

Online Appendix

A Properties in Large Menus

A.1 Conditional probability of choosing an item from a menu with K items

Define $\mathbf{Y}' = \{Y_1, \dots, Y_k, \dots, Y_K = 0\}$ as some combination of items that excludes, without loss of generality, the K -th item, but is otherwise equal to $\mathbf{Y} = \{Y_1, \dots, Y_k, \dots, Y_K = 1\}$. According to Equation 4, given the utility specification in Equation 5, and following basic principles of probability, the conditional probability of $Y_K = 1$ then is:

$$\begin{aligned}
 \Pr(Y_K = 1|\mathbf{Y}') &= \frac{\Pr(\mathbf{Y})}{\Pr(\mathbf{Y}')} = \frac{\Pr(Y_K = 1, \mathbf{Y}')}{\Pr(\mathbf{Y}')} = \\
 &= \frac{\exp\left(\sum_{k=1}^K \beta_{i,k} Y_{i,k} + \sum_{k=1}^K \sum_{k'=k+1}^K \theta_{k,k'} Y_{i,k} Y_{i,k'}\right)}{\exp\left(\sum_{k=1}^K \beta_{i,k} Y_{i,k} + \sum_{k=1}^K \sum_{k'=k+1}^K \theta_{k,k'} Y_{i,k} Y_{i,k'}\right) + \exp\left(\sum_{k=1}^{K-1} \beta_{i,k} Y_{i,k} + \sum_{k=1}^{K-1} \sum_{k'=k+1}^{K-1} \theta_{k,k'} Y_{i,k} Y_{i,k'}\right)} \\
 &= \frac{\exp\left(\sum_{k=1}^{K-1} \beta_{i,k} Y_{i,k} + \beta_{i,K} + \sum_{k=1}^{K-1} \sum_{k'=k+1}^{K-1} \theta_{k,k'} Y_{i,k} Y_{i,k'} + \sum_{k=1}^{K-1} \theta_{k,K} Y_{i,k}\right)}{\exp\left(\sum_{k=1}^K \beta_{i,k} Y_{i,k} + \sum_{k=1}^K \sum_{k'=k+1}^K \theta_{k,k'} Y_{i,k} Y_{i,k'}\right) + \exp\left(\sum_{k=1}^{K-1} \beta_{i,k} Y_{i,k} + \sum_{k=1}^{K-1} \sum_{k'=k+1}^{K-1} \theta_{k,k'} Y_{i,k} Y_{i,k'}\right)} \\
 &= \frac{\exp\left(\beta_{i,K} + \sum_{k=1}^{K-1} \theta_{k,K} Y_{i,k}\right)}{\exp\left(\beta_{i,K} + \sum_{k=1}^{K-1} \theta_{k,K} Y_{i,k}\right) + 1} \quad (1)
 \end{aligned}$$

Here, $\Pr(\mathbf{Y}')$ in the first line is defined as $\Pr(\mathbf{Y}', Y_K = 1) + \Pr(\mathbf{Y}', Y_K = 0)$. Remarkably, the resulting conditional probability is independent of the attractiveness, and the substitutive or complementary nature of rejected items in this menu, and independent of the item specific attractiveness of conditioning arguments, i.e., independent of $\sum_{k=1}^{K-1} \beta_k$ given choice \mathbf{Y}' .

Denote the direct effect of other items in \mathbf{Y}' on the conditional choice probability of item K , i.e., $\sum_{k=1}^{K-1} \theta_{k,K} Y_{i,k}$ as $\Upsilon(\mathbf{Y}', \Theta)$. The correspondence between this direct effect and (collective) substitution (complementarity) between item K and items in \mathbf{Y}' is the same as in the simple case of a two-item menu developed in the body of the paper.

A.2 Marginal probability of choosing an item from a menu with K items and cross-derivative

Next we show how changing the attractiveness of items in \mathbf{Y}' , i.e., changing $\sum_{k=1}^{K-1} \beta_k$ affects the marginal probability of choosing item K defined as $\Pr(Y_K = 1) = \Pr(Y_K = 1, \mathbf{Y}') + \Pr(Y_K = 1, \{\emptyset\})$. We will write the deterministic components of utility derived

from the choices of the outside good, only Y_K , \mathbf{Y}' , and $\mathbf{Y} = \{Y_K, \mathbf{Y}'\}$ as $U(\emptyset)$, $U(Y_K)$, $U(\mathbf{Y}')$, and $U(\mathbf{Y})$, respectively. As before let $\Upsilon = \sum_{k=1}^{K-1} \theta_{k,K} Y_{i,k}$ and define the utility interactions in choice $U(\mathbf{Y}')$ as $\varrho = \sum_{k=1}^{K-1} \sum_{k'=k+1}^{K-1} Y_k Y_{k'} \theta_{k,k'}$. Note that $U(\emptyset) = 0$, $U(Y_K) = \beta_K$, $U(\mathbf{Y}') = \sum_{k=1}^{K-1} Y_k \beta_k + \varrho$, and finally $U(\mathbf{Y}) = \beta_K + \sum_{k=1}^{K-1} Y_k \beta_k + \Upsilon + \varrho = U(Y_K) + U(\mathbf{Y}') + \Upsilon$. We can now compactly express the marginal probability of choosing item K as defined before as:

$$\Pr(Y_K = 1) = \frac{e^{U(Y_K)} + e^{U(Y_K) + U(\mathbf{Y}') + \Upsilon}}{1 + e^{U(Y_K)} + e^{U(\mathbf{Y}')} + e^{U(Y_K) + U(\mathbf{Y}') + \Upsilon}} \quad (2)$$

We are interested in obtaining the degree to which the attractiveness of other items influences the marginal probability of choosing item K . In other words, we need to compute $\partial \Pr(Y_K = 1) / \partial \left(\sum_{k=1}^{K-1} \beta_k \right)$. We note that since $\partial U_{\mathbf{Y}'} / \partial \left(\sum_{k=1}^{K-1} \beta_k \right) = 1$, we have $\partial \Pr(Y_K = 1) / \partial \left(\sum_{k=1}^{K-1} \beta_k \right) = \partial \Pr(Y_K = 1) / \partial U_{\mathbf{Y}'}$.

