 1995) approach notwithstanding, many recognize that it is very difficult to estimate both preferences and the distribution of preferences over consumers from only aggregate data.

Econometricians have focused considerable attention on the problem of "endogeneity" in demand estimation. With unobservable product characteristics or other drivers of demand, not all price variation may be usable for preference inference. In these situations, econometricians advocate a variety of instrumental variable and other techniques which further restrict the usable portion of price variation. In contrast, conjoint data has all product characteristics well-specified and the levels of price and characteristics are chosen by experimental design. Therefore, there is no "endogeneity" problem in conjoint analysis - all variation in both product attributes and price is exogenous and usable in preference estimation. This is both a great virtue as well as a limitation of conjoint analysis. Conjoint analysis avoids the problems that plague valid demand inference with observational data but at the cost of requiring that all product attributes are well-specified and can be described to the survey respondents in a meaningful way. Conjoint researchers have shown great inventiveness in describing product features but always face the limitation that the analysis is limited to the features included. Respondents are typically instructed to assume that all unspecified features are constant across the choice alternatives presented in the conjoint survey. In some situations, this assumption can seem strained.

It should be emphasized that conjoint analysis can only estimate demand. Conjoint survey data, alone, cannot be used to compute market equilibrium outcomes such as market prices or equilibrium product positioning in characteristics space. Clearly, supply assumptions and cost information must be added to conjoint data in order to compute equilibrium quantities. Conjoint practitioners have chosen the unfortunate term "market simulation" to describe demand predictions based on conjoint analyses. These practices are not simulations of any market outcome and must be understood for their limitations as we discuss below.

The purpose of this chapter in the Handbook of the Economics of Marketing is to provide an introduction to the economic foundations of modern conjoint analysis. We begin in Section 2 with a discussion of the economic justification for conjoint analysis, examining economic models of choice that can be adapted for conjoint analysis.

Section 3 discusses economic measures of value that can be derived from a conjoint study. Section 4 discusses survey requirements for conducting valid conjoint analysis, covering topics such as the use of screening questions and the task of selecting and describing product attributes. Section 5 considers aspects of current practice in conjoint analysis that are not supported by economic theory and should be avoided. Section 6 provides evidence that demand data and conjoint data provide similar estimates of feature value. Section 7 offers concluding remarks.

## 2 Conjoint analysis

Conducting a valid conjoint analysis requires both a valid procedure for collecting conjoint data as well as a valid model for analysis. Conjoint data is not the same as stated preference data in that respondents in conjoint surveys are not simply declaring what they know about their utility or, equivalently, their willingness to pay for products and their features. Instead, respondents react to hypothetical purchase scenarios involving products, which may not currently exist in the marketplace, that are specifically described and priced. Respondents provide expected demand quantities across a variety of buying scenarios in which product attributes and prices change. Well-executed conjoint surveys can approximate revealed preference data available from sources like shopper panels by instructing respondents to think about a specific buying context and focus on a subset of product alternatives.

The economic foundation of conjoint analysis rests on using valid economic models for analyzing this data. In this chapter we discuss two models based on direct utility maximization – the discrete choice model based on extreme value errors that leads to the standard logit specification, and a volumetric demand model that allows for both corner and interior solutions as well as continuous demand quantities. In both specifications, coefficients associated with the marginal utility of an offering are parameterized in terms of product characteristics.

### 2.1 Discrete choices

A simple model of choice involves respondents selecting just one choice alternative. Examples include demand for durable goods, such as a car, where only one good is demanded. Models for the selection of one good are referred to as discrete choice models (Manski et al., 1981) and can be motivated by a linear[1] direct utility function:

$$u(\mathbf{x}, z) = \sum_k \psi_k x_k + \psi_z z \tag{1}$$

[1] The linearity assumption is only an approximation in that we expect preferences to exhibit satiation. However, since demand quantities are restricted to unit demand, satiation is not a consideration.

where $x_k$ denotes the demand quantities, constrained to be either zero or one for the choice alternative $k$, $z$ is an outside good, and $\psi_k$ is the marginal utility of consuming the $k$th good. Respondents are assumed to make choices to maximize their utility $u(\mathbf{x}, z)$ subject to budget constraint:

$$\sum_k p_k x_k + z = E \tag{2}$$

where the price of the outside good is set to \$1.00. The outside good represents the decision to allocate some or all of the budgetary allotment $E$ outside of the $k$ goods in the choice task, including the option to delay a purchase. The budgetary allotment is the respondent's upper-limit of expenditure for the goods under study and is not meant to indicate his or her annual income.

An additive error term is introduced into the model specification for each good to allow for factors affecting choice that are known to the respondent but not observed by the researcher. Assuming distributional support for the error term on the real line allows the model to rationalize any pattern of respondent choices. There are some error realizations, however unlikely, that lead to an observed choice being utility maximizing. The utility for selecting good $j$ is therefore equal to:

$$u\left(x_j = 1, z = E - p_j\right) = \psi_j + \psi_z\left(E - p_j\right) + \varepsilon_j \tag{3}$$

and the utility for selecting the 'no-choice' option is:

$$u\left(x = 0, z = E\right) = \psi_z E + \varepsilon_z \tag{4}$$

Choice probabilities are obtained from the utility expressions by integrating over regions of the error space that coincide with a particular choice having highest utility. Assuming extreme value EV(0,1)[2] errors leads to the familiar logit choice probability:

$$\begin{aligned}
\Pr(j) &= \Pr\left(\psi_j + \psi_z\left(E - p_j\right) + \varepsilon_j > \psi_k + \psi_z\left(E - p_k\right) + \varepsilon_k \text{ for any } k \neq j\right) \\
&= \frac{\exp\left[\psi_j + \psi_z\left(E - p_j\right)\right]}{\exp\left[\psi_z(E)\right] + \sum_k \exp\left[\psi_k + \psi_z\left(E - p_k\right)\right]} \\
&= \frac{\exp\left[\psi_j - \psi_z p_j\right]}{1 + \sum_k \exp\left[\psi_k - \psi_z p_k\right]}
\end{aligned} \tag{5}$$

The discrete choice model is used extensively in conjoint applications because of its computational simplicity. Observed demand is restricted to two points {0, 1} for each of the inside goods ($x$) and the constant marginal utility assumption leads to the

[2] The cdf of the EV(0,1) error term is F(x) = exp[-exp(-x)] with location parameter equal to zero and scale parameter equal to one.

budgetary allotment dropping out of the expression for the choice probability.[3] The marginal utility for the outside good, $\psi_z$, is interpreted as a price coefficient and measures the disutility for paying a higher price. Choice alternatives with prices that are larger than the budgetary allotment are screened-out of the logit choice probability. That is, only goods for which $p_k \leq E$ are included in the logit probability expression (see Pachali et al., 2017). Finally, expected demand is equal to the choice probability, which is convenient for making demand predictions.

Conjoint analysis uses the discrete choice model by assuming marginal utility, $\psi_j$, is a linear function of brand attributes:

$$\psi_j = \mathbf{a}_\mathbf{j}'\beta \tag{6}$$

where $\mathbf{a_j}$ denotes the attributes of good $j$. The attributes are either coded using a dummy variable specification where one of the attribute-levels is selected as the null level of the attribute, or using effects-coding that constrain the sum of coefficients to equal zero for an attribute (Hardy, 1993). For either specification, conjoint analysis measures the value of changes among the levels of an attribute. Since the valuation of the attribute-levels are jointly determined through Eq. (6), the marginal utilities for a respondent are comparable across the attributes and features included in the analysis. The marginal utility of a product feature can therefore be compared to the utility of changes in the levels of other attributes, including price.

## 2.2 Volumetric choices

A more general model for conjoint analysis is one that introduces non-linearities into the utility function (Allenby et al., 2017):

$$u(\mathbf{x}, z) = \sum_k \frac{\psi_k}{\gamma} \ln(\gamma x_k + 1) + \ln(z) \tag{7}$$

where $\gamma$ is a parameter that governs the rate of satiation of the good. The marginal utility for the inside and outside goods is:

$$\begin{aligned} u_j &= \frac{\partial u(\mathbf{x}, z)}{\partial x_j} = \frac{\psi_j}{\gamma x_j + 1} \\ u_z &= \frac{\partial u(\mathbf{x}, z)}{\partial z} = \frac{1}{z} \end{aligned} \tag{8}$$

The marginal utility of the good is equal to $\psi_j$ when $x_j = 0$ and decreases as demand increases ($x_j > 0$) and as the satiation parameter ($\gamma$) increases. The general solution to maximizing the utility in (7) subject to the budget constraint in (2) is to employ the

3 Constant marginal utility, however, is not required to obtain a discrete choice model as shown by Allenby and Rossi (1991), wherein the budgetary allotment does not drop out.

Kuhn-Tucker (KT) optimality conditions:

$$
\begin{array}{llllll}
\text{if} \quad x_k > 0 & \text{and} & x_j > 0 & \text{then} & \lambda = \dfrac{u_k}{p_k} = \dfrac{u_j}{p_j} & \text{for all } k \text{ and } j \\
\text{if} \quad x_k > 0 & \text{and} & x_j = 0 & \text{then} & \lambda = \dfrac{u_k}{p_k} > \dfrac{u_j}{p_j} & \text{for all } k \text{ and } j
\end{array} \tag{9}
$$

Assuming that $\psi_j = \exp[a'_j\beta + \varepsilon_j]$ and solving for $\varepsilon_j$ leads to the following expression for the KT conditions:

$$\varepsilon_j = g_j \quad \text{if} \quad x_j > 0 \tag{10}$$

$$\varepsilon_j < g_j \quad \text{if} \quad x_j = 0 \tag{11}$$

where

$$g_j = -\mathbf{a}'_j\beta + \ln(\gamma x_j + 1) + \ln\left(\frac{p_j}{E - \mathbf{p}'\mathbf{x}}\right) \tag{12}$$

The assumption of $i.i.d.$ extreme-value errors, i.e., $\text{EV}(0, \sigma)$,[4] results in a closed-form expression for the probability that $R$ of $N$ goods are chosen. The error scale ($\sigma$) is identified in this model because price enters the specification without a separate price coefficient. We assume there are $N$ choice alternatives and $R$ items are chosen:

$$x_1, x_2, \ldots, x_R > 0, \qquad x_{R+1}, x_{R+2}, \ldots, x_N = 0.$$

The likelihood $\ell(\theta)$ of the model parameters is proportional to the probability of observing $n_1$ chosen goods ($n_1 = 1, \ldots, R$) and $n_2$ goods with zero demand ($n_2 = R + 1, \ldots, N$). The contribution to the likelihood of the chosen goods is in the form of a probability density because of the equality condition in (10) while the goods not chosen contribute as a probability mass because of the inequality condition in (11). We obtain the likelihood by evaluating the joint density of model errors at $g_j$ for the chosen goods and integrating the joint density to $g_i$ for the goods that are not chosen:

$$
\begin{aligned}
\ell(\theta) &\propto p(x_{n_1} > 0, x_{n_2} = 0|\theta) \\
&= |J_R| \int_{-\infty}^{g_N} \cdots \int_{-\infty}^{g_{R+1}} f(g_1, \ldots, g_R, \varepsilon_{R+1}, \ldots, \varepsilon_N) d\varepsilon_{R+1}, \ldots, d\varepsilon_N \\
&= |J_R| \left\{ \prod_{j=1}^{R} \frac{\exp(-g_j/\sigma)}{\sigma} \exp\left(-e^{-g_j/\sigma}\right) \right\} \left\{ \prod_{i=R+1}^{N} \exp\left(-e^{-g_i/\sigma}\right) \right\} \\
&= |J_R| \left\{ \prod_{j=1}^{R} \frac{\exp(-g_j/\sigma)}{\sigma} \right\} \exp\left\{ -\sum_{i=1}^{N} \exp(-g_i/\sigma) \right\}
\end{aligned}
$$

---

[4] $F(x) = \exp[-e^{-x/\sigma}]$.

where $f(\cdot)$ is the joint density for $\varepsilon$ and $|J_R|$ is the Jacobian of the transformation from random-utility error ($\varepsilon$) to the likelihood of the observed data ($x$), i.e., $|J_R| = |\partial\varepsilon_i/\partial x_j|$. For this model, the Jacobian is equal to:

$$|J_R| = \prod_{k=1}^{R}\left(\frac{\gamma}{\gamma x_k + 1}\right)\left\{\sum_{k=1}^{R}\frac{\gamma x_k + 1}{\gamma}\cdot\frac{p_k}{E - \mathbf{p}'\mathbf{x}} + 1\right\}$$

The expression for the likelihood of the observed demand vector $x_t$ is seen to be the product of $R$ "logit" expressions multiplied by the Jacobian, where the purchased quantity, $x_j$ is part of the value ($g_j$) of the choice alternative. To obtain the standard logit model of discrete choice we set $R = 1$, set the scale of the error term to one ($\sigma = 1$), and allow the expression for $g_j$ to include a price coefficient (i.e., $g_j = -a_j'\beta - \psi_z p_j$). The Jacobian equals one for a discrete choice model because demand ($x$) enters the KT conditions through the corner solutions only corresponding to mass points.

