

# Outside Good Uncertainty

## Ordinal Dual Response in Choice-Based Conjoint Analysis

Prachi Bhalerao

Dan Yavorsky

Geoffery Zheng

2024-09-30

We provide a behavioral model in which consumers are able to identify their most-preferred option among a set of alternatives, but face uncertainty in their evaluation of the outside good. The model rationalizes the use of a dual-response framework in choice-based conjoint analysis and extends the evaluation of the purchase decision from binary to ordinal. A simulation study provides estimation routines for aggregate and Hierarchical Bayesian versions of the model, and demonstrates the model's superior performance compared with the binary and heuristic approaches currently used by practitioners. We also provide an empirical example from a choice-based conjoint analysis conducted in partnership with a consumer insights consultancy.

## 1 Introduction

Motivate with:

### 1. Anecdotes:

- Quote stats on cart abandonment rates. Consumers can pick a preferred bluetooth speaker out of many brands (or a preferred color out of many options from the same brand), but then they may hesitate making the purchase.
- no unusual reaction from consumers if you ask them "how likely are you to purchase?" Uncertainty about making a purchase "feels natural."

### 2. Psych findings:

- [possibly use] decoy effect: low purchase prob when coke machine by itself, but high coke and pepsi purchase prob when machines side-by-side,

### 3. Practice:

- Brazell et al 2006 paper

- Sawtooth’s implementation in Lighthouse Studio application

Describe: The dual-response framework asks for an immediate comparison (which of these things in front of you do you like best) and a hypothetical future scenario (if this product was available to purchase in the near future, would you purchase it). We provide a behavioral model that rationalizes use of the dual-response framework

## 2 Model

### 2.1 Set Up

Consumer  $i = 1, \dots, N$  derives utility from good  $j \in \mathcal{J}_i = \{0, 1, 2, \dots, J_i\}$  with utility  $u_{ij}$  given by

$$u_{ij} = h(\mathbf{x}_j, \beta_i) + \eta_{ij}.$$

where  $\mathbf{x}_j$  is a length- $P$  column vector of good characteristics,  $\beta_i$  is a length- $P$  column vector of consumer-specific taste parameters, and  $\eta_{ij}$  encapsulates factors known to the consumer but unobserved by the researcher. We take  $h(\mathbf{x}_j, \beta_i) = \mathbf{x}_j' \beta_i$  but this specification is not required.

The “zero”-th (or “outside”) good is special and is associated with a zero vector of good characteristics ( $\mathbf{x}_0 = \mathbf{0}$ ) such that  $h(\mathbf{x}_0, \beta_i) = 0$ . Consumers observe  $\mathbf{x}_j$  and  $\eta_{ij}$  for all “inside” goods ( $j \in \mathcal{J}_i^+ := \{1, 2, \dots, J_i\}$ ), but they *do not* observe  $\eta_{i0}$ .

The consumer first reports her preferred inside good, which is given by

$$j_i^* = \operatorname{argmax}_{j \in \mathcal{J}^+} u_{ij}.$$

The consumer then reports a value  $y_i$  on a discrete qualitative scale  $w \in \mathcal{W} = \{1, \dots, W\}$  to reflect the probability  $p_i$  that she would purchase  $j^*$ , i.e., that she prefers good  $j_i^*$  to the outside good  $j = 0$ .

Let  $F$  denote the cumulative distribution function associated with  $p_i$ . Apportion the range of  $F$  into a number  $W$  of intervals  $\phi_w = [\alpha_{w-1}, \alpha_w)$ , where  $\alpha_0 = 0$  and  $\alpha_W = 1$ . We assume that consumers share the definition of the qualitative scale and thus understand these intervals, whereas  $\alpha_w$  for  $0 < w < W$  are unobserved by the researcher and are to be estimated.

Let  $u_i^*$  denote the utility of good  $j_i^*$  for consumer  $i$ . Consumers know  $u_i^*$ , but do not know  $\eta_{i0}$ , and thus the purchase probability from the consumer’s perspective is given by

$$p_i = \Pr(\eta_{i0} < u_i^*) = F(u_i^*).$$

The act of reporting the interval  $y_i = w$  into which  $p_i$  falls, reveals that  $p_i \in \phi_{w(i)} = [\alpha_{w-1}, \alpha_w)$  and is therefore equivalent to reporting that

$$u_i^* \in [F^{-1}(\alpha_{w(i)-1}), F^{-1}(\alpha_{w(i)})].$$

## 2.2 Probability Specification

We model  $\eta_{ij} \stackrel{iid}{\sim} \text{Gumbel}(0, 1)$  with  $\eta_{ij}$  assumed to be independent of both  $\mathbf{x}_j$  and  $\eta_{ik}$  for  $k \neq j$ ,<sup>1 2</sup>