Taking the derivative of Equation 2 w.r.t. $U_{\mathbf{Y}'}$ we get

$$\begin{aligned} \frac{\partial \Pr(Y_k = 1)}{\partial U(\mathbf{Y}')} = & \left[\left(1 + e^{U(Y_K)} + e^{U(\mathbf{Y}')} + e^{U(Y_K) + U(\mathbf{Y}') + \Upsilon} \right) \left(e^{U(Y_K) + U(\mathbf{Y}') + \Upsilon} \right) - \right. \\ & \left. \left(e^{U(Y_K)} + e^{U(Y_K) + U(\mathbf{Y}') + \Upsilon} \right) \left(e^{U(\mathbf{Y}')} + e^{U(Y_K) + U(\mathbf{Y}') + \Upsilon} \right) \right] \times \\ & \left(1 + e^{U(Y_K)} + e^{U(\mathbf{Y}')} + e^{U(Y_K) + U(\mathbf{Y}') + \Upsilon} \right)^{-2} \end{aligned}$$

which simplifies to

$$\frac{\partial \Pr(Y_k = 1)}{\partial U(\mathbf{Y}')} = \frac{\left(e^{U(Y_K) + U(\mathbf{Y}') + \Upsilon} \right) - \left(e^{U(Y_K) + U(\mathbf{Y}')} \right)}{\left(1 + e^{U(Y_K)} + e^{U(\mathbf{Y}')} + e^{U(Y_K) + U(\mathbf{Y}') + \Upsilon} \right)^2}$$

Then it is straightforward to see that

$$\text{sign} \left(\frac{\partial \Pr(Y_K = 1)}{\partial \left(\sum_{k=1}^{K-1} Y_k \beta_k \right)} \right) = \text{sign} \left(\frac{\partial \Pr(Y_k = 1)}{\partial U(\mathbf{Y}')} \right) = \text{sign}(\Upsilon)$$

Thus, the results developed in Section 2.2 for the special case of a two-item menu generalize to arbitrarily large menus and, in an obvious manner, to any two disjunctive sets of items in a menu.

B Convergence Proof for the Generative Model

B.1 Ergodicity

Let there be two bundles \mathbf{Y} and \mathbf{Y}' each of which is a vector of K binary elements. In our framework, the transition from one bundle to another in one step is only possible if they differ by at most one element. In all other cases the transition probability is zero. For our discussion we will assume, for now, that the bundles differ only in the k -th element, so that

$$\mathbf{Y} = (Y_k = 1, \mathbf{Y}_{-k})$$

$$\mathbf{Y}' = (Y_k = 0, \mathbf{Y}_{-k})$$

The transition kernel then can be written as

$$\mathbb{P}_{\mathbf{Y} \rightarrow \mathbf{Y}'} = \lambda_k \Pr(Y_k = 0 | \mathbf{Y}_{-k})$$

$$\mathbb{P}_{\mathbf{Y}' \rightarrow \mathbf{Y}} = \lambda_k \Pr(Y_k = 1 | \mathbf{Y}_{-k})$$

where the λ_k are the attention probabilities and the second terms reflect the conditional probability of choosing the k -th element given the other items in the bundle. It is immediately obvious that the transition probabilities are such that the consumer can move from a given bundle to any other bundle that differs on one item in one step with positive probability.

The more general transition from $\widetilde{\mathbf{Y}}$ to $\widetilde{\mathbf{Y}'}$ that differ on more than one element is a little more involved. We first note that since the consumer can move to one-item different bundle in one step, there will always exist a finite number of steps $m + 1$ in which the consumer can move from $\widetilde{\mathbf{Y}}$ to $\widetilde{\mathbf{Y}'}$. We will define the transition probability between $\widetilde{\mathbf{Y}}$ and $\widetilde{\mathbf{Y}'}$ as follows

$$\mathcal{T}(\widetilde{\mathbf{Y}} \rightarrow \widetilde{\mathbf{Y}'}) = \mathbb{P}_{\mathbf{Y}_m \rightarrow \widetilde{\mathbf{Y}'}} \left(\prod_{l=1}^{m-1} \mathbb{P}_{\mathbf{Y}_l \rightarrow \mathbf{Y}_{l+1}} \right) \mathbb{P}_{\widetilde{\mathbf{Y}} \rightarrow \mathbf{Y}_1}$$

Since $\mathcal{T}(\widetilde{\mathbf{Y}} \rightarrow \widetilde{\mathbf{Y}'}) > 0$ everywhere and $E(m) < \infty$ it follows that the Markov chain induced by the transition probabilities defined above is irreducible and aperiodic. Further, since the chain is also defined over a finite state-space it is also positive (Harris) recurrent. It follows that the Markov chain is ergodic and has a unique stationary distribution.

B.2 Proof that Equation 1 is the (unique) Stationary Distribution of the Generative Model

We take advantage of the correspondence between time-reversibility, i.e., detailed balance, and stationarity in Markov processes. Thus, we need to show that

$$\mathbb{P}_{\mathbf{Y} \rightarrow \mathbf{Y}'} \pi_{\mathbf{Y}} = \mathbb{P}_{\mathbf{Y}' \rightarrow \mathbf{Y}} \pi_{\mathbf{Y}'}, \quad (3)$$

where $\mathbb{P}_{\mathbf{Y} \rightarrow \mathbf{Y}'}$ is as defined earlier and

$$\begin{aligned} \pi_{\mathbf{Y}} &= \Pr(\mathbf{Y}^{(s)} = \mathbf{Y}) \\ &= \frac{\exp(U(\mathbf{Y}; \Psi))}{\sum_{\mathbf{Y}^{\circ} \in \mathcal{Y}} \exp(U(\mathbf{Y}^{\circ}; \Psi))}, \end{aligned}$$

as defined in Equation 1. First we note that the transitions can only change one item at a time. Since the probability of all other moves is zero, Equation 3 is satisfied in those cases trivially. Now let

$$\begin{aligned} \Delta U &= U(Y_k^{(s)} = 1, \mathbf{Y}_{-k}^{(s-1)}; \Psi) - U(Y_k^{(s)} = 0, \mathbf{Y}_{-k}^{(s-1)}; \Psi) \\ &= U(\mathbf{Y}; \Psi) - U(\mathbf{Y}'; \Psi) \end{aligned}$$