Variation in the specification of the choice model utility function and budget constraint results in different values of ($g_j$) and the Jacobian $|J_R|$, but not to the general form of the likelihood, i.e.,

$$p(x|\theta) = |J_R|\left\{\prod_{j=1}^{R} f(g_j)\right\}\left\{\prod_{i=R+1}^{N} F(g_i)\right\} \tag{13}$$

The analysis of conjoint data arising from either discrete or volumetric choices proceeds by relating utility parameters ($\psi_j$) to product attributes ($a_j$) as in Eq. (6). It is also possible to specify non-linear mappings from attributes to aspects of the utility function as discussed, for example, in Kim et al. (2016).

## 2.3 Computing expected demand

Demand ($\mathbf{x}_{ht}$) for subject $h$ at time $t$ is a function of parameters of the demand model ($\theta_h$), a realization of the vector of error terms ($\varepsilon_{ht}$), characteristics of the available choice set ($A_t$), and prices ($\mathbf{p}_t$). We need to obtain expected demand quantities for deriving various measures of economic value. For the discrete choice model, expected demand is expressed as a choice probability, while expected demand for the volumetric demand model does not have a closed-form solution. We therefore introduce D, the function of demand for one realization of model parameters $\theta_h$ and one realization of the error term $\varepsilon_{ht}$, given the characteristics of the set of alternatives $\mathbf{A}_t$ and corresponding prices $\mathbf{p}_t$:

$$\mathbf{x}_{ht} = \mathrm{D}\left(\theta_h, \varepsilon_{ht}|\mathbf{A}_t, \mathbf{p}_t\right) \tag{14}$$

Here, $\theta_h = \left\{\beta_h, \psi_{h,z} = \beta_{hp}\right\}$ for the discrete choice model and $\theta_h = \{\beta_h, \gamma_h, E_h, \sigma_h\}$ for the volumetric model. Expected demand is obtained by integrating out the error term and posterior distribution of model parameters $\theta_h$. First, consider integrating over the error term:

$$E\left(\mathbf{x}_{ht}|\theta_h\right) = \int_{\varepsilon_{ht}} \mathrm{D}\left(\theta_h, \varepsilon_{ht}|\mathbf{A}_t, \mathbf{p}_t\right) p(\varepsilon_{ht}) d\varepsilon_{ht} \tag{15}$$

For volumetric choices, there is no closed form solution for integrating out $\varepsilon_{ht}$ and numerical methods are used to obtain expected demand by simulating a large number of realizations of $\varepsilon$ and computing D for each. Expected demand is the average of D values. One way to compute D is to use a general-purpose algorithm for maximization, such as `constrOptim` in R. A more efficient algorithm is provided in the appendix.

## 2.4 Heterogeneity

Data from a conjoint study can be characterized as being wide and shallow – i.e., wide in the sense that there may be 1000 or more respondents included in the study and shallow in the sense that each respondent provides, at most, about 8-16 responses to the choice task. This type of panel structure is commonly found in marketing applications involving consumer choices. We therefore employ hierarchical models, or a random-effect specification, to deal with the lack of data at the individual respondent-level. The lower level of the hierarchical model applies the direct utility model to a specific respondent's choice data, and the upper level of the model incorporates heterogeneity in respondent coefficients.

Respondent heterogeneity can be incorporated into conjoint analysis using a variety of random-effect models. The simplest and most widely used is a Normal model for heterogeneity. Denoting all individual-level parameters $\theta_h$ for respondent $h$ we have:

$$\theta_h \sim \mathrm{Normal}(\bar{\theta}, \Sigma) \tag{16}$$

where $\theta_h = \left\{\beta_h', \gamma_h, E_h, \sigma_h\right\}$ for the volumetric model, and $\theta_h = \left\{\beta_h, \beta_{ph}\right\}$ for the discrete choice model. The price coefficient $\beta_{ph}$ is referred to as $(\psi_z)$ in Eq. (5).

Estimation of this model is easily done using modern Bayesian (Monte Carlo Markov chain) methods as discussed in Rossi et al. (2005). More complicated models for heterogeneity involve specifying more flexible distributions. One option is to specify a mixture of Normal distributions:

$$\theta_h \sim \sum_k \varphi_k \mathrm{Normal}(\bar{\theta}_k, \Sigma_k) \tag{17}$$

(here $\varphi_k$ are the mixture probabilities) or to add covariates ($z$) to the mean of the heterogeneity distribution as in a regression model:

$$\theta_h \sim \mathrm{Normal}(\Gamma z_h, \Sigma) \tag{18}$$

The parameters $\bar{\theta}$, $\Gamma$, and $\Sigma$ are referred to as hyper-parameters ($\tau$) because they describe the distribution of other parameters. Covariates in the above expression might include demographic variables or other variables collected as part of the conjoint survey, e.g., variables describing reasons to purchase. Alternatively, one could use a non-parametric distribution of heterogeneity as discussed in Rossi (2014). The advantage of employing Bayesian estimators for models of heterogeneity is that they provide access to individual-level parameters $\theta_h$ in addition to the hyper-parameters ($\tau$) that describe their distribution.

It should be emphasized that the hyper-parameters $\tau$ can be shown to be consistently estimated as the sample size, or number of respondents, in the conjoint survey increases. However, the individual-level estimates of $\theta_h$ cannot be shown to be consistent because the number of conjoint question for any one respondent is constrained to be small. Therefore inferences for conjoint analysis should always be based on the hyper-parameters and not on the individual-level estimates.

## 2.5 Market-level predictions

In order to make statements about a population of customers, we need to aggregate the results from the hierarchical model by integrating over the distribution heterogeneity. These statements involve quantities of interest $\mathcal{Z}$, such as expected demand ($\mathbf{x}_t$), or derived quantities based on the demand model such as willingness-to-pay, price elasticities, and associated confidence intervals. We can approximate the posterior distribution of $\mathcal{Z}$ by integrating out the hyper-parameters of the distribution of heterogeneity ($\tau$) and model error ($\varepsilon_{ht}$):

$$p\left(\mathcal{Z}|Data\right) = \int_{\tau}\int_{\theta_h}\int_{\varepsilon_{ht}} p\left(\mathcal{Z}|\theta_h, \varepsilon_{ht}\right) p\left(\theta_h|\tau\right) p\left(\tau\right) p(\varepsilon_{ht}) d\varepsilon_{ht} d\theta_h d\tau \qquad (19)$$

The integration is usually done numerically. Posterior realizations of $\tau$ are available when using a Markov chain Monte Carlo estimation algorithm, and can be used to generate individual-level draws of $\theta_h$. These draws, along with draws of the model error term, can then be used to evaluate posterior estimates of $\mathcal{Z}$. The distribution of evaluations of $\mathcal{Z}$ can be viewed as its posterior distribution.

## 2.6 Indirect utility function

The indirect utility function is the value function of the maximum attainable utility of a utility maximization problem. It is useful to define the indirect utility function for the volumetric demand case, as it is the basis for computing economic measures such as willingness-to-pay. Conditional on the characteristics of the choice alternatives $A$, parameters $\theta_h$ and a realization of the vector error terms $\varepsilon_{ht}$, the indirect utility function is defined in terms of optimal demand $(\mathbf{x}^*_{ht}, z^*_{ht})$:

$$V\left(\mathbf{p}_t, E|A, \theta_h, \varepsilon_{ht}\right) = u\left(\mathbf{x}^*_{ht}, z^*_{ht}|\mathbf{p}_t, \theta_h, \varepsilon_{ht}\right) = u\left(\mathrm{D}\left(\theta_h, \varepsilon_{ht}|A, \mathbf{p}_t\right)\right) \qquad (20)$$

Eq. (20) can be evaluated by first determining optimal demand and then computing the corresponding value of the direct utility function (Eq. (7)). There is no closed form solution for the value function since there is no closed-form solution to $D$. Moreover, the error term $\varepsilon_{ht}$ and individual-level parameters $(\theta_h)$ need to be integrated out to obtain an expression for expected indirect utility. We do this by using the structure of the heterogeneity distribution, where uncertainty in the hyper-parameters $(\tau)$ induces variation on the distribution of individual-level parameters $(\theta_h)$:

$$V(\mathbf{p}_t, E|\mathbf{A}_t) = \int_\tau \int_{\theta_h} \int_{\varepsilon_{ht}} u\left(\mathrm{D}\left(\theta_h, \varepsilon_{ht}|\mathbf{A}_t, \mathbf{p}_t\right) p\left(\theta_h|\tau\right) p\left(\tau\right)\right) p(\varepsilon_{ht}) d\varepsilon_{ht} d\theta_h d\tau \tag{21}$$

## 3 Measures of economic value

The economic value of a product feature requires an assessment of consumer welfare and its effect on marketplace demand and prices. Measuring the value of a product feature often begins by expressing part-worths $(\beta_h)$ in monetary terms by dividing by the price coefficient, i.e., $\beta_h/\beta_p$. This monetization of utility is useful when summarizing results across respondents because utility, by itself, is not comparable across consumers. A monetary conversion of the part-worth of a product feature, however, is not sufficient for measuring economic value because it does not consider the effects of competitors in the market.

A criticism of the simple monetization of utility is that consumers never have to pay the maximum they are willing in the marketplace. A billionaire may be willing to pay a large amount of money to attend a sporting event, but does not end up doing so because of the availability of tickets sold by those with lower WTP. Firms in a competitive market are therefore not able to capture all of the economic surplus of consumers. Consumers can switch to other providers and settle for alternatives that are less attractive but still worth buying. Below we investigate three measures of economic value that incorporate the effects of competing products.

In theory, WTP can be the monetary value of an entire option or of a feature-change (or improvement) to a given product (Trajtenberg, 1989). As discussed below, WTP for an entire choice option involves the addition of a new model error term for the new product, while a new error term is not added when assessing the effects of a feature-change. The perspective in this chapter is focused on the feature change/improvement.

### 3.1 Willingness to pay (WTP)

WTP is a demand-based estimate of monetary value. The simple monetization of utility to a dollar measure (e.g., $\beta_h/\beta_{hp}$) does not correspond to a measure of consumer welfare unless there is only one good in the market and consumers are forced to select it. The presence of alternatives with non-zero choice probabilities means that the maximum attainable utility from a transaction is affected by more than one good. Increasing the number of available choice alternatives increases the expected maximum utility a consumer can derive from a marketplace transaction, and ignoring the effect of competitive products leads to an misstatement of consumer welfare. Any measurement of the economic value of a product feature cannot be made in isolation of the set of available alternatives because it is not known, a priori, which product will be chosen (Lancsar and Savage, 2004).

The evaluation of consumer welfare needs to account for private information held by the consumer at the time of choice. This information is represented as the error term in the model, whose value is not realized until the respondent is confronted with a choice. Consumer welfare is determined by the maximum attainable utility of a transaction and cannot be based on the likelihood, or probability that an alternative is utility maximizing. Choice probabilities do not indicate the value, or utility arising from a transaction. Welfare should be measured in terms of the expected maximum attainable utility, ($E[\max(u(.))]$), where the maximization operator is taken over the choice alternatives and the expectation operator is taken over error realizations.

The effect of competitive offers can be taken into account by considering respondent choices among all the choice alternatives in the conjoint study. Let $A$ be a $J \times K$ matrix that define the set of products in a choice set, where $J$ is the number of choice alternatives and $K$ is the number of product features under study. The rows of the choice set matrix, $a_j$ indicates the features of the $j$th product in the choice set. Similarly, let $A^*$ be a choice matrix similar to $A$ except that one of its rows is different indicating a different set of features for one of the products. Typically, just one element in the row differs when comparing $A$ to $A^*$ because the WTP measure typically focuses on what respondents are willing to pay for an enhanced version of one of the attributes.