From the consumer's perspective, her purchase probability is

$$p_i = \Pr(\eta_{i0} < u_i^*) = F_{\text{Gumbel}(0,1)}(u_i^*)$$

and, before she reports  $y_i$ , the probability of reporting any particular value  $w$  on the ordinal scale  $\Pr(y_i = w)$  is given by

$$\begin{aligned} \Pr(y_i = w) &= F_{\text{Gumbel}(0,1)}\left(F_{\text{Gumbel}(0,1)}^{-1}(\alpha_w)\right) - F_{\text{Gumbel}(0,1)}\left(F_{\text{Gumbel}(0,1)}^{-1}(\alpha_{w-1})\right) \\ &= \alpha_w - \alpha_{w-1}. \end{aligned}$$

From the researcher's perspective, the unobserved likelihood of purchase is

$$p_i = \Pr(\eta_{i0} - u_i^* < 0) = \Pr(\varepsilon_i < 0)$$

where  $\varepsilon_i = \eta_{i0} - u_i^*$ . As is well known,<sup>3</sup>  $u_i^*$  follows a Gumbel distribution with location parameter  $\bar{\mu}_i$  and scale parameter 1, where

$$\bar{\mu}_i = \ln \left( \sum_{j \in \mathcal{J}_i} \exp(\mathbf{x}_j' \boldsymbol{\beta}_i) \right). \quad (1)$$

and so  $\varepsilon_i$  follows a Logistic distribution with location parameter  $\bar{\mu}_i$  and scale parameter 1. The unobserved choice probability is

$$p_i = F_{\text{Logistic}(\bar{\mu}_i, 1)}(\varepsilon_i).$$

Therefore, the probability that consumer  $i$  reports  $y_i = w$  is

$$\begin{aligned} \Pr(u_i^* \text{ s.t. } p_i \in \phi_{w(i)}) &= F_{\text{Logistic}(\bar{\mu}, 1)}\left(F_{\text{Gumbel}(0,1)}^{-1}(\alpha_{w(i)})\right) - F_{\text{Logistic}(\bar{\mu}, 1)}\left(F_{\text{Gumbel}(0,1)}^{-1}(\alpha_{w(i)-1})\right) \\ &= \frac{1}{1 - \exp(\bar{\mu}) \ln \alpha_{w(i)}} - \frac{1}{1 - \exp(\bar{\mu}) \ln \alpha_{w(i)-1}} \end{aligned}$$

<sup>1</sup>This can be motivated by a framework in which  $u_0 = 0$  and  $u_{ij} = \mathbf{x}_j' \boldsymbol{\beta}_i + \eta_{ij} - \eta_0$  for  $j \in \mathcal{J}_i^+$ . Here,  $\eta_0$  captures the consumer's uncertainty about their future tastes. Given that utilities are ordinal and  $\eta_0$  is a common shock, it plays no role in the choice among the most preferred inside good  $j^* \in \mathcal{J}_i^+$ .

<sup>2</sup>We note that, while the parametrization of  $\eta_0 \sim \text{Gumbel}(0, 1)$  preserves symmetry among the  $J + 1$  goods and is thus a natural choice, the framework can easily accommodate an alternative distribution for  $\eta_0$ . For example, one could use an affine function of individual characteristics to accommodate individual-level variation in the propensity to prefer the outside good.

<sup>3</sup>See, e.g., (McFadden 1981) or (Cardell 1997).

### 3 Simulation Study

1. Agg MNL

- Simulate data ( $\beta_i = \beta$ ), then:
- recover parameters
- compare to dichotomized dual-response
- compare to constant-prob (if  $W = 5$  set  $\alpha$ 's such that  $\Pr(y_i = w) = 1/5$  )
- or, instead of  $1/W$  do the observed frequency

1. HB MNL

- Same 4 things as above

1. Assessing  $\alpha$  cut-points

- May need to try extreme values for some  $\alpha$ 's
- What happens when different consumers have different interpretation of the  $W$  categories such that  $\alpha$ 's are not the same for all consumers?

### 4 Empirical Analysis

add

### 5 Discussion

Differs from (Brazell et al. 2006)

### 6 Conclusion

add

## References

- Brazell, Jeff D., Christopher G. Diener, Ekaterina Karniouchina, William L. Moore, Valérie Séverin, and Pierre-Francois Uldry. 2006. "The No-Choice Option and Dual Response Choice Designs." *Marketing Letters* 17 (4): 255–68. <https://doi.org/10.1007/s11002-006-7943-8>.
- Cardell, N. Scott. 1997. "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity." *Econometric Theory* 13 (2): 185–213. <https://doi.org/10.1017/s0266466600005727>.
- McFadden, Daniel L. 1981. "Structural Discrete Probability Models Derived from Theories of Choice." In, 198–272. The MIT Press.