Then,

$$\begin{aligned} \mathbb{P}_{\mathbf{Y} \rightarrow \mathbf{Y}'} \pi_{\mathbf{Y}} &= \lambda_k \times \Pr(Y_k = 0 | \mathbf{Y}_{-k}) \frac{\exp(U(\mathbf{Y}; \Psi))}{\sum_{\mathbf{Y}^{\circ} \in \mathcal{Y}} \exp(U(\mathbf{Y}^{\circ}; \Psi))} \\ &= \lambda_k \times \frac{1}{1 + \exp(\Delta U)} \times \frac{\exp(U(\mathbf{Y}; \Psi))}{\mathcal{Z}} \\ &= \lambda_k \times \frac{1}{1 + \exp(\Delta U)} \times \frac{\exp(U(\mathbf{Y}; \Psi))}{\mathcal{Z}} \times \frac{\exp(U(\mathbf{Y}'; \Psi))}{\exp(U(\mathbf{Y}'; \Psi))} \\ &= \lambda_k \times \frac{1}{1 + \exp(\Delta U)} \times \frac{\exp(U(\mathbf{Y}; \Psi) + U(\mathbf{Y}'; \Psi) - U(\mathbf{Y}'; \Psi))}{\mathcal{Z}} \\ &= \lambda_k \times \frac{1}{1 + \exp(\Delta U)} \times \frac{\exp(\Delta U + U(\mathbf{Y}'; \Psi))}{\mathcal{Z}} \\ &= \lambda_k \times \frac{\exp(\Delta U)}{1 + \exp(\Delta U)} \times \frac{\exp(U(\mathbf{Y}'; \Psi))}{\mathcal{Z}} \\ &= \lambda_k \times \Pr(Y_k = 1 | \mathbf{Y}'_{-k}) \frac{\exp(U(\mathbf{Y}'; \Psi))}{\sum_{\mathbf{Y}^{\circ} \in \mathcal{Y}} \exp(U(\mathbf{Y}^{\circ}; \Psi))} \\ &= \mathbb{P}_{\mathbf{Y}' \rightarrow \mathbf{Y}} \pi_{\mathbf{Y}'} \quad \square \end{aligned}$$

The last line follows from the fact that

$$\mathbf{Y}_{-k} = \mathbf{Y}'_{-k}$$

Since the distribution across bundles satisfies detailed balance, it follows that it is the stationary distribution of the Markov chain defined by the transition kernel of the generative model introduced in Section 2.1. In addition, since we have established that the chain is ergodic, this distribution is unique.

C Sampling \mathbf{Y}_i^c

The exchange algorithm requires that we sample from Equation 1. Direct sampling requires evaluating the normalizing constant $\mathcal{Z}_t(\Psi_i^c)$ which the exchange algorithm is designed to avoid.

C.1 Gibbs Sampling

One way to obtain draws of $\mathbf{Y}_{i,t}$ without evaluating $\mathcal{Z}_t(\Psi_i^c)$ is to “Gibbs-through” the conditional distributions defined in Equation 1. Note that this Gibbs-sampler is that special case of the generative model introduced in Section 2.1 that updates each conditional item choice with probability one in one Gibbs iteration. Because these conditional distributions are in binomial logit form, probabilities can be computed fast. The Gibbs-sampler updates elements in $\mathbf{Y}_{i,t} = (Y_{i,t,1}, \dots, Y_{i,t,k}, \dots, Y_{i,t,K})$ one at a time conditional on $\mathbf{Y}_{i,t,-k}$.

1. Initialize $\mathbf{Y}_{i,t}^{(m-1)}$ as the observed vector of choices.
2. Sample from $Y_{i,t,k} \sim \text{Bernoulli}(\Pr(Y_{i,t,k} | \Psi_i^c, \mathbf{Y}_{i,t,-k}^{(m-1)}))$
3. Replace the k -th element in $\mathbf{Y}_{i,t}^{(m-1)}$ by the draw just obtained and proceed to generating $Y_{i,t,k+1}$ as in step number 2. Repeat steps 2 and 3 until all K elements in $\mathbf{Y}_{i,t}^{(m-1)}$ have been updated. Then set $\mathbf{Y}_{i,t}^{(m)} = \mathbf{Y}_{i,t}^{(m-1)}$. This completes one Gibbs cycle.
4. Return to step 2 until convergence from the initial condition.

A drawback of this approach is that it remains unclear how often the Gibbs-cycle needs to be repeated before a draw equivalent to direct sampling from Equation 1 is obtained. An

elegant way around this problem is the perfect sampler proposed by Propp and Wilson (1996) that we describe next.

C.2 Perfect Sampling

The pseudo code we provide here implements a version of Propp and Wilson (1996) “coupling from the past” algorithm. We provide `Rcpp`-code and an `R`-run file as part of the electronic Appendix. For simplicity we omit the subscripts i (for individual) and t (for choice task) in what follows.

1. Initialize all elements of the auxiliary vector \mathbf{Y}^c as “?”

2. For all $k = 1, \dots, K$

Define $p(k)_{max} = \max_{\mathbf{Y}_{-k}} \Pr(Y_k^c = 1 | \mathbf{Y}_{-k})$

Define $p(k)_{min} = \min_{\mathbf{Y}_{-k}} \Pr(Y_k^c = 1 | \mathbf{Y}_{-k})$

3. For $l = -T$ to 0:

For $k = 1, \dots, K$:

(a) Draw u^* from $U[0,1]$

(b) Set $Y_k^{c(l)}$ to 1 if $p_{min} > u^*$

Set $Y_k^{c(l)}$ to 0 if $p_{max} < u^*$

Set $Y_k^{c(l)}$ to ? if $p_{max} > u^* > p_{min}$

(c) If $Y_k^{c(l)} \neq ?$, update the corresponding conditioning argument in $p(j)_{max}$ and $p(j)_{min}$ for all $j \neq k$.

4. Check coalescence: If all elements of $\mathbf{Y}^{c(0)}$ are determined, i.e., no “?”s remain, accept \mathbf{Y}^c as a draw from the target distribution.¹ If not, the time-loop is run for $l = -2T$ to 0, and then again for $l = -2(2T)$ to 0, and so on until coalescence.

To initialize p_{max} and p_{min} in step 2 above, we rely on the full conditional probabilities, whereby p_{max} refers to the choice configuration that maximizes a particular conditional probability and p_{min} refers to the choice configuration that minimizes it. These configurations are easily determined as a function of utility interactions Θ . p_{min} (p_{max}) results

¹When this happens, the distinction between \mathbf{Y}^c and the different conditioning arguments in $p(k)_{max}$ and $p(k)_{min}$, $k = 1, \dots, K$ vanishes such that $p(k)_{max} = p(k)_{min}$ for all k , and we are back to Gibbs-sampling.

from conditioning on choosing everything that acts as a substitute for (complement to) the k -th product, i.e., is connected to product k via a negative (positive) element in Θ .