The maximum attainable utility for a given choice set is defined in terms of the indirect utility function:

$$V(\mathbf{p}, E|\mathbf{A}) = \max_{\mathbf{x}} u(\mathbf{x}, z|\mathbf{A}) \qquad \text{subject to} \qquad \mathbf{p}'\mathbf{x} \leq E \tag{22}$$

WTP is defined as the compensating value required to make the utility derived from the feature-poor set, $A$, equal to the utility derived from the feature-rich set, $A^*$:

$$V(\mathbf{p}, E + \text{WTP}|\mathbf{A}) = V(\mathbf{p}, E|\mathbf{A}^*) \tag{23}$$

### 3.1.1 WTP for discrete choice

The expected maximum attainable monetized utility, or consumer welfare, for a logit demand model can be shown to be equal to (Small and Rosen, 1981; Manski et al., 1981; Allenby et al., 2014b):

$$W\left(\mathbf{A}, \mathbf{p}, \theta_h = \{\beta_h, \beta_{hp}\}\right) = E + \ln\left[\sum_{j=1}^{J} \exp\left(\beta_h' \mathbf{a}_j^* - \beta'_{hp} p_j\right)\right] \Big/ \beta_{hp} \tag{24}$$

This expression measures the benefit of spending some portion of the budgetary allotment, $E$, on one of the choice alternatives. As the attractiveness of the inside goods declines, the consumer is more likely to select the outside good and save his or her money. Thus, the lower bound of consumer welfare confronted with an exchange is their budgetary allotment $E$. The improvement in welfare provided by a feature enhancement can be obtained as the difference of the maximum attainable utility of the enriched and original set of alternatives.

$$\mathrm{WTP} = \ln\left[\sum\nolimits_{j=1}^{J} \exp\left(\beta_h' \mathbf{a}_j^* - \beta'_{hp}\mathbf{p}_j\right)\right] \Big/ \beta_{hp} - \ln\left[\sum\nolimits_{j=1}^{J} \exp\left(\beta_h' \mathbf{a}_j - \beta'_{hp}\mathbf{p}_j\right)\right] \Big/ \beta_{hp} \tag{25}$$

Ofek and Srinivasan (2002) define an alternative measure, which they refer to as the market value of an attribute improvement (MVAI) based on the change in price needed to restore the aggregate choice share of a good in a feature-rich ($A^*$) relative to a feature-poor ($A$) choice set. Their calculation, however, does not monetize the change in consumer welfare, or utility, gained by the feature enhancement because choice probabilities do not provide a measure of either. Moreover, the MVAI measure is not a "market" measure of value as there is no market equilibrium calculation. MVAI also only applies to continuous attributes.

The gain in utility is a function of both the observable product features and the unobserved error realization that are jointly maximized in a marketplace transaction, thus requiring a consideration of the alternative choices available to a consumer because, prior to observing the choice, the values of the error realizations are not known. Our proposed WTP measure monetizes the expected improvement in the maximized utility that comes from feature-enhanced choices.

### 3.1.2 WTP for volumetric choice

For the volumetric choice model, WTP is the additional budgetary allotment amount that is necessary for restoring the indirect utility of the feature-rich set given the feature-poor set. Conditional on realizations of $\varepsilon_{ht}$ and $\theta_h$, WTP can be obtained by numerically solving Eq. (26) for WTP:

$$V\left(\mathbf{p}, E + \mathrm{WTP}|\mathbf{A}, \theta_h, \varepsilon_{ht}\right) - V\left(\mathbf{p}, E|\mathbf{A}^*, \theta_h, \varepsilon_{ht}\right) = 0 \tag{26}$$

Computation of indirect utility $V()$ has been described in Section 2.6. In the volumetric model, subjects are compensated for the loss in utility in a demand space that extends beyond unit demand. WTP therefore depends on purchase quantities, and are expected to be larger when the affected products have larger purchase quantities.

## 3.2 Willingness to buy (WTB)

WTB is an alternative measure of economic value based on the expected increase in demand for an enhanced offering. It is similar to the measure proposed by Ofek and Srinivasan (2002) but does not attempt to monetize the change in share due to a feature enhancement. Instead, economic value is determined by calculating expected increase in revenue or profit due to a feature enhancement, using WTB as an input to that calculation. The increase in demand due to the improved feature is calculated for one offering, holding fixed all of the other offerings in the market.

### 3.2.1 WTB for discrete choice

In a discrete choice model, WTB is defined in terms of the change in market share that can be achieved by moving from a diminished to an improved feature set:

$$\mathrm{WTB} = \mathrm{MS}(j|\mathbf{p}, \mathbf{A}^*) - \mathrm{MS}(j|\mathbf{p}, \mathbf{A}) \tag{27}$$

It is computed as the increase in choice probability for each respondent in the survey, which is then averaged to produce an estimate of the change in aggregate market share. As mentioned above in Section 2.5, statements about the market level behavior require integration over uncertainty in model hyper-parameters ($\tau$), the resulting distribution of individual-level coefficients ($\theta_h$), and the model error ($\varepsilon$).

### 3.2.2 WTB for volumetric choice

We can express WTB as a change in absolute sales or as a change in market share. Since there is no closed form for demand, $D$, we suggest first simulating a set of realizations of $\varepsilon_{ht}$, then compute demand for the initial $A$ and changed feature set $A^*$ conditional on the same set of $\varepsilon_{ht}$ and $\theta_h$ realizations. As shown in Section 2.3, we do this by numerically solving for $D$ and evaluating the change in demand for the two feature sets:

$$\mathrm{WTB}_{\text{sales}} = \int_\varepsilon \mathrm{D}_j\left(\theta, \varepsilon | A^*, \mathbf{p}\right) - \mathrm{D}_j\left(\theta, \varepsilon | A, \mathbf{p}\right) \tag{28}$$

$$\mathrm{WTB}_{\text{share}} = \int_\varepsilon \frac{\mathrm{D}_j\left(\theta, \varepsilon | A^*, \mathbf{p}\right)}{\mathrm{D}\left(\theta, \varepsilon | A^*, \mathbf{p}\right)} - \frac{\mathrm{D}_j\left(\theta, \varepsilon | A, \mathbf{p}\right)}{\mathrm{D}\left(\theta, \varepsilon | A, \mathbf{p}\right)} \tag{29}$$

## 3.3 Economic price premium (EPP)

The EPP is a measure of feature value that allows for competitive price reaction to a feature enhancement or introduction. An equilibrium is defined as a set of prices and accompanying market shares which satisfy the conditions specified by a particular concept of equilibrium. In our discussion below, we employ a standard Nash Equilibrium concept for differentiated products using a discrete choice model of demand, not the volumetric version of conjoint analysis.

The calculation of an equilibrium price premium requires additional assumptions beyond those employed in a traditional, discrete choice conjoint study:

- The demand specification is a standard heterogeneous logit that is linear in the attributes, including prices.
- Constant marginal cost for the product.
- Single product firms. i.e., each firm has just one offering.
- Firms cannot enter or exit the market after product enhancement takes place.
- Firms engage in a static Nash price competition.

The first assumption can be easily replaced by any valid demand system, including the volumetric demand model discussed earlier. One can also consider multi-product firms.

The economic value of a product feature enhancement to a firm is the incremental profits that it will generate:

$$\Delta\pi = \pi(p^{\text{eq}}, m^{\text{eq}}|A^*) - \pi(p^{\text{eq}}, m^{\text{eq}}|A) \tag{30}$$

where $\pi$ are the profits associated with the equilibrium prices and shares given a set of competing products defined by the attribute matrix $A$.

The EPP is the increase in profit maximizing prices of an enhanced product given some assumptions about costs and competitive offerings. Each product provider, one at a time, determines their optimal price given the prices of other products and the cost assumptions. This optimization is repeated for each provider until an equilibrium is reached where it is not in anyone's interest to change prices any further. Equilibrium prices are computed for the offerings with features set at the lower level for all the attributes $(A)$, and then recomputed for the offerings set to its higher level $(A^*)$. EPP introduces the concept of price competition in the valuation of product features assuming static Nash price competition.

In a discrete choice setting, firm profits is

$$\pi(p_j|p_{-j}) = MS(j|p_j, p_{-j}, A)(p_j - c_j) \tag{31}$$

where $p_j$ is the price of good $j$, $p_{-j}$ are the prices of other goods, and $c_j$ is the marginal cost of good $j$. The first-order conditions of the firm are:

$$\frac{\partial\pi}{\partial p_j} = \frac{\partial}{\partial p_j} MS(j|p_j, p_{-j}, A)(p_j - c_j) + MS(j|p_j, p_{-j}, A) \tag{32}$$

The Nash equilibrium is a root of the system of equations defined by the first-order conditions for all $J$ firms. If we define:

$$h(\mathbf{p}) = \begin{bmatrix} h_1(\mathbf{p}) = \frac{\partial\pi}{\partial p_1} \\ h_2(\mathbf{p}) = \frac{\partial\pi}{\partial p_2} \\ \vdots \\ h_J(\mathbf{p}) = \frac{\partial\pi}{\partial p_J} \end{bmatrix} \tag{33}$$

then the equilibrium price vector $\mathbf{p}^*$ is a zero of the function $h(\mathbf{p})$.

## 4 Considerations in conjoint study design

Conjoint studies rely on general principles of survey design (Diamond, 2000) in collecting data for analysis. This includes the consideration of an appropriate sampling frame, screening questions, and conjoint design to provide valid data for analysis. Of critical concern in designing a conjoint survey is to ensure that questions are posed in an unambiguous and easy-to-understand manner. Specifically, the conjoint attributes and their levels should be specified so as not to bias preference measurement and introduce unnecessary measurement error due to confusion or lack of understanding. Typically, conjoint design is informed by qualitative research prior to the drafting of the conjoint survey questionnaires.

Questionnaires should also be pre-tested to determine if the survey is too long and if the information intended to be collected in the survey questions are understandable to survey respondents. For example, a survey designed to understand consumer demand for smartphones might include questions about product usage in form of data and text charges. Not all respondents may have a clear understand of the term ‘data,’ and this question might need to be phrased in terms of wording used in current advertisements (e.g., ‘talk’ and text). Records should be kept of both the qualitative and pre-testing phases of the survey design. In particular, revisions in the questionnaire should be documented.

Modern conjoint studies are often administered using Internet panels where respondents are recruited to join a survey. Internet panel providers such as Survey Sampling International (https://www.surveysampling.com) or GfK (https://www.gfk.com) maintain a population of potential respondents who are emailed invitations to participate in a survey. Respondents usually do not know the exact nature of the survey (e.g., a survey about laundry detergent) which helps to reduce self-selection bias. Participants are then asked to complete a survey that typically begins with a series of demographic and screening questions to help establish the representativeness of the initial sample and to exclude respondents who do not qualify for inclusion in the survey. As discussed below, the general population is generally not the intended target of the survey because not all people have the interest or ability to answer the survey questions. Next is a series of questions that document attitudes, opinions and behaviors related to the focal product category. A glossary section follows in which attributes and attribute levels of the conjoint study are defined, following by the conjoint choice tasks. Additional classification variables for the analysis of subpopulations of interest are included at the end of the survey. We consider each of these survey components in detail below.

## 4.1 Demographic and screening questions

The incidence of product usage varies widely across product categories, and screening questions help ensure that only qualified respondents are surveyed. Respondents in a conjoint study should represent a target market of potential buyers who have preexisting interests in the product category, often identified as people who have recently purchased in the product category and those who report they are considering a purchase in the near future. These individuals are referred to as “prospects” in marketing textbooks, defined as individuals with the willingness and ability to make purchases (Kotler, 2012).

The purpose of the screening questions is to remove respondents who either lack the expertise to provide meaningful answers to the survey questions, or who are not planning on making a purchase in the near future. The presence of screening questions creates some difficulty in assessing the representativeness of the survey sample. Demographic variables are available to profile the general population, but not prospects in a specific product category. Some product categories appeal to younger consumers and others to older consumers. Claiming that the resulting sample is demographically representative therefore relies on obtaining a representative sample prior to screening respondents out of the survey.

There is a recent conjoint literature on incentive-alignment techniques used to induce greater effort among respondents so that their responses provider a 'truer' reflection of their actual preferences (Ding et al., 2005; Ding, 2007; Yang et al., 2018). The idea behind these techniques is that respondents will be more likely to provide thoughtful responses when offered an incentive that is derived from their choices, such a being awarded some version of the product under study that is predicted to give them high utility. In this literature, it is common to apply conjoint analysis to student samples in a lab setting. Improvements in predicting hold-out choices of these respondents are offered as evidence of increased external validity. However, a test of external validity needs to relate choice experiments to actual marketplace behavior. Shoppers in an actual marketplace setting may face time constraints and distractions after a busy workday that differ from a lab setting. Motivating lab subjects through incentive alignment does not necessarily lead to more realistic marketplace predictions and inferences.

In industry-grade conjoint studies, respondents are screened so that conjoint surveys are relevant. Some panel providers can screen respondents based on past purchases from a given category to ensure familiarity with the product category. Respondents thus have the opportunity to help improve products relevant to them, so there is little reason to assume they would receive utility from lying. Moreover, it is often not possible to conduct incentive-aligned studies because Internet panel providers do not allow additional incentives to be offered to panel members. At worst, respondents may respond in a careless manner. Respondents with low in-sample model fit can be screened out of the sample as discussed in Allenby et al. (2014b), Section 6. We show below that conjoint estimates can be closely aligned with marketplace behavior without the use of incentive-alignment by screening for people who actively purchase in the product category.