The idea behind this approach is to implicitly define as many Markov Chains as we have products in a particular menu. Each chain is characterized by its upper and lower (conditional) probability, i.e., the starting values that yield the product specific minimal and maximal conditional choice probabilities. In step 2a and 2b we sample from these different Markov chains. If $p_{min} > u^*$, we have that a particular product is chosen even under the least favorable configuration of conditioning arguments. Conversely, if $p_{max} < u^*$ we have that a particular product is not chosen even under the most favorable configuration of conditioning arguments. In either case, we can use the outcome to update the set of conditioning arguments in all initially different Markov chains. This is because the same draw of u^* would have resulted in the same update across all chain due to the definition of p_{min} and p_{max} . As conditioning arguments are updated, the sets of conditioning arguments originally differentiating the chains become more similar until all chains “coalesce”, i.e., share the same values of conditioning arguments. Note that once this happens $p_{min} = p_{max}$, and the algorithm becomes a standard Gibbs sampler, however, with the assurance that draws necessarily come from the invariant target distribution. In other words, the standard Gibbs sampler runs one Markov chain initialized at some arbitrary value and simulates to convergence “into the future”. The perfect sampler based on the “coupling from the past”-paradigm, in contrast, directly operationalizes convergence as the coalescence of as many Markov chains as there are products in the menu.

In our case study we rely on the exchange algorithm in combination with perfect sampling from Equation 1. In simulation studies not reported here, we found both Gibbs-sampling and perfect sampling of \mathbf{Y}_i^c to work in practice. However, we prefer not to worry about the number of Gibbs-cycles required for convergence to $\mathbf{Y}_{i,t}^c \sim \text{Pr}_t(\Psi_i^c)$ because even small amounts of approximation bias at the individual level can translate into substantial overall bias in hierarchical models that pool information across respondents. Nevertheless, the Gibbs-sampler may be the better choice when the goal is to estimate probabilities, which requires larger samples of $\mathbf{Y}_{i,t}$, because each individual draw from the perfect sampler is more computationally expensive. We use the Gibbs-sampler when “simulating forward” from the model for predictive validation and to compute counterfactual results in our case study.

D MCMC Sampling

Unless noted otherwise, individual sampling steps are implemented in plain R.

0. Initialize β_i , Θ_i , for all $i = 1, \dots, N$, and δ at zero, and σ_ϵ at 1.
1. Draw from $\mathbf{z}_i | \delta, \sigma_\epsilon, \mathbf{W}, ind_i$ for all $i = 1, \dots, N$ using the inverse CDF-transformation. Here ind_i refers to a vector of classifications of elements in Θ_i into S , I , or C initialized randomly (see Equation 18).
2. Draw from $\delta, \sigma_\epsilon | \{\mathbf{z}_i\}, \mathbf{W}, prior$ using standard conjugate results and weakly informative priors $\delta \sim N(0, \mathbf{A}^{-1})$ and $\sigma_\epsilon \sim IG(\nu, s)$, where $\mathbf{A} = 0.01$, $\nu = 3$, $s = 1$.
3. Compute prior classification probabilities $\Pr(S, I, C) | \delta, \sigma_\epsilon, \mathbf{W}$, see Equations 20 through 24.
4. Draw from $\bar{\beta}, \mathbf{V}_\beta | \{\beta_i\}, prior$. We use standard conjugate results to update hyperparameters $\bar{\beta}, \mathbf{V}_\beta$ based on weakly informative priors:

$$\begin{aligned} \bar{\beta} &\sim N(0, \mathbf{A}^{-1}) \\ \mathbf{V}_\beta &\sim IW(\nu, \mathbf{V}) \end{aligned} \tag{4}$$

where $\mathbf{A} = 0.01$, $\nu = m + 3$, $\mathbf{V} = \nu \mathbf{I}$, and m is the number of elements in $\bar{\beta}$.

5. Draw from $\bar{\theta}, \mathbf{V}_\theta | \{ind_i, \theta_i\}, prior$. See Section 4.2 for the subjective prior parameter setting. We again rely on standard conjugate results.
6. Draw from $\Psi_i, ind_i | \bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, \Pr(S, I, C), \mathbf{Y}_i$. This step employs the MH algorithm for the joint update of individual parameters, $\Psi_i = (\beta_i, \Theta_i)$ and ind_i , where ind_i refers to the vector of classifications of (unique) individual elements $\theta_{i,k,k'}$ in Θ_i . A joint update is required because given the classification of an element $\theta_{i,k,k'}$ into S , I , or C , the hierarchical prior $\theta_{i,k,k'}$ is highly informative.
 - (a) Propose a classification candidate from $ind_i^c | \Pr(S, I, C), p^*$, where p^* is the probability of attempting the reclassification of a (unique) individual elements $\theta_{i,k,k'}$ in Θ_i . This probability is independently generated as $p^* \sim Beta(1, 5)$ at each draw for each respondent, and corresponds to a stochastically determined step-size. If an element of ind_i becomes a candidate for a re-classification attempt, the proposal is generated from its hierarchical prior $\Pr(S, I, C)_{k,k'}$ as

a function of $\boldsymbol{\delta}$, σ_ϵ and $\mathbf{w}_{k,k'}$ (see Equations 20 through 24). In turn, ind_i then determines the structure of the prior for Θ_i .

- (b) Propose a candidate value Ψ_i^c from $\Psi_i|\bar{\boldsymbol{\beta}}, \mathbf{V}_\beta, \bar{\boldsymbol{\theta}}, \mathbf{V}_\theta, ind_i^c, \mathbf{Y}_i$ using auxiliary data augmentation (coded using `Rcpp`). The resulting proposal ensures concordance between ind_i^c and Ψ_i^c which facilitates the large jumps in the parameters space implied by θ -reclassifications.

We generate auxiliary latent variables $\zeta_{i,t,k}$ for item k in menu t for each individual respondent i from a density denoted by $h(\zeta_i|\Psi_i, \mathbf{Y}_i)$ that we explain next. Depending on whether the item was chosen ($Y_{i,t,k} = 1$) or rejected ($Y_{i,t,k} = 0$), we generate $\zeta_{i,t,k}$ by truncating below (above) zero a t -distributed variable with 10 degrees of freedom, mean $\beta_{i,k} + \beta_{price,i} p_{t,k} + \mathbf{Y}'_{i,t} (\Theta_i + \Theta'_i)^k$, and variance $\frac{\pi^2}{3}$, where $(\Theta_i + \Theta'_i)^k$ corresponds to the k -th column of $(\Theta_i + \Theta'_i)$ (see Equation 6), such that $\mathbf{Y}'_{i,t} (\Theta_i + \Theta'_i)^k$ measures the contribution of $\mathbf{Y}'_{i,t,-k}$ to the conditional probability $\Pr(Y_{i,t,k}|\Psi_i, \mathbf{Y}_{i,t,-k})$. The auxiliary latent variables can be thought of as approximate data augmentation for the PL (Equation 17) which implies logistically distributed latent variables (see Equation 1).

We then regress the auxiliary variables $\zeta_{i,k,t}$ on item specific constants and prices, and the respective $\mathbf{Y}'_{i,t,-k}$ as illustrated next using a three-item menu as example. With three items in total, Θ_i contains three unique θ -elements:

$$\Theta_i = \begin{pmatrix} 0 & \theta_{i,1,2} & \theta_{i,1,3} \\ 0 & 0 & \theta_{i,2,3} \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \theta_{i,2,1} \\ \theta_{i,3,1} \\ \theta_{i,3,2} \end{pmatrix}$$

The regression equation then is

$$\begin{pmatrix} \zeta_{i,1,1} \\ \zeta_{i,1,2} \\ \zeta_{i,1,3} \\ \vdots \\ \vdots \\ \zeta_{i,T,1} \\ \zeta_{i,T,2} \\ \zeta_{i,T,3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & p_{1,1} & Y_{i,1,2} & Y_{i,1,3} & 0 \\ 0 & 1 & 0 & p_{1,2} & Y_{i,t,1} & 0 & Y_{i,1,3} \\ 0 & 0 & 1 & p_{1,3} & 0 & Y_{i,1,1} & Y_{i,1,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & p_{T,1} & Y_{i,T,2} & Y_{i,T,3} & 0 \\ 0 & 1 & 0 & p_{T,2} & Y_{i,T,1} & 0 & Y_{i,T,3} \\ 0 & 0 & 1 & p_{T,3} & 0 & Y_{i,T,1} & Y_{i,T,2} \end{pmatrix} \begin{pmatrix} \beta_{i,1} \\ \beta_{i,2} \\ \beta_{i,3} \\ \beta_{i,price} \\ \theta_{i,2,1} \\ \theta_{i,3,1} \\ \theta_{i,3,2} \end{pmatrix} + \begin{pmatrix} \epsilon_{i,1,1} \\ \epsilon_{i,1,2} \\ \epsilon_{i,1,3} \\ \vdots \\ \vdots \\ \epsilon_{i,T,1} \\ \epsilon_{i,T,2} \\ \epsilon_{i,T,3} \end{pmatrix}$$

The first and the central second ‘‘OLS-moments’’ of Ψ_i from this regression are combined with the multivariate normal hierarchical prior $p(\Psi_i|\bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i^c)$ implied by the proposed classifications in ind_i^c to derive the location and scale parameter of a multivariate T proposal distribution with 10 degrees of freedom denoted $g(\Psi_i|\bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i^c, \zeta_i)$.

- (c) Draw auxiliary data \mathbf{Y}_i^c as described in Appendix C.2 (coded using `Rcpp`).
- (d) Finally, we substitute into Equation 15 to obtain the probability of accepting the joint move from (Ψ_i, ind_i) to $(\Psi_i^c, ind_i^c, \mathbf{Y}_i^c)$:

$$\begin{aligned} \alpha(\Psi_i, ind_i \rightarrow \Psi_i^c, ind_i^c, \mathbf{Y}_i^c) = \\ \min \left(1, \frac{p(\Psi_i^c|\bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i^c) \Pr(ind_i^c | \Pr(S, I, C))}{p(\Psi_i|\bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i) \Pr(ind_i | \Pr(S, I, C))} \times \right. \\ \frac{g(\Psi_i|\bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i, \zeta_i) h(\zeta_i|\Psi_i^c, \mathbf{Y}_i) \Pr(ind_i(\mathbf{I}(p^*)) | \Pr(S, I, C))}{g(\Psi_i^c|\bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i^c, \zeta_i) h(\zeta_i|\Psi_i, \mathbf{Y}_i) \Pr(ind_i^c(\mathbf{I}(p^*)) | \Pr(S, I, C))} \times \\ \left. \prod_{t=1}^T \frac{\ell^*(\mathbf{Y}_{i,t}|\Psi_i^c) \ell^*(\mathbf{Y}_{i,t}^c|\Psi_i)}{\ell^*(\mathbf{Y}_{i,t}|\Psi_i) \ell^*(\mathbf{Y}_{i,t}^c|\Psi_i^c)} \right) \end{aligned} \quad (5)$$

The vertically separated factors on the left hand side of Equation 5 in turn correspond to the (hierarchical) prior of (Ψ_i, ind_i) (see Equations 19 and 4, and Equations 20 through 24), the proposal distributions just described, and the likelihood contribution. The factors $h(\zeta_i|\Psi_i^c, \mathbf{Y}_i)$ and $(h(\zeta_i|\Psi_i, \mathbf{Y}_i))^{-1}$ in the third line of Equation 5 ensure the detailed balance between the move from the current state (Ψ_i, ind_i) — via ζ_i — to the candidate state, $(\Psi_i^c, ind_i^c, \mathbf{Y}_i^c)$, and the move in the opposite direction.

Finally, the notation $(\mathbf{I}(p^*))$ indicates the subset of elements in ind_i eligible for a potential reclassification in a particular iteration of the MCMC. All other elements in ind_i deterministically retain their old value in this iteration. Because we use the hierarchical prior as proposal, and elements in ind_i are independently distributed a priori, Equation 5 simplifies to:

$$\begin{aligned}
& \alpha(\Psi_i, ind_i \rightarrow \Psi_i^c, ind_i^c, \mathbf{Y}_i^c) = \\
& \min \left(1, \frac{p(\Psi_i^c | \bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i^c) g(\Psi_i | \bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i, \zeta_i) h(\zeta_i | \Psi_i^c, \mathbf{Y}_i)}{p(\Psi_i | \bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i) g(\Psi_i^c | \bar{\beta}, \mathbf{V}_\beta, \bar{\theta}, \mathbf{V}_\theta, ind_i^c, \zeta_i) h(\zeta_i | \Psi_i, \mathbf{Y}_i)} \times \right. \\
& \quad \left. \prod_{t=1}^T \frac{\ell^*(\mathbf{Y}_{i,t} | \Psi_i^c) \ell^*(\mathbf{Y}_{i,t}^c | \Psi_i)}{\ell^*(\mathbf{Y}_{i,t} | \Psi_i) \ell^*(\mathbf{Y}_{i,t}^c | \Psi_i^c)} \right)
\end{aligned} \tag{6}$$

Implementation note We ran our MCMC code that combines R and Rcpp as explained above on a Windows Desktop PC with an Intel Xeon CPU E5-1660 v4 3.20 GHz processor and 64 GB of RAM. The overall run-time to obtain the estimates for MCMhet reported in the paper is 6.5 days (157h). We started with 100,000 iterations of the computationally cheaper model using the pseudo-likelihood (42h) to navigate into the neighborhood of good posterior support and then ran the chain that combines perfect sampling with the exchange algorithm for another 350,000 iterations (115h). Posterior calculations are based on the last 150,000 draws.² The computational bottleneck is the update of individual parameters, involving perfect sampling and the automatic construction of proposal densities. Thus, there is room for speeding up the code by orders-of-magnitudes by parallelizing the update of individual level parameters (step 6 above).