## 4.2 Behavioral correlates

Respondents participating in a conjoint study are often queried about their attitudes and opinions related to the product category. A study of smartphones would include questions about current payment plans, levels of activity (voice, text, Internet), and other electronic devices owned by the respondent. Questions might also involve details of their last purchase, competitive products that were considered, and specific product features that the respondent found attractive and/or frustrating to use.

The purpose of collecting this data is two-fold: i) it encourages the respondent to think about the product and recall aspects of a recent or intended purchase that are important to them; and ii) it provides an opportunity to establish the representativeness of the sample using data other than demographics. The behavioral correlates serve to 'warm up' the respondent so that they engage in the choice tasks with a specific frame of reference. It also provides information useful in exploring antecedents and consequences of product purchase through their relationship to estimates of the conjoint model.

Behavioral covariates are often reported for products by various media, and these reports can be used as benchmarks for assessing sample representativeness. For example, trade associations and government agencies report on the number of products or devices owned by households, and the amount of time people spend engaged in various activities. There are also syndicated suppliers of data that can be used to assess product penetration, product market shares, and other relevant information to demand in the product category under study. Behavioral covariates are likely to be more predictive of product and attribute preferences than simple demographic information and, therefore, may be of great value in establishing representativeness.

## 4.3 Establishing representativeness

In many instances, a survey is done for the purpose of projecting results to a larger target population. In particular, a conjoint survey is often designed to project demand estimates based on the survey sample to the larger population of prospective buyers. In addition, many of the inference methods used in the analysis of survey data are based on the assumption that the sample is representative or, at least, acquired via probability sampling methods.

All sampling methods start with a target population, a sample frame (a particular enumeration of the set of possible respondents to be sampled from), and a sampling procedure. One way of achieving representativeness is to use sampling procedures that ensure representativeness by construction. For example, if we want to construct a representative sample of dentists in the US, we would obtain a list of licensed dentists (the sample frame) and use probability sampling methods to obtain our sample. In particular, simple random sampling (or equal probability of selection) would produce a representative sample with high probability. The only way in which random sampling would not work is if the sample size was small. Clearly, this approach is ideal for populations for which there are readily available sample frames.

The only barrier to representativeness for random samples is potential non-response bias. In the example of sampling dentists (regarding a new dental product, for example), it would be relatively easy to construct a random sample but the survey response rate could be very low (less than 50 per cent). In these situations, there is the possibility of a non-response bias, i.e. that those who respond to the survey have different preferences than those who do not respond. There is no way to assess the magnitude of non-response bias except to field a survey with a higher response rate. This puts a premium on well-designed and short surveys and careful design of adequate incentive payments to reduce non-response.

However, it is not always possible to employ probability-based sampling methods. Consider the problem of estimating the demand for a mid-level SUV prototype.

Here we use conjoint because this prototype is not yet in the market. Enumerating the target population of prospective customers is very difficult. One approach would be to start with a random sample of the US adult population and then screen this sample to only those who are in the market for a new SUV. The sample prior to screening is sometimes called the "inbound" sample. If this sample is representative, then clearly any resulting screened sample will also be representative, unless there is a high non-response rate to the screening questions.

Of course, this approach relies on starting with representative sample of the US adult population. There are no sample frames for this population. Instead, modern conjoint practitioners use internet panels maintained by various suppliers. These are not random samples but, instead, represent the outcome of a continuous operation by the supplier to "harvest" the email addresses of those who are willing to take surveys. Internet panel providers often tout the "quality" of their panels but, frequently, this is not a statement about the representativeness of the sample but merely that the provider undertakes precautions to prevent fraud of various types including respondents who are bots or who live outside of the US.

Statisticians will recognize that the internet panels offered commercially are what are called "convenience" samples and there is no assurance of representativeness. This means that it is incumbent on the researcher who uses an internet panel to establish representativeness by providing affirmative evidence of the representativeness of their sample. It is our view that, with adequate affirmative evidence, samples that are based on internet panels can be used as representative.

Internet panel providers are aware of the problem of establishing representativeness and have adopted a variety of approaches and arguments. The first argument is that their panel may be very large (exceeding 1 million). The argument here is that this makes it more difficult for the internet panel to be skewed toward any particular subpopulation. This is not a very strong argument given that there are some 250 million US adults.

Internet panels tend to be over-represented in older adults and under-representative of the extremes of the income and education distribution. To adjust for possible non-representativeness, internet panel providers use "click-balancing." Internet panel members are surveyed at regular intervals regarding their basic demographic (and many other) characteristics. The practice of "click-balancing" is used to ensure that the "inbound" sample is representative by establishing quotas. For example, if census data establishes that the US adult population is 51 per cent female and 49 per cent male, then internet provider establishes quotas of male and female respondents. Once over quota, the internet provider rejects potential respondents. Typically, click-balancing is only used to impose quotas for age, sex, and region, even though many internet providers have a wealth of other information which can be used to implement click-balancing.

Statisticians will recognize this approach as quota sampling. Quota sampling cannot establish representativeness unless the quantities that are measured by the survey are highly correlated with whatever variables are used to balance. If I click-balance only on age and gender, my conjoint demand estimates could be very nonrepresentative unless product and attribute preferences are highly correlated with age and gender. This is unlikely in any real world application.

Our view is that one should measure a set of demographic variables that are most likely to be related to preference but also to measure a set of behavioral correlates. For example, we might want to include ownership of current make and model cars to establish a rough correspondence between our inbound sample and the overall market share by make or type of car available from sources such as JD Power. We might also look at residence type, family composition, and recreational activities as potential behavioral correlates for the SUV prototype sample. We could use our conjoint survey constrained to only existing products to simulate market shares which should be similar to the actual market shares for the set of products in the simulation.

In short, we recognize that, for some general consumer products, probability samples are difficult to implement and that we must resort to the use of internet panels. However, we do not believe that click-balancing on a handful of demographic variables is sufficient to assert representativeness.

### 4.4 Glossary

The glossary portion of a survey introduces the respondent to the product features and levels included in the choice tasks. It is important that product attributes are described factually, in simple terms, and not in terms of the benefits that might accrue to a consumer. Doing so would possibly educate the respondent about possible uses of the product that were not previously known and threaten the validity of the study. For example, a conjoint study of digital point-and-shoot camera might include the attribute "WiFi enabled." Benefits associated with this attribute include easy downloading and storage of pictures, and the ability to quickly post photos on social media platforms such as Facebook or Instagram. However, not all respondents in a conjoint survey may use social media, and may not make the connection between the attribute "WiFi enabled" and benefits that accrue from its use. Moreover, there are many potential benefits of posting photos on social media, including telling others that you are safe, that you care about them, or that the photograph represents something you value. Including such a message is problematic for a conjoint study because the attribute description extends beyond a description of the product and introduces an instrumentation bias (Campbell and Stanley, 1963) into the survey instrument.

The utility that respondents have for the attributes and features of a product depends on their level of knowledge of the attributes and how they can be useful to them. In some cases, firms may anticipate an advertising campaign to inform and educate consumers about the advantage of a particular attribute, and may want to include information on the benefits that can accrue from its use. However, incorporating such analysis into a conjoint study is problematic because it assumes that consumer attention and learning in the study is calibrated to levels of attention and learning in the marketplace.

A challenge in constructing an effective glossary is getting respondents to understand differences among the levels of an attribute being investigated. A study of

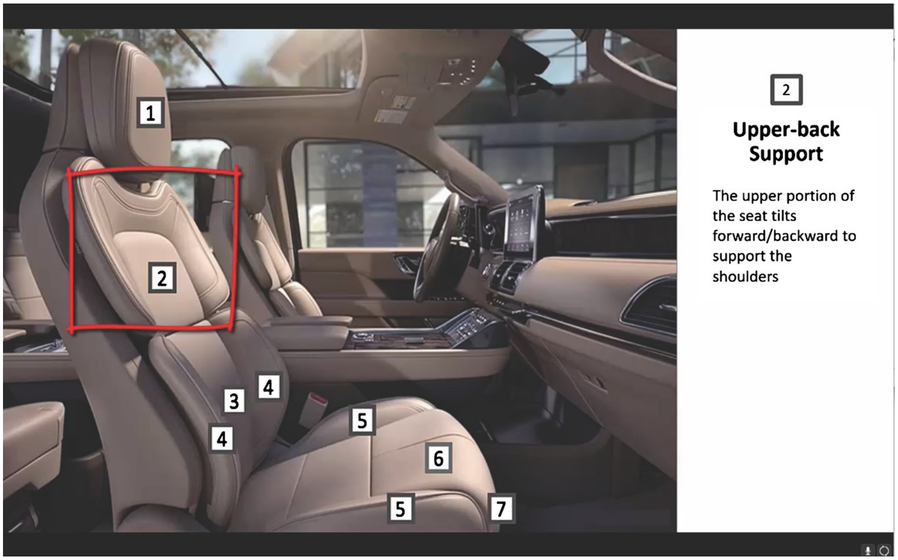


**FIGURE 1**

Screenshot of luxury car seat glossary.

automotive luxury car seats, for example, may include upgrades such as power head restraints, upper-back support, and independent thigh support (Kim et al., 2017). Respondents need to pay careful attention to the glossary to understand exactly how these product features work and avoid substituting their own definitions. An effective method of accomplishing this is to produce a video in which the attributes are defined, and requiring respondents to watch the video before proceeding to the choice task. A screenshot explaining upper-back support is provided in Fig. 1.

## 4.5 Choice tasks

The simplest case of conjoint analysis involves just two product features – brand and price – because every marketplace transaction involves these features. A brand-price analysis displays an array of choice alternatives (e.g., varieties of yogurt, different 12-packs of beer) and prices, and asks respondents to select the alternative that is most preferred. The purpose of a brand-price conjoint study is to understand the strength of brand preferences and its relationship to prices. That is, a brand-price analysis allows for the share prediction of existing offerings at different prices. It should be emphasized that a conjoint design with only brand and price is likely to produce unrealistic results as there are few products that can be characterized only two features, however important.

The inclusion of non-price attributes to a conjoint study allows analysis to expand beyond brands and prices. An example of a choice task to study features of digit cameras is provided in Fig. 2 (Allenby et al., 2014a,b). Product attributes are listed

| Scenario 1 of 16 | Camera 1 | Camera 2 | Camera 3 | Camera 4 | |
|---|---|---|---|---|---|
| Brand | Canon Powershot | Panasonic Lumix | Sony Cyber-shot | Nikon COOLPIX | |
| Megapixels | 16 | 16 | 16 | 16 | |
| Optical Zoom | 10x | 4x | 10x | 4x | |
| Video | Full HD Video (1080p) with Stereo Microphone | HD Video (720p) | HD Video (720p) | Full HD Video (1080p) with Stereo Microphone | None of these |
| Swivel Screen | No | Yes | No | Yes | |
| Wifi | Yes | Yes | No | Yes | |
| Price | $79 | $279 | $379 | $179 | |
| Which of these digital cameras do you prefer? | | | | | |

**FIGURE 2**

Example choice task.

on the left side of the figure, and attribute levels are provided in the cells of the feature grid. Brand and price are present along with the other product features. Also included on the right side of the grid is the 'no-choice' option, where respondents can indicate they would purchase none of the products described.

The choice task illustrated in Fig. 2 illustrates a number of aspects of conjoint analysis. First, the choice task does not need to include all brands and offerings available in the marketplace. Just four brands of digital cameras are included in the digital camera choice task, with the remaining cameras available for purchase represented by the no-choice option. Second, the choice task also doesn't need to include all the product features that are present in offerings. The brand name serves as a proxy for the unmentioned attributes of a brand in a conjoint study, and consumer knowledge of these features is what gives the brand name its value. For example, in a brand-price conjoint study, the brand name 'Budweiser' stands for a large number of taste attributes that would be difficult to enumerate in a glossary.

Researchers have found that breaking the conjoint response into two parts results in more accurate predictions of market shares and expected demand (Brazell et al., 2006). The two-part, or dual response is illustrated in Fig. 3. The first part of the response asks the respondent to indicate their preferred choice option and the second part asks if the respondent would really purchase their most preferred option. The advantage of this two-part response is that it slows down the respondent so that they think through the purchase task, and results in a higher likelihood of a respondent selecting the no-choice option. Research has shown that this two-part response leads to more realistic predictions of market shares.

Survey respondents are asked to express their preference across multiple choice tasks in a conjoint study. Statistical experimental design principles are used to make sure that the attribute levels are rotated across the choice options so that the part-worths can be estimated from the data. Statistical experimental design principles are used to design the choice tasks so that the data are informative about the part-worths (Box et al., 1978).

| Scenario 1 of 16 | Camera 1 | Camera 2 | Camera 3 | Camera 4 |
|---|---|---|---|---|
| Brand | Canon Powershot | Panasonic Lumix | Sony Cyber-shot | Nikon COOLPIX |
| Megapixels | 16 | 16 | 16 | 16 |
| Optical Zoom | 10x | 4x | 10x | 4x |
| Video | Full HD Video (1080p) with Stereo Microphone | HD Video (720p) | HD Video (720p) | Full HD Video (1080p) with Stereo Microphone |
| Swivel Screen | No | Yes | No | Yes |
| Wifi | Yes | Yes | No | Yes |
| Price | $79 | $279 | $379 | $179 |
| Which of these digital cameras do you prefer? | | | | |
| If the camera you selected were being offered would you actually purchase it? | Yes<br>No | | | |

**FIGURE 3**

Dual-response choice task.

The economic foundation of conjoint analysis rests on the assumption that consumers have well-defined preferences for offerings that are recalled when responding to the choice task. That is, consumer utility is not constructed based on the choice set, but is recalled from memory. This assumption is contested in the psychological choice literature (Lichtenstein and Slovic, 2006) where various framing effects have been documented (Roe et al., 2001). However, by screening respondents for inclusion in a conjoint study, the effects of behavioral artifacts of choice are largely reduced. Screened respondents, who are familiar with the product category and understand the product features, are more likely to have well-developed preferences and are less likely to construct their preferences at the time of decision. We discuss this issue again below when discussing the robustness of conjoint results.

As discussed earlier, conjoint analysis can be conducted for decisions that involve volumetric purchases using utility functions that allow for the purchase of multiple goods. The responses can be non-zero for multiple choice alternatives corresponding to an interior solution to a constrained maximization problem. Fig. 4 provides an illustration of a volumetric choice task for the brand-price conjoint study discussed by Howell et al. (2015).

## 4.6 Timing data

While samples based on internet panels are not necessarily representative, the internet format of survey research provides measurement capabilities that can be used to determine the validity of the data. In addition to timing the entire survey, researchers can measure the time spent reading and absorbing the glossary and on each of the choice tasks. This information can and should be gathered in the pre-test stage as well as the fielding of the final survey. The questionnaire can be reformulated if there are pervasive problems with attention or "speeding." Sensitivity analyses with respect


**FIGURE 4**

Volumetric choice task.

to inclusion of respondents who appear to be giving the survey little attention are vital to establishing the credibility of results based on a conjoint survey which has been observed by many practitioners to be viewed by at least some respondents as tedious. Clearly, there is a limit to the amount of effort any one respondent is willing to devote to even the most well-crafted conjoint survey and conjoint researchers should bear this in mind before designing unnecessarily complex or difficult conjoint tasks.

## 4.7 Sample size

For a simple estimator such as a sample mean or sample proportion, it is a relatively simple matter to undertake sample size computations. That is, with some idea of the variance of observations, sample sizes sufficient to reduce sampling error or posterior uncertainty below a specific margin can be determined. Typically, for estimation of sample proportions, sample sizes of 300-500 are considered adequate.

However, a conjoint survey based on a valid economic formulation is designed to allow for inference regarding market demand and various other functions of demand such as equilibrium prices. In these contexts, there are no rules of thumb that can easily be applied to establish adequate sample sizes. In a full Bayesian analysis, the posterior distribution of any summary of the conjoint data, however complicated, can easily be constructed using the draws from the predictive posterior distribution of the respondent level parameters (Eq. (19)). This posterior distribution can be used to assess the reliability or statistical information in a given conjoint sample. However, analytical expressions for these quantities are not typically not available. All we know is that the posterior distribution tightens (at least asymptotically) at the rate of the square root of the number of respondents. This does not help us plan sample sizes, in advance, for any specific function of demand parameters. The only solution to this problem is to perform a pilot study and scale the sample off of this study in such a way as to assure a given margin of error or posterior interval. This can be done by assuming that the posterior standard error will tighten at rate $\sqrt{N}$.

Our experience with equilibrium pricing applications is that sample sizes considerably larger than what is often viewed by conjoint practitioner as adequate are required. Many conjoint practitioners assume that a sample size of 500-1000 with a 10-12 conjoint tasks will be adequate. This may be on the low side for more demanding computations.

We hasten to add that many conjoint practitioners do not report any measures of statistical reliability for the quantities that they estimate. Given the ease of constructing the posterior predictive distribution of any quantity of interest, there is no excuse for failure to report a measure of uncertainty.

# 5 Practices that compromise statistical and economic validity

While conjoint originated as a method to estimate customer preferences or utility, many practitioners of conjoint have created alternative methodologies which either invalidate statistical inference or compromise the economic interpretation of conjoint results. As long as there are well-formulated and practical alternatives, it is our view that there is no excuse for using a method that is not statistically valid. Whether or not conjoint methods should be selected so that the results can be interpreted as measuring a valid economic model of demand is more controversial. A pure predictivist point of view is that any procedure is valid for predicting demand as long as it predicts well. This point of view opens the range of possible conjoint specifications to any specification that can predict well. Ultimately, conjoint is most useful in predicting or simulating demand in market configurations which differ from that observed in the real world. Here we are expressing our faith that specifications derived from valid utility structures will ultimately prevail in a prediction context where the state of the world is very different from that which currently prevails.

## 5.1 Statistical validity

There are two threats to the statistical validity of conjoint analyses: 1) the use of estimation methods which have not been shown to provide consistent estimators and 2) improper methods for imposing constraints on conjoint part-worths.

### 5.1.1 Consistency

A consistent estimator is an estimator whose sampling distribution collapses on the true parameter values as the sample size grows infinitely large. Clearly, this is a very minimal property. That is, there are consistent but very inefficient estimators. The important point is that when choosing an estimation procedure, we should only select from the set of procedures that are consistent. The only reason one might resort to using a procedure whose consistency has not been established is if there are no practical consistent alternatives. Even then, our view is that any analysis based on the procedure for which consistency cannot be verified is that this must be termed tentative at best.

However, in conjoint, we do not have to resort to the use of unverified procedures since all Bayes procedures are consistent unless the prior is dogmatic in the sense of putting zero probability on a non-trivial portion of the parameter space. Moreover, Bayes procedures are admissible which means that it will be very difficult to find a procedure which dominates Bayes methods in either estimation or prediction (Bernardo and Smith, 2000).

In spite of these arguments in favor of only using methods for which consistency can be demonstrated, there has been work in marketing on conjoint estimators (see Toubia et al., 2004) that propose estimators based on minimization of some criterion function. It is clear that, for a given criterion function, there is a minimum which will be an estimator. However, what is required to establish consistency is a proof that the criterion function used to derive the estimator converges (as the sample size grows to infinity) to a function with a minimum at the true parameter value. This cannot be established by sampling experiments alone as this convergence is required over the entire parameter space. The fact that an estimator works "well" in a few examples in finite samples does not mean that the estimator has good finite sample or asymptotic properties.

We should note that there are two dimensions of conjoint panel data ($N$ – number of respondents and $T$ – the number of choice tasks). We are using large $N$ and fixed $T$ asymptotics in the definition of consistency. Estimates of respondent-level part-worths are likely to be unreliable as they are based only on a handful of observations and an informative prior, and cannot be shown to be consistent if $T$ is fixed.

A common practice in conjoint analysis is to use respondent-level coefficients to predict the effects of changes to product attributes and price. An advantage of Bayesian methods of estimating conjoint models is its ability to provide individual-level estimates of part-worths ($\theta_h$) in Eq. (16) in addition to the hyper-parameters that describe the distribution of part-worths ($\tau$). Using the individual-level estimates for predicting the effects of product changes on sales, however, is problematic because of the shallow nature of the data used in conjoint analysis. Respondents provide at most about 16 responses to the choice tasks before becoming fatigued, and while these estimates may be consistent in theory, they are not consistent in practice because of data limitations.

The effect of using individual-level estimates is to under-state the confidence associated with predicted effects. That is, the confidence intervals are too large when using the individual-level estimates. This is because uncertainty in the individual-level estimates is due to two factors – uncertainty in the hyper-parameters and uncertainty arising from the individual-level data. As the sample size increases in a conjoint study, it is only possible to increase the number of respondents $N$, not the number of observations per respondent $T$. As a result, the individual-level estimates will always reflect a large degree of uncertainty, even when the hyper-parameters are accurately estimated. The accurate measurement of uncertainty in predictions from conjoint analysis must be based on model hyper-parameters as shown in Eq. (21).

### 5.1.2 Using improper procedures to impose constraints on partworths

The data used in conjoint analysis can be characterized as shallow in the sense that there are few observations per respondent. Individual-level parameters ($\theta_h$) can therefore be imprecisely estimated and can violate standard economic assumptions. Price coefficients, for example, are expected to be negative in that people should want to pay less for a good than more, and attribute-levels may correspond to an ordering where consumers should want more of a feature or attribute holding all else constant.

There are two approaches for introducing this prior information into the analysis. The first is to reparameterize the likelihood so that algebraic restrictions on coefficients are enforced. For example, the price coefficient in Eq. (5) can be forced to be negative through the transformation:

$$\psi_z = -\exp(\beta_p)$$

and estimating $\beta_p$ unrestricted. Alternatively, constraints can be introduced through the prior distribution as discussed by Allenby et al. (1995). Sign constraints as well as monotonicity can be imposed automatically using our R package, **bayesm**.

It is a common practice in conjoint analysis to impose sign restrictions by simply zeroing out the estimates or to use various "tying" schemes for ordinal constraints in which estimates are arbitrarily set equal to other estimates to enforce the ordinal constraints. This is an incoherent practice which violates Bayes theorem and, therefore, removes the desirable properties of the Bayes procedures.

## 5.2 Economic validity

There are four threats to the economic validity of conjoint analyses:

1. Using conjoint specifications contrary to valid utility or indirect utility functions.
2. Various sorts of self-explicated conjoint which violate utility theory.
3. Comparison of part-worths across respondents.
4. Attempts to combine conjoint and rankings data.

### 5.2.1 Non-economic conjoint specifications

We have emphasized that conjoint specifications should be derived from a valid direct utility function (see (1) or (7)). For discrete-choice conjoint, the respondent choice probabilities are a standard logit function of a linear index where prices enter in levels not in logs. It is common for practitioners to enter prices as a sequence of dummy variables, each for a different level of price used in the conjoint design. The common explanation is that the dummy variable specification is "non-parametric." This approach is not only difficult to reconcile with economic theory but also opens the investigator for a host of violations of the reasonable economic assumption that indirect utility is monotone in price (not to mention convex). In general, utility theory imposes a certain discipline on investigators to derive the empirical specification from a valid utility function.

In the volumetric context, the problems are even worse as the conjoint specifications are often completely ad hoc. This ad-hocery arises, in part, from the lack of software to implement an economically valid volumetric conjoint – a state of affairs we trying to remedy by our new R package, **echoice**.[5]

### 5.2.2 Self-explicated conjoint

Some researchers advocate using self-explicated methods of estimating part-worths (see, for example, Srinivasan and Park, 1997; Netzer and Srinivasan, 2011). In some forms of self-explicated conjoint, respondents are asked to rate the relative importance of a product feature on some sort of integer valued scale, typically 5 or 7 points. In other versions of self-explicated conjoint, separate measures of the relative importance and desirability of product attributes are combined in an arbitrary way to form an estimate of a part-worth. There are many ways in which these procedures violate both economic and statistical principles. Outside of the demand context (as in choice-based or volumetric conjoint), there is no meaning to the "importance" or "desirability" of features. The whole point of a demand experiment is to infer a valid utility function from demand responses. No one knows what either "importance" or "desirability" means, including the respondents. The scales used are only ordinal and therefore cannot be converted to a utility scale which is an interval scale. In short, there is no way to transform or convert relative importance or desirability to a valid part-worth. Finally, as there is no likelihood function (or error term) in these models, it is impossible to analyze the statistical properties of self-explicated procedures.

### 5.2.3 Comparing raw part-worths across respondents

Conjoint part-worths provide a measure of the marginal utility associated with changes in the levels of product attributes. The utility measure obtained from a conjoint analysis allows for the relative assessment of changes in the product attribute-levels, including price. However, utility is not a measure that is comparable across respondents because it is only intended to reflect the preference ordering of a respondent, and a preference ordering can be reflected by any monotonic transformation of the utility scale. That is, the increase or decrease in utility associated with changes in the attribute-levels are person-specific, and cannot be used to make statements that one respondent values changes in the levels of an attribute more than another.