E IndepCPE - Numerical Illustrations

In our empirical study we estimate credibly negative cross-price effects between products that are clearly identified as strong substitutes on average in the population by the MCM (see Section 5).

To numerically illustrate the misspecification inherent to the IndepCPE that causes this result, we first illustrate the role of price variation in the context of estimating cross-price effects in an individual level choice model without heterogeneity, i.e., looking at one consumer's choices only. We then broaden the investigation to the case of a hierarchical model with heterogeneity.

²MvP takes about the same time to run because of the need to augment dependently distributed utilities.

Estimating cross-price effects in an individual level choice model

We simulate 100 choices from two-item menus comprising alternatives A and B from the MCM with price parameter equal to -1 and alternative specific constants equal to 3 and 5 such that B is the more preferred brand overall. The utility interaction parameter θ is set to -10 implying relatively strong substitution between A and B given the magnitude of alternative specific constants.

Across the 100 two-item menus, prices of alternative A vary uniformly on the grid $[0.2, 0.4, 0.6, \dots, 4]$. Prices of alternative B are constructed according to $p_B = p_A + 2$, such that the expected indirect utilities of both alternatives are exactly equal in each of the 100 menus. Thus, our simulated respondent is, before realizations of the error draws, always indifferent between A and B .

We then create additional data sets using the same setting except for how we generate B 's prices. In these data sets we generate B 's prices as $p_B = p_A + 2 + \xi$, where ξ is distributed iid across menus according to $\xi \sim N(0, \sigma)$. The ξ -draws serve to break the perfect (expected) utility balance between A and B .

Table 1: Posterior Means (posterior SD) of IndepCPE Parameters

Noise	β_{price}	β_A	β_B	$\beta_{cp}^{(B \rightarrow A)}$	$\beta_{cp}^{(A \rightarrow B)}$
0	0.253 (0.857)	-0.337 (1.772)	0.086 (1.734)	-0.178 (0.848)	-0.517 (0.888)
0.1	-0.096 (0.715)	-0.081 (1.517)	0.118 (1.483)	0.008 (0.711)	-0.267 (0.754)
0.2	-0.101 (0.605)	0.950 (1.349)	-0.208 (1.288)	-0.272 (0.607)	0.009 (0.634)
0.3	-0.175 (0.471)	0.884 (1.116)	0.030 (1.016)	-0.219 (0.472)	0.043 (0.515)
0.4	-0.308 (0.377)	0.222 (0.917)	0.390 (0.904)	0.009 (0.373)	0.122 (0.407)
0.5	-0.554 (0.333)	-0.369 (0.841)	0.889 (0.797)	0.275 (0.322)	0.408 (0.384)
0.9	-0.773 (0.219)	-0.388 (0.662)	1.182 (0.657)	0.398 (0.208)	0.607 (0.277)
2	-1.010 (0.164)	-0.440 (0.526)	1.087 (0.610)	0.523 (0.124)	0.882 (0.282)
random	-1.327 (0.231)	-1.722 (0.805)	2.054 (0.693)	1.127 (0.297)	1.153 (0.293)

Note: Based on simulated data

Table 1 summarizes posterior means (standard deviations) of IndepCPE parameters, i.e., the price coefficient, alternative specific constants, the cross effect from B 's price on the indirect utility of A , and vice versa. The rows in Table 1 correspond to different data sets and the first column indicates the standard deviation of the ξ -displacement of B 's price in the data set. Thus, the first row corresponds to the case where the respondent

always is exactly indifferent between A and B , before realizations of the error terms. The last row represents the case where we draw p_A and p_B completely randomly from the predefined set of prices. We see that IndepCPE fails to recover positive cross-price effects when a respondent is indifferent or about indifferent between A and B . As the independent variation in p_B increases, estimated cross-price effects become more positive and the uncertainty in parameters decreases. However, even when the price variation in the data is sufficient for cross-price effects to reflect substitution between A and B , the model fails at generating substitutive effects in individual menus where prices make the respondent indifferent between the two brands.

Estimating cross-price effects in a hierarchical choice model

Next, we replicate the findings from our empirical case study, i.e., large and significantly negative cross-price effects between substitutes in the context of a hierarchical model and a larger menu. The intuition for the counter-intuitive negative cross-price effects is that these help rationalize the fact that only one alternative is chosen in a set comprised of substitutes.

We generate individual choices from the MCM with five items in a menu (A , B , C , D , and E) that are perfect substitutes (all $\theta_i = -100$), such that at maximum one item in a menu can be chosen. We generate 200 respondents with individual alternative specific constants $\beta_{i,k}$ distributed *iid* as $N(3, 16)$ and price-sensitivity $\beta_{i,price} \sim N(-1.3, 0.25)$. Each respondent solves 20 choice tasks. Each choice task (menu) consists of all five inside brands and the outside good. Prices are drawn independently from the grid $[0.2, 0.4, 0.6, \dots, 4]$ with equal probability. We then fit the IndepCPE with a standard multivariate normal hierarchical prior distribution to the simulated data.

Tables 2 through 4 report posterior means of the hyper parameters in the IndepCPE. Posterior standard deviations are in parentheses. Despite the fact that all items are essentially perfect substitutes for all respondents, inferred cross-price effects are negative, and almost all of them (17 out of 20) are credibly different from zero. The results illustrate how IndepCPE is likely to fail staggeringly at even reflecting substitution and complementarity in applications.

Table 2: Hierarchical Prior Means and Standard Deviations (posterior SD):
Price-Sensitivity and Alternative Specific Constants

	Mean	Standard Deviation
β_{price}	-0.002 (0.056)	0.581 (0.036)
β_A	-1.254 (0.440)	1.667 (0.315)
β_B	-1.706 (0.413)	1.704 (0.327)
β_C	-1.365 (0.560)	2.153 (0.522)
β_D	-1.780 (0.443)	2.001 (0.421)
β_E	-1.289 (0.425)	1.840 (0.352)

Note: Based on simulated data

Table 3: Hierarchical Prior Means of Cross-Price Effects (posterior SD)

	A	B	C	D	E
A →	- (0.130)	-0.437 (0.130)	-0.325 (0.143)	-0.345 (0.123)	-0.414 (0.143)
B →	-0.397 (0.137)	- (0.130)	-0.517 (0.143)	-0.157 (0.119)	-0.325 (0.139)
C →	-0.224 (0.164)	-0.423 (0.150)	- (0.143)	-0.221 (0.150)	-0.644 (0.191)
D →	-0.527 (0.135)	-0.228 (0.117)	-0.267 (0.128)	- (0.123)	-0.506 (0.140)
E →	-0.454 (0.135)	-0.279 (0.126)	-0.494 (0.140)	-0.302 (0.123)	- (0.123)