[5] Development version available at https://bitbucket.org/ninohardt/echoice/.

Making utility comparisons across respondents requires the monetization of utility to express the value of changes in the levels of an attribute on a scale common to all respondents. While the pain or gain of changes to a product attribute is not comparable across people, the amount they would be willing to pay is comparable across respondents. The WTP, WTB, and EPP measures discussed above provide a coherent metric for summarizing the results of conjoint analysis across respondents.

### 5.2.4 Combining conjoint with other data

One could argue that there are covariates predictive of preferences that can be measured outside the conjoint portion of the survey. The proper way to combine this data with conjoint choice or demand data is to employ a hierarchical model specification in which individual-level parameters are related to each other in the upper-level of the model hierarchy. The upper-level model in conjoint analysis typically is used to describe cross-sectional variation in part-worth estimates using these covariates to model observed heterogeneity. Individual-level coefficients from other models, calibrated on other datasets, could be used as variables to describe the cross-sectional variation of part-worths (Dotson et al., 2018). Combining data in this way automatically weights the datasets and can improve the precision of the part-worth estimates.

A disturbing practice is the combination of conjoint choice data with Max-Diff rankings data. Practitioners are well aware that conjoint surveys are not popular with respondents but that a standard Max-Diff exercise is found to be very easy for respondents. Max-Diff is a procedure to rank (by relative importance) any set of attributes or products. The Max-Diff procedure breaks the task of ranking all in the set down into a sequence of smaller and more manageable tasks which consist of picking the most and least "important" from a small set of alternatives. A logit-style model can be used to analyze Max-Diff data to provide a ranking for each respondent. It is vital to understand that rankings are not utility weights and rankings only have ordinal properties. The exact coefficients used to implement the ranking are irrelevant and have no value. That is to say, the ranking of 3 things represented by (1, 3, 2) is the same as (10, 21, 11). There is no meaning to the intervals separating values or to ratios of values. This is true even setting aside the thorny question of what "importance" means. There are no trade-offs in Max-Diff analysis, so there is no way to predict choice or demand behavior on the basis of the Max-Diff derived rankings. Unfortunately, some practitioners have taken to scaling Max-Diff ranking coefficients and interpreting these scaled coefficients as part-worths. They are not. There is no way of combining Max-Diff and conjoint data in a coherent fashion. The only way to do so is the regard "importance" as the same as utility and to use the Max-Diff results as the basis of an informative prior used in analysis of conjoint data.

## 6 Comparing conjoint and transaction data

The purpose of conducting a conjoint analysis is to predict changes in demand for changes in a product's configuration and its price. Conjoint analysis provides an experimental setting for exploring these changes when revealed preference data from the marketplace lacks sufficient variation for making inferences and predictions. A natural question to ask is the degree to which conjoint analysis provides reliable estimates of changes that would occur. In this chapter we have argued that there are many requirements of conducting a valid conjoint study, beginning with the use of a valid economic model for conducting inference and the use of rigorous methods in data collection to make sure the respondent is qualified to provide answers and understands the choice task.

We investigate the consistency of preferences and implications between conjoint and transaction data using a dataset containing marketplace transactions and conjoint responses from the same panelists. Frequent buyers in the frozen pizza product category were recruited to provide responses to a conjoint study in which frozen pizza attribute-levels changed. The fact that participants were frequent purchasers in the product category made them ideal respondents for the conjoint survey in the sense that they are known to be prospects who were well acquainted with the category.

Most attempts to reconcile results from stated and revealed preferences data have tended to focus on aggregate predictors of demand and market shares. Examples range from the study of vegetables (Dickie et al., 1987) and grocery items (Burke et al., 1992; Ranjan et al., 2017) that are frequently purchased, to infrequently purchased goods such as automobiles (Brownstone et al., 2000). Lancsar and Swait (2014) provide an overview of studies across different disciplines. We start by investigating the consistency of estimated parameters, including estimates of marginal utility. We then assess the extent to which market demand predictions can be made using conjoint-based estimates. Finally, we show estimates of measures of economic value.

## 6.1 Preference estimates

The conjoint study choice task involved six frozen pizza offerings and included a ‘no-choice’ option. The transaction data comprised 103 unique product offerings and included attributes such as brand name (e.g., DiGiorno, Red Baron, and Tombstone), crust (i.e., thin, traditional, stuffed, rising), and toppings (e.g., pepperoni, cheese, supreme). The volumetric demand model arising from non-linear utility (Eq. (7)) was used to estimate the model parameters because pizza purchases typically involve the purchase of multiple units and varieties. Table 1 provides a list of product attributes.

The attributes and attribute-levels describing the 103 UPCs were used to design the conjoint study. Among the 297 households in an initial transaction dataset, 181 households responded to a volumetric conjoint experiment and had more than 5 transactions in the 2-year period. Qualifying respondents in this way ensures that they are knowledgeable about the category and typical attribute levels. In each of the 12 choice tasks, respondents choose how many units of each of the 6 product alternatives they would purchase the next time they are looking to buy frozen pizza. A sample choice

**Table 1** Attributes.

| Brand | Size for | Crust | Topping type | Topping spread | Cheese |
|---|---|---|---|---|---|
| DiGiorno | One | Thin | Pepperoni | Moderate | No claim |
| Frescetta (Fr) | Two (FT) | Traditional (TC) | Cheese (C) | Dense (DT) | Real (RC) |
| Red Baron (RB) | | Stuffed (SC) | Vegetarian (V) | | |
| Private Label (Pr) | | Rising (RC) | Surpreme (Sr) | | |
| Tombstone (Tm) | | | PepSauHam (PS) | | |
| Tony's (Tn) | | | Hawaii (HI) | | |

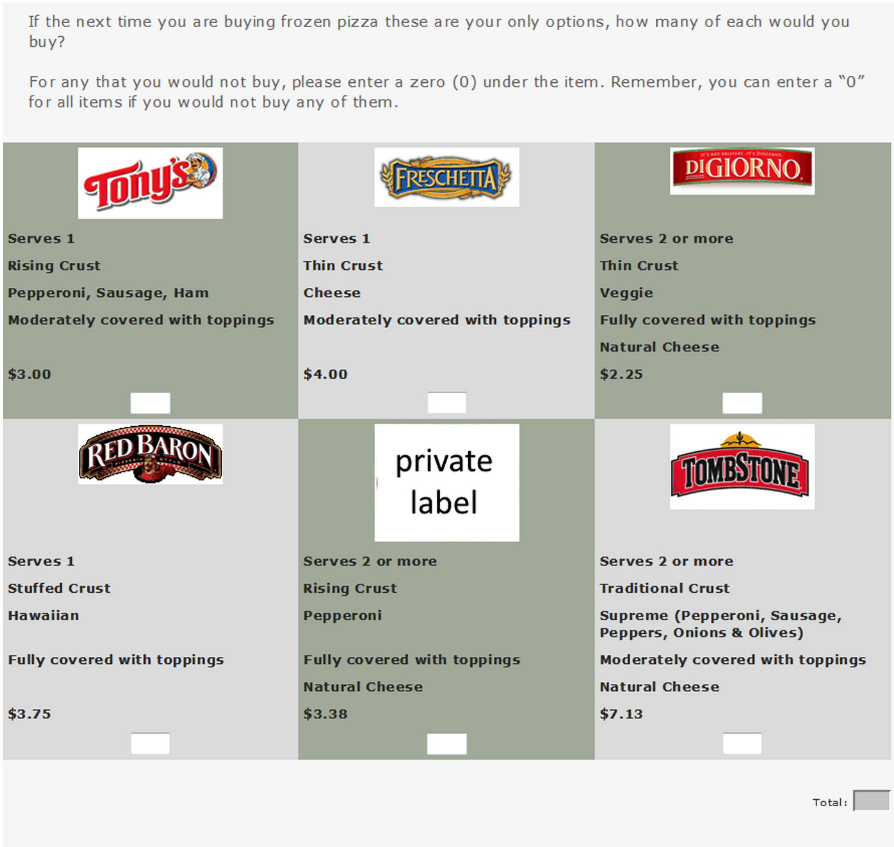


**FIGURE 5**

Choice task.

task is shown in Fig. 5. Price levels were chosen in collaboration with the sponsoring company to mimic the actual price range.

**Table 2** Estimated parameters (volumetric model). Mean of random-effects distribution $\bar{\theta}$.

| | | Conjoint | Transaction |
|---|---|---|---|
| | $\beta_0$ | **−2.91** (0.11) | **−4.66** (0.18) |
| Brand | Frescetta | **−0.35** (0.09) | −0.19 (0.11) |
| | Red Baron | **−0.66** (0.10) | **−0.38** (0.12) |
| | Private Label | **−0.72** (0.10) | **−0.85** (0.15) |
| | Tombstone | **−0.80** (0.11) | **−0.63** (0.16) |
| | Tony's | **−1.29** (0.13) | **−2.05** (0.17) |
| Size | Serves two | **0.64** (0.07) | **1.22** (0.08) |
| Crust | Traditional | 0.11 (0.07) | **0.31** (0.07) |
| | Stuffed | −0.04 (0.08) | −0.04 (0.15) |
| | Rising | 0.07 (0.08) | **0.26** (0.07) |
| Topping Type | Cheese | **−0.40** (0.12) | **−0.57** (0.09) |
| | Vegetarian | **−1.12** (0.16) | **−0.96** (0.14) |
| | Surpreme | **−0.30** (0.11) | **−0.39** (0.09) |
| | PepSauHam | −0.13 (0.09) | **−0.22** (0.08) |
| | Hawaii | **−1.14** (0.15) | **−0.99** (0.21) |
| Topping | Dense | 0.06 (0.05) | −0.02 (0.06) |
| Cheese | Real | 0.10 (0.05) | −0.01 (0.10) |
| | $\ln\gamma$ | **−0.50** (0.08) | **−2.07** (0.09) |
| | $\ln E$ | **3.57** (0.07) | **2.89** (0.06) |
| | $\ln\sigma$ | **−0.57** (0.05) | **−0.87** (0.05) |

*Boldfaced parameters signify that the 95% posterior credible interval of the estimate does not include zero. Standard Deviations printed in parentheses.*

We use dummy coding in our model estimation, where the first level of each attribute in Table 1 are the reference levels. The vector of 'part-worths' $\beta$ includes a baseline coefficient $\beta_0$, which represents the value of an inside good relative to the outside good. The remaining elements of $\beta$ refer to the dummy coefficients. We use the volumetric demand model described in Section 2.2. Individual-level parameters are given by the vector of 'part-worths' $\beta$, the rate of satiation of inside goods $\gamma$, the alloted budget $E$, and the scale of the error term $\sigma$. The latter three parameters are log-transformed to ensure positivity. A multivariate normal distribution of heterogeneity is assumed with default (diffuse) priors. Parameter estimates for the conjoint and transaction data are provided in Table 2. The left side of the table reports estimates from the conjoint data, and estimates from the transaction data are displayed on the right side of the table. All estimates are based on models with Type 1 extreme value error terms.

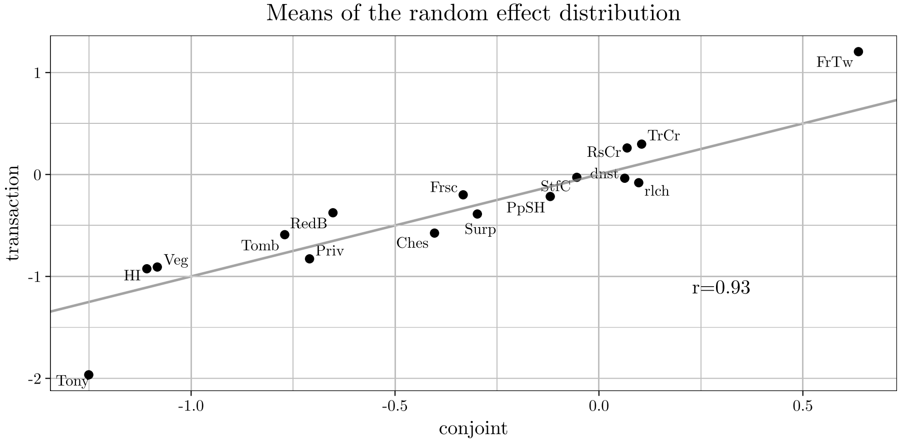


**FIGURE 6**

Comparison of part-worths (volumetric model).

We find general agreement among the estimates except for the inside good intercept ($\beta_0$), the estimated rate of satiation ($\gamma$), the budgetary allotment ($E$), and the scale of the error terms ($\sigma$). That is, the relative value of the brand names and product attribute-levels are approximately the same when estimated with either conjoint or transaction data. Fig. 6 provides a plot of the mean of the random-effects distribution for the two datasets. Average part-worth estimates are plotted close to the 45 degree line indicating that the estimates are similar.

There are a number of potential reasons for the difference in estimates of the brand intercept ($\beta_0$), satiation parameter ($\gamma$), and the budgetary allotment ($E$). Respondents in the conjoint task were told to consider their choices the next time they were buying frozen pizza, while data from the transaction data conditioned on some purchase in the frozen pizza category. Most respondents in the dataset purchased frozen pizza less than 10 times over a two year period, and including data in the analysis from shopping occasions in which they did not make a frozen pizza purchase implicitly assumes that they were in the market for pizza on each and every occasion. We therefore excluded shopping occasions from the transaction data in which pizza wasn't purchased so that it would mimic the conjoint data where respondents were told to consider their next purchase occasion. There is no way of knowing for certain when shoppers considered making a frozen pizza purchase, but did not, in the revealed preference (transaction) data and this discrepancy is partly responsible for the difference in the brand intercept ($\beta_0$) estimates.

Conjoint data cannot account for variation in the context of purchase and consumption, and this limitation may lead to differences in estimates of satiation ($\gamma$) and budgetary allotment ($E$). For example, frozen pizza may occasionally be purchased for social gatherings, which may not be taken into account when providing conjoint responses, resulting in an estimate of the budgetary allotment that is too high for the

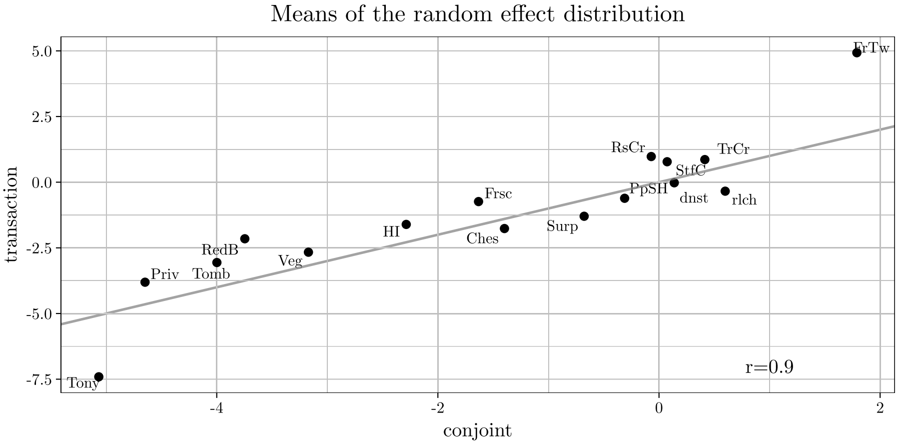


**FIGURE 7**

Comparison of monetized part-worths (discrete choice model).

typical conjoint transaction. The over-estimation of $E$ may also affect the estimate of the rate of satiation. Another consequence of the hypothetical nature of conjoint tasks is that respondents may apply a larger budget when making allocations because they aren't actually spending their own money. We find that the conjoint data reflect lesser satiation and a greater budgetary data than that found in revealed preference data.

We also investigate a discrete choice approximation to the demand data by 'exploding' the volumetric data to represent a series of discrete choices. When $q$ units of a good are purchased, it is interpreted as $q$ transactions with a discrete choice for that good. This allows for the estimation of a discrete choice model on the volumetric conjoint and transaction data. However, this practice results in no consumption of an 'outside good' in the transaction data because only nonzero quantities are observed. We estimate a hierarchical Logit model with a multivariate normal distribution of heterogeneity and relatively diffuse priors. The price coefficient $\beta_p$ is re-parameterized to ensure negativity of the price coefficient.

Estimated coefficients are shown in Table 3. Monetized estimates of the part-worths, obtained by dividing the part-worth by the price coefficient, are compared in Fig. 7, where we use the corresponding means of the random effects distribution. We find close agreement of part-worth estimates from the transaction and conjoint data.

## 6.2 Marketplace predictions

We next compare marketplace predictions from conjoint and transaction data by aggregating across the 103 UPCs to obtain brand-level demand estimates. Parameter estimates from the volumetric conjoint and transaction data are used to produce two

**Table 3** Estimated parameters (discrete choice model). Mean of random-effects distribution $\bar{\theta}$.

| | | Conjoint | Transaction |
|---|---|---|---|
| | $\beta_0$ | **3.67** (0.31) | |
| Brand | Frescetta | **−0.25** (0.11) | **−0.58** (0.23) |
| | Red Baron | **−0.60** (0.11) | **−1.61** (0.27) |
| | Private Label | **−0.73** (0.11) | **−2.90** (0.38) |
| | Tombstone | **−0.60** (0.11) | **−2.22** (0.43) |
| | Tony's | **−0.80** (0.13) | **−5.45** (0.47) |
| Size | Serves two | **0.28** (0.07) | **3.63** (0.30) |
| Crust | Traditional | 0.04 (0.08) | **0.59** (0.15) |
| | Stuffed | −0.01 (0.08) | **0.45** (0.25) |
| | Rising | −0.03 (0.08) | **0.71** (0.16) |
| Topping type | Cheese | −0.22 (0.12) | **−1.33** (0.21) |
| | Vegetarian | **−0.52** (0.13) | **−2.09** (0.35) |
| | Surpreme | −0.12 (0.11) | **−0.96** (0.24) |
| | PepSauHam | −0.06 (0.09) | **−0.46** (0.19) |
| | Hawaii | **−0.38** (0.13) | **−1.60** (0.24) |
| Topping | Dense | 0.02 (0.06) | −0.03 (0.11) |
| Cheese | Real | 0.09 (0.06) | −0.16 (0.25) |
| | $\ln \beta_p$ | **−1.93** (0.20) | **−0.33** (0.10) |

*Boldfaced parameters signify that the 95% posterior credible interval of the estimate does not include zero. Standard Deviations printed in parentheses.*

forecasts for each brand that are displayed in Fig. 8 for the volumetric demand model and Fig. 9 for the logit model with 'exploded' choices. We find that the demand curves are roughly parallel in Fig. 8 with predictions from the conjoint data consistently higher than that based on the transaction data. The reason for the shift is due to differences in the estimate of the brand coefficient ($\beta_0$) which we attribute to differences in the treatment of the 'no-choice' option. That is, the smaller estimated brand intercept in the transaction data results in lower estimates of demand.

While the level of demand is estimated to be different in Fig. 8, changes in demand for changes in price and other product attributes is approximately the same because the demand curves are roughly parallel to each other and because the part-worth coefficients enter the Kuhn-Tucker conditions in Eq. (12) linearly. The consistency of the estimates from conjoint and transaction data is observed more readily observed when we convert the volumetric predictions to shares in Fig. 9.

It is useful to remember that the purpose of conjoint analysis is to predict changes in marketplace demand as a result of changes in the formulation of marketplace of-

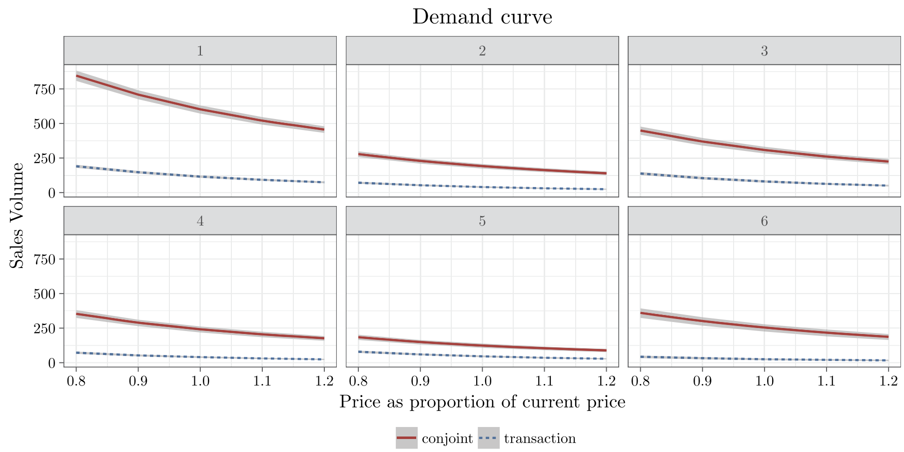


**FIGURE 8**

Comparison of predictions (volumetric model).

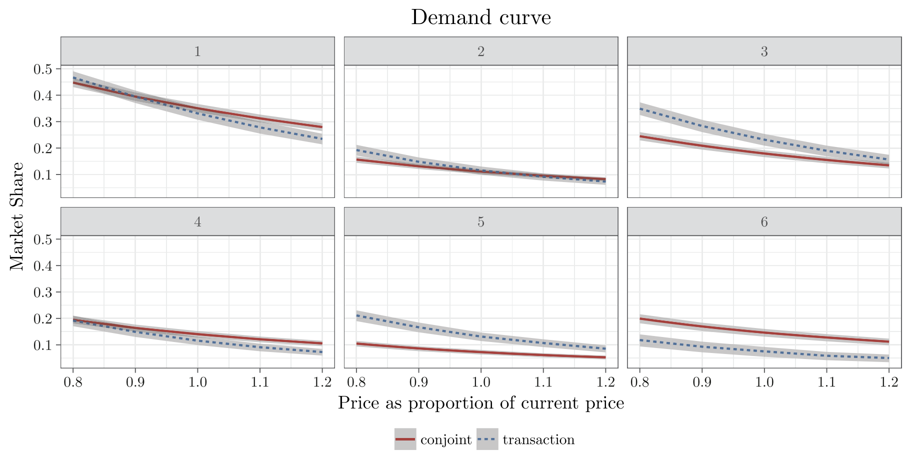


**FIGURE 9**

Comparison of predictions converted to shares (volumetric model).

ferings. Even though the aggregate demand curves displayed in Figs. 8 and 9 are seen to be vertically translated, predictions of the change in volume and change in share are closer to each other.

## 6.3 Comparison of willingness-to-pay (WTP)

We compute the consumer willingness-to-pay (WTP) for the family-sized (‘for two’) attribute of a DiGiorno pepperoni pizza with rising crust and a real cheese claimed

**Table 4** Willingness-to-pay estimates (volumetric and logit), 'for-two' attribute.

| | | WTP | | | p-WTP | |
|---|---|---|---|---|---|---|
| | | Conjoint | Transaction | With $\beta_{0\text{trans}}$ | Conjoint | Transaction |
| Logit | mean | 0.034 | 0.027 | | 0.782 | 3.315 |
| | median | 0.029 | 0.026 | | 0.776 | 3.278 |
| | perc25 | −0.009 | 0.021 | | 0.704 | 3.185 |
| | perc75 | 0.074 | 0.031 | | 0.859 | 3.430 |
| Volumetric | mean | 0.113 | 0.013 | 0.023[a] | | |
| | median | 0.070 | 0.002 | 0.009[a] | | |
| | perc25 | 0.020 | 0.000 | 0.001[a] | | |
| | perc75 | 0.144 | 0.015 | 0.026[a] | | |

[a] *Volumetric WTP estimates based on conjoint data except for brand intercepts, which are substituted from transaction data estimates.*

attribute. As discussed above, a true WTP measure includes the effects of alternative products that consumers could consider as alternatives if a particular attribute is unavailable, and is not simply a monetary rescaling of the 'for two' part-worth. Ignoring the effect of competitive offerings will over-state the value of product attributes because it ignores the other options consumers have available to them as they make their purchases.

We compare estimates of pseudo willingness-to-pay (p-WTP) based on part-worth monetization (i.e., $\beta_h/\beta_{hp}$), to estimates of willingness-to-pay (WTP) based on compensating valuation calculations (see Eqs. (25) and (26)). Table 4 shows estimates for the Logit and Volumetric demand models, using parameter estimates based on conjoint and transaction data. The top portion of Table 4 reports results for the logit model with 'exploded' data, and the bottom portion of the table pertains to the volumetric demand model. WTP estimates are reported on the left side of the table, and pseudo-WTP estimated are reported on the right side.

We find that the WTP estimates for the logit model are much smaller than p-WTP estimates because of the large number of choice alternatives present in the marketplace. The absence of a 'for-two' DiGiorno pepperoni pizza creates an economic loss that is worth, on average about three cents using the WTP statistic as opposed to either 78 cents or $3.31 based on the p-WTP estimate. The loss in utility is much smaller in the WTP calculation because consumers can recognize that they can purchase a different option to generate their utility and are not constrained to purchase the same good. Moreover, we find that estimates of WTP for the logit model is about three cents using conjoint estimates of the part-worths, and about two cents using the transaction data estimates. These estimates are not statistically different from each other.