Note: Based on simulated data

Table 4: Hierarchical Prior Standard Deviations of Cross-Price Effects (posterior SD)

	A	B	C	D	E
A →	- (0.088)	0.929 (0.088)	0.941 (0.099)	0.900 (0.086)	0.925 (0.094)
B →	0.942 (0.093)	- (0.088)	0.922 (0.094)	0.909 (0.082)	0.889 (0.083)
C →	1.033 (0.115)	0.964 (0.096)	- (0.094)	0.997 (0.105)	1.137 (0.138)
D →	0.920 (0.095)	0.839 (0.075)	0.854 (0.079)	- (0.079)	0.915 (0.095)
E →	0.934 (0.091)	0.883 (0.087)	0.917 (0.087)	0.875 (0.081)	- (0.081)

Note: Based on simulated data

F Supplemental Tables

Table 5: Hierarchical Prior Standard Deviations:
Heterogeneity in Price Sensitivity and Alternative Specific Constants (posterior SD)

	MCMhet	MCMhom	Indep	MvP	IndepCPE
Price	0.33 (0.01)	0.30 (0.01)	0.31 (0.01)	0.10 (0.01)	0.66 (0.03)
Xbox 360	4.38 (0.27)	3.79 (0.26)	4.42 (0.31)	2.23 (0.16)	4.82 (0.79)
Xbox One	6.42 (0.50)	5.20 (0.42)	5.77 (0.51)	2.76 (0.21)	7.49 (0.79)
Xbox Kinect	2.68 (0.25)	2.65 (0.23)	2.99 (0.30)	1.72 (0.15)	5.46 (0.68)
Xbox Wheel	2.72 (0.29)	2.77 (0.25)	3.27 (0.36)	1.81 (0.15)	4.38 (0.91)
PS3	4.81 (0.28)	3.90 (0.23)	4.33 (0.29)	2.16 (0.12)	7.78 (0.96)
PS4	8.71 (0.48)	7.43 (0.55)	8.29 (0.56)	3.81 (0.21)	11.86 (0.99)
PS Eye	2.39 (0.24)	2.27 (0.17)	2.54 (0.20)	1.56 (0.11)	5.36 (0.62)
PS Move	2.95 (0.23)	3.07 (0.22)	3.39 (0.26)	1.81 (0.12)	8.37 (1.41)
PS Wheel	2.79 (0.20)	3.10 (0.25)	3.53 (0.29)	2.03 (0.16)	8.09 (0.74)
Wii	5.44 (0.43)	4.94 (0.38)	6.00 (0.51)	2.74 (0.23)	7.58 (0.73)
Wii U	5.62 (0.37)	5.22 (0.41)	6.03 (0.47)	2.83 (0.21)	7.90 (1.01)
Wii Wheel	3.84 (0.31)	3.32 (0.31)	3.88 (0.34)	1.98 (0.15)	5.35 (0.72)
Wii Motion	3.25 (0.25)	3.22 (0.23)	3.78 (0.30)	2.07 (0.16)	5.66 (0.59)

Table 6: Correlation Matrix of Individual Preferences (posterior SD) - MCMhet

	Price	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel
Xbox 360	-0.42 (0.05)	1											
Xbox One	-0.42 (0.06)	0.45 (0.07)	1										
Xbox Kinect	-0.17 (0.08)	0.19 (0.11)	0.24 (0.11)	1									
Xbox Wheel	-0.21 (0.06)	0.17 (0.09)	0.04 (0.10)	0.65 (0.10)	1								
PS3	-0.39 (0.06)	0.87 (0.02)	0.32 (0.10)	-0.09 (0.12)	0.04 (0.09)	1							
PS4	-0.38 (0.05)	0.40 (0.07)	0.78 (0.04)	-0.06 (0.12)	-0.08 (0.10)	0.47 (0.06)	1						
PS Eye	-0.13 (0.07)	0.18 (0.09)	-0.08 (0.09)	0.48 (0.11)	0.57 (0.08)	0.19 (0.11)	0.11 (0.09)	1					
PS Move	-0.28 (0.06)	0.20 (0.08)	0.36 (0.08)	0.54 (0.09)	0.54 (0.08)	0.20 (0.08)	0.48 (0.06)	0.70 (0.06)	1				
PS Wheel	-0.15 (0.06)	0.00 (0.09)	0.03 (0.10)	0.35 (0.10)	0.71 (0.08)	0.06 (0.09)	0.22 (0.09)	0.61 (0.07)	0.69 (0.06)	1			
Wii	-0.38 (0.06)	0.63 (0.06)	0.10 (0.12)	0.08 (0.10)	0.24 (0.09)	0.59 (0.06)	-0.02 (0.08)	0.09 (0.10)	-0.06 (0.10)	-0.05 (0.10)	1		
Wii U	-0.48 (0.05)	0.26 (0.08)	0.63 (0.06)	0.14 (0.10)	0.19 (0.10)	0.24 (0.09)	0.52 (0.07)	0.05 (0.09)	0.27 (0.12)	0.14 (0.10)	0.50 (0.09)	1	
Wii Wheel	-0.27 (0.06)	0.22 (0.08)	-0.01 (0.10)	0.54 (0.11)	0.79 (0.06)	0.10 (0.08)	-0.23 (0.08)	0.33 (0.09)	0.30 (0.08)	0.49 (0.09)	0.48 (0.06)	0.31 (0.09)	1
Wii Motion	-0.31 (0.06)	0.19 (0.08)	0.03 (0.09)	0.53 (0.08)	0.64 (0.07)	0.06 (0.09)	-0.24 (0.08)	0.28 (0.11)	0.26 (0.11)	0.28 (0.12)	0.59 (0.06)	0.46 (0.07)	0.83 (0.04)