WTP estimates based on the volumetric demand model are less consistent, with estimates based on the conjoint data equal to eleven cents versus one cent for the transaction data. This difference is due to differences in the estimated baseline intercept coefficient ($\beta_0$). When the transaction data intercept is substituted for the conjoint intercept, the estimated WTP for the conjoint data reduces from eleven cents to two cents. Thus, overall, we find that estimates based on the conjoint data are slightly higher, but not significantly higher, than those based on the transaction data for both models once the difference is the baseline intercept is aligned.

## 7 Concluding remarks

Conjoint analysis is an indispensable tool for predicting the effects of changes to marketplace offerings when observational data does not exist to inform an analysis. This occurs when predicting sales of new products and product configurations with new attribute-levels and their combinations and when existing products have little price variation. Conjoint data reflects simulated choices among a set of competitive products for which respondents are assumed to have well-defined preferences. Analysis is based on a combination of economic and statistical principles that afford inferences and predictions about the expected sales of new offerings.

We argue in this chapter that a valid conjoint analysis requires a valid model and constructs for inference and valid data. We discuss two economic models for analysis based on random-utility theory for discrete choice and volumetric demand, and discuss alternative measures (WTP, WTB, and EPP) of economic value. We demonstrate that these measures are consistently estimated using either conjoint or transaction data using a conjoint analysis conducted on scanner panelists in the frozen pizza category. Part-worth estimates of product features are shown to be approximately the same, and forecasts of the change in demand for changes in attributes such as price are found to be similar.

Some model parameters, however, are not consistently estimated across stated and revealed preference data. The largest discrepancy involves the baseline brand intercept ($\beta_0$) for the no-choice option, which is difficult to align, because conjoint studies ask respondents to think about the next time they will be making a purchase in the category while revealed preference cannot exactly identify these shopping trips. Shopping trips limited to purchases in the category ignore occasions when shoppers may be in the market but decide not to purchase because prices are not sufficiently low, and the collection of all shopping trips to the store contain instances where consumers are not in the market for items in the category. The difference in the estimated baseline intercept results in a vertical translation of demand curves and heightened measures of economic willingness-to-pay.

This chapter demonstrates that a properly designed conjoint study, using valid economic models to analyze the data, can produce accurate estimates of economic demand in the marketplace. We identify practices that should be avoided, and demonstrate that ad-hoc estimates of value, such as the pseudo willingness-to-pay (p-WTP), provide poor estimates of the economic value of product features. Additional research is needed to better design conjoint studies to obtain valid estimates of budgetary allotments and the rate of satiation of purchase quantities.

## Technical appendix: Computing expected demand for volumetric conjoint

The algorithm described here first determines the optimal amount of the outside good $z_{ht}$, which then allows computing the corresponding inside good quantities $\mathbf{x}_{ht}$.

From Eqs. (8) and (9) we have that:

$$\frac{p_j}{z} = u_j \quad \text{if } x_j > 0 \tag{A.1}$$

$$\frac{p_j}{z} \geq u_j \quad \text{if } x_j = 0 \tag{A.2}$$

At the optimum, $u_i/p_i = u_j/p_j$ for the $R$ goods with non-zero demand. Solving for $x$ yields an equation for optimal demand quantities for the inside goods:

$$x_k = \frac{\psi_k z - p_k}{\gamma p_k} \tag{A.3}$$

Substituting Eq. (A.3) into the budget constraint (2) yields:

$$z = \frac{\gamma + E\sum_{k=1}^{R} p_k}{\gamma + \sum_{k=1}^{R} \psi_k} \qquad \text{if } R > 0 \tag{A.4}$$

$$z = E \qquad \text{if } R = 0 \tag{A.5}$$

Re-arranging (A.1) yields the following for $z$:

$$zs = \frac{p_j}{\psi_j} \quad \text{if } x_j > 0 \tag{A.6}$$

$$zs \leq \frac{p_j}{\psi_j} \quad \text{if } x_j = 0 \tag{A.7}$$

where

$$s = \frac{1}{\gamma x_j + 1}$$

The algorithm needs $R$ iterations to complete. At each step $k$, we compute the corresponding quantity $x_k$ and $z$, as if $R = k$. Then checking Eqs. (A.6) and (A.7) will determine if the breakpoint has been reached. To implement this, let:

$$\rho_i = \frac{p_i}{\psi_i} \quad \text{for } 1 \leq i \leq K \tag{A.8}$$

$$\rho_0 = 0 \tag{A.9}$$

$$\rho_{K+1} = \infty \tag{A.10}$$

and order the values $\rho_i$ in ascending order so that $\rho_i \leq \rho_{i+1}$ for $1 \leq i \leq K$. Then, $z > \rho_k$ implies $z > \rho_i$ for $i \leq k$. At the optimum, $x_i > 0$ for $1 \leq k \leq K$, $x_i = 0$ for

$k < i \leq K$, and $\rho_k < z < \rho_{k+1}$. The algorithm is guaranteed to stop at optimal $z$ and $0 \leq k \leq K$. The steps are as follows:

1. $a \longleftarrow \gamma E, b \longleftarrow \gamma, k \longleftarrow 0$
2. $z \longleftarrow a/b$
3. while $z \leq \rho_k$ or $z > \rho_{k+1}$:
   (a) $k \longleftarrow k + 1$
   (b) $a \longleftarrow a + \rho_k$
   (c) $b \longleftarrow b + \psi_k$
   (d) $z \longleftarrow a/b$

Once the algorithm terminates, we can insert optimal $z$ into Eq. (A.3) to compute the optimal inside good quantities $\mathbf{x}$.

## References


Allenby, Greg M., Arora, Neeraj, Ginter, James L., 1995. Incorporating prior knowledge into the analysis of conjoint studies. Journal of Marketing Research, 152–162.

Allenby, Greg M., Brazell, Jeff, Howell, John R., Rossi, Peter E., 2014a. Valuation of patented product features. The Journal of Law and Economics 57 (3), 629–663.

Allenby, Greg M., Brazell, Jeff D., Howell, John R., Rossi, Peter E., 2014b. Economic valuation of product features. Quantitative Marketing and Economics 12 (4), 421–456.

Allenby, Greg M., Kim, Jaehwan, Rossi, Peter E., 2017. Economic models of choice. In: Handbook of Marketing Decision Models. Springer, pp. 199–222.

Allenby, Greg M., Rossi, Peter E., 1991. Quality perceptions and asymmetric switching between brands. Marketing Science 10 (3), 185–204.

Allenby, Greg M., Rossi, Peter E., 1998. Marketing models of consumer heterogeneity. Journal of Econometrics 89 (1–2), 57–78.

Becker, Gordon M., DeGroot, Morris H., Marschak, Jacob, 1964. Measuring utility by a single-response sequential method. Behavioral Science 9 (3), 226–232.

Bernardo, José M., Smith, Adrian F.M., 2000. Bayesian Theory. Wiley.

Berry, Steven, Levinsohn, James, Pakes, Ariel, 1995. Automobile prices in market equilibrium. Econometrica, 841–890.

Box, George E.P., Hunter, William Gordon, Hunter, J. Stuart, et al., 1978. Statistics for Experimenters. John Wiley and Sons, New York.

Brazell, Jeff D., Diener, Christopher G., Karniouchina, Ekaterina, Moore, William L., Séverin, Válerie, Uldry, Pierre-Francois, 2006. The no-choice option and dual response choice designs. Marketing Letters 17 (4), 255–268.

Brownstone, David, Bunch, David S., Train, Kenneth, 2000. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. Transportation Research Part B: Methodological 34 (5), 315–338.

Burke, Raymond R., Harlam, Bari A., Kahn, Barbara E., Lodish, Leonard M., 1992. Comparing dynamic consumer choice in real and computer-simulated environments. Journal of Consumer Research 19 (1), 71–82.

Campbell, Donald T., Stanley, Julian C., 1963. Experimental and Quasi-Experimental Designs for Research. Houghton Mifflin Company, Boston.

Diamond, Shari Seidman, 2000. Reference guide on survey research. In: Reference Manual on Scientific Evidence, pp. 221–228.

Dickie, Mark, Fisher, Ann, Gerking, Shelby, 1987. Market transactions and hypothetical demand data: a comparative study. Journal of the American Statistical Association 82 (397), 69–75.

Ding, Min, 2007. An incentive-aligned mechanism for conjoint analysis. Journal of Marketing Research 44 (2), 214–223.

Ding, Min, Grewal, Rajdeep, Liechty, John, 2005. Incentive-aligned conjoint analysis. Journal of Marketing Research 42 (1), 67–82.

Dotson, Marc, Büschken, Joachim, Allenby, Greg, 2018. Explaining preference heterogeneity with mixed membership modeling. https://doi.org/10.2139/ssrn.2758644.

Green, Paul E., Rao, Vithala R., 1971. Conjoint measurement for quantifying judgmental data. Journal of Marketing Research, 355–363.

Hardy, Melissa A., 1993. Regression with Dummy Variables, vol. 93. Sage.

Howell, John R., Lee, Sanghak, Allenby, Greg M., 2015. Price promotions in choice models. Marketing Science 35 (2), 319–334.

Kim, Dong Soo, Bailey, Roger A., Hardt, Nino, Allenby, Greg M., 2016. Benefit-based conjoint analysis. Marketing Science 36 (1), 54–69.

Kim, Hyowon, Kim, Dongsoo, Allenby, Greg M., 2017. Benefit Formation and Enhancement. Working paper. Fisher College of Business, The Ohio State University.

Kotler, Philip, 2012. Kotler on Marketing. Simon and Schuster.

Lancsar, Emily, Savage, Elizabeth, 2004. Deriving welfare measures from discrete choice experiments: inconsistency between current methods and random utility and welfare theory. Health Economics 13 (9), 901–907.

Lancsar, Emily, Swait, Joffre, 2014. Reconceptualising the external validity of discrete choice experiments. PharmacoEconomics 32 (10), 951–965.

Lichtenstein, Sarah, Slovic, Paul, 2006. The Construction of Preference. Cambridge University Press.

Louviere, Jordan J., Hensher, David A., 1983. Using discrete choice models with experimental design data to forecast consumer demand for a unique cultural event. Journal of Consumer Research 10 (3), 348–361.

Luce, R. Duncan, Tukey, John W., 1964. Simultaneous conjoint measurement: a new type of fundamental measurement. Journal of Mathematical Psychology 1 (1), 1–27.

Manski, Charles F., McFadden, Daniel, et al., 1981. Structural Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge, MA.

Matzkin, Rosa L., 1991. Semiparametric estimation of monotone and concave utility functions for polychotomous choice models. Econometrica, 1315–1327.

Matzkin, Rosa L., 1993. Nonparametric identification and estimation of polychotomous choice models. Journal of Econometrics 58 (1–2), 137–168.

Netzer, Oded, Srinivasan, Visvanathan, 2011. Adaptive self-explication of multiattribute preferences. Journal of Marketing Research 48 (1), 140–156.

Ofek, Elie, Srinivasan, Venkataraman, 2002. How much does the market value an improvement in a product attribute? Marketing Science 21 (4), 398–411.

Orme, Bryan K., 2010. Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Research Publishers.

Pachali, Max J., Kurz, Peter, Otter, Thomas, 2017. The perils of ignoring the budget constraint in single-unit demand models. https://doi.org/10.2139/ssrn.3044553.

Ranjan, Bhoomija, Lovett, Mitchell J., Ellickson, Paul B., 2017. Product launches with new attributes: a hybrid conjoint-consumer panel technique for estimating demand. https://doi.org/10.2139/ssrn.3045379.

Roe, Robert M., Busemeyer, Jermone R., Townsend, James T., 2001. Multialternative decision field theory: a dynamic connectionist model of decision making. Psychological Review 108 (2), 370.

Rossi, P.E., 2014. Bayesian Semi-Parametric and Non-Parametric Methods with Applications to Marketing and Micro-Econometrics. Princeton University Press.

Rossi, Peter E., Allenby, Greg M., McCulloch, Robert, 2005. Bayesian Statistics and Marketing. John Wiley and Sons Ltd.

Small, Kenneth A., Rosen, Harvey S., 1981. Applied welfare economics with discrete choice models. Econometrica, 105–130.

Srinivasan, Vinay, Park, Chan Su, 1997. Surprising robustness of the self-explicated approach to customer preference structure measurement. Journal of Marketing Research, 286–291.

Toubia, Olivier, Hauser, John R., Simester, Duncan I., 2004. Polyhedral methods for adaptive choice-based conjoint analysis. Journal of Marketing Research 41 (1), 116–131.

Trajtenberg, Manuel, 1989. The welfare analysis of product innovations, with an application to computed tomography scanners. Journal of Political Economy 97 (2), 444–479.

Yang, Cathy L., Toubia, Olivier, de Jong, Martijn G., 2018. Attention, information processing and choice in incentive-aligned choice experiments. Journal of Marketing Research 55 (6), 783–800.