Table 7: Correlation Matrix of Individual Preferences (posterior SD) - MCMhom

	Price	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel
Xbox 360	-0.40 (0.05)	1											
Xbox One	-0.37 (0.06)	0.51 (0.08)	1										
Xbox Kinect	-0.15 (0.06)	0.29 (0.08)	0.34 (0.09)	1									
Xbox Wheel	-0.11 (0.07)	0.26 (0.08)	0.08 (0.10)	0.60 (0.07)	1								
PS3	-0.39 (0.05)	0.81 (0.03)	0.35 (0.09)	0.02 (0.09)	0.07 (0.08)	1							
PS4	-0.35 (0.06)	0.30 (0.08)	0.73 (0.05)	-0.04 (0.09)	-0.20 (0.08)	0.45 (0.07)	1						
PS Eye	-0.09 (0.06)	0.16 (0.09)	-0.02 (0.10)	0.45 (0.08)	0.28 (0.08)	0.28 (0.08)	0.08 (0.08)	1					
PS Move	-0.18 (0.06)	0.19 (0.09)	0.35 (0.08)	0.62 (0.07)	0.46 (0.08)	0.23 (0.08)	0.34 (0.07)	0.56 (0.06)	1				
PS Wheel	-0.12 (0.06)	0.07 (0.09)	0.04 (0.09)	0.19 (0.09)	0.64 (0.06)	0.17 (0.08)	0.22 (0.07)	0.26 (0.08)	0.49 (0.07)	1			
Wii	-0.31 (0.06)	0.54 (0.07)	-0.11 (0.10)	0.14 (0.09)	0.23 (0.09)	0.52 (0.07)	-0.16 (0.09)	0.24 (0.09)	-0.09 (0.09)	0.04 (0.09)	1		
Wii U	-0.41 (0.06)	0.29 (0.10)	0.49 (0.08)	0.23 (0.10)	0.09 (0.10)	0.21 (0.10)	0.36 (0.08)	0.11 (0.09)	0.07 (0.09)	0.03 (0.09)	0.46 (0.08)	1	
Wii Wheel	-0.18 (0.06)	0.22 (0.08)	0.00 (0.10)	0.52 (0.08)	0.78 (0.05)	0.02 (0.08)	-0.30 (0.08)	0.16 (0.08)	0.26 (0.08)	0.47 (0.07)	0.44 (0.07)	0.37 (0.08)	1
Wii Motion	-0.24 (0.06)	0.21 (0.09)	-0.04 (0.09)	0.54 (0.08)	0.50 (0.07)	0.08 (0.08)	-0.22 (0.08)	0.28 (0.09)	0.34 (0.08)	0.22 (0.08)	0.60 (0.05)	0.42 (0.08)	0.67 (0.05)

Table 8: Correlation Matrix of Individual Preferences (posterior SD) - Indep Model

	Price	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel
Xbox 360	-0.34 (0.06)	1											
Xbox One	-0.31 (0.08)	0.38 (0.10)	1										
Xbox Kinect	-0.08 (0.07)	0.53 (0.07)	0.41 (0.10)	1									
Xbox Wheel	-0.02 (0.06)	0.48 (0.07)	0.20 (0.09)	0.67 (0.06)	1								
PS3	-0.33 (0.06)	0.74 (0.04)	0.12 (0.10)	0.10 (0.09)	0.17 (0.08)	1							
PS4	-0.30 (0.06)	0.05 (0.09)	0.56 (0.08)	-0.15 (0.10)	-0.25 (0.08)	0.22 (0.08)	1						
PS Eye	-0.06 (0.06)	0.10 (0.08)	0.04 (0.10)	0.24 (0.09)	0.20 (0.08)	0.31 (0.08)	0.36 (0.06)	1					
PS Move	-0.14 (0.06)	0.11 (0.08)	0.33 (0.08)	0.26 (0.09)	0.21 (0.08)	0.28 (0.08)	0.60 (0.05)	0.60 (0.06)	1				
PS Wheel	-0.06 (0.06)	0.06 (0.08)	0.07 (0.08)	0.03 (0.09)	0.41 (0.07)	0.24 (0.08)	0.46 (0.06)	0.52 (0.06)	0.65 (0.05)	1			
Wii	-0.30 (0.07)	0.49 (0.07)	-0.22 (0.13)	0.12 (0.10)	0.17 (0.08)	0.45 (0.07)	-0.27 (0.08)	0.01 (0.08)	-0.25 (0.08)	-0.13 (0.08)	1		
Wii U	-0.38 (0.07)	0.13 (0.10)	0.40 (0.10)	0.12 (0.10)	0.09 (0.08)	0.02 (0.10)	0.20 (0.08)	0.03 (0.08)	0.03 (0.08)	-0.02 (0.07)	0.41 (0.08)	1	
Wii Wheel	-0.12 (0.07)	0.25 (0.08)	-0.06 (0.10)	0.39 (0.08)	0.64 (0.05)	0.03 (0.07)	-0.42 (0.07)	0.01 (0.08)	-0.10 (0.08)	0.15 (0.07)	0.53 (0.06)	0.53 (0.07)	1
Wii Motion	-0.19 (0.06)	0.18 (0.07)	-0.13 (0.09)	0.31 (0.08)	0.37 (0.07)	0.01 (0.08)	-0.33 (0.07)	0.02 (0.08)	-0.05 (0.08)	-0.02 (0.07)	0.69 (0.04)	0.61 (0.05)	0.77 (0.04)

Table 9: Correlation Matrix of Individual Preferences (posterior SD) - MvP

	Price	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel
Xbox 360	-0.44 (0.06)	1											
Xbox One	-0.52 (0.08)	0.36 (0.08)	1										
Xbox Kinect	-0.02 (0.09)	0.42 (0.08)	0.31 (0.09)	1									
Xbox Wheel	0.04 (0.07)	0.41 (0.07)	0.14 (0.09)	0.55 (0.07)	1								
PS3	-0.49 (0.06)	0.67 (0.05)	0.19 (0.09)	0.08 (0.08)	0.14 (0.07)	1							
PS4	-0.49 (0.06)	0.10 (0.08)	0.50 (0.07)	-0.16 (0.09)	-0.25 (0.08)	0.24 (0.07)	1						
PS Eye	-0.01 (0.08)	0.09 (0.08)	-0.02 (0.09)	0.22 (0.08)	0.20 (0.09)	0.23 (0.07)	0.28 (0.07)	1					
PS Move	-0.16 (0.07)	0.10 (0.08)	0.21 (0.08)	0.23 (0.08)	0.17 (0.08)	0.23 (0.07)	0.51 (0.05)	0.52 (0.05)	1				
PS Wheel	0.01 (0.07)	0.04 (0.08)	-0.03 (0.09)	0.06 (0.10)	0.35 (0.08)	0.16 (0.07)	0.37 (0.07)	0.44 (0.07)	0.56 (0.06)	1			
Wii	-0.42 (0.05)	0.44 (0.07)	0.00 (0.09)	0.08 (0.10)	0.15 (0.09)	0.40 (0.06)	-0.18 (0.07)	-0.05 (0.07)	-0.23 (0.07)	-0.18 (0.07)	1		
Wii U	-0.51 (0.07)	0.13 (0.09)	0.39 (0.09)	0.07 (0.09)	0.07 (0.08)	0.06 (0.09)	0.23 (0.08)	0.02 (0.08)	0.02 (0.08)	-0.05 (0.08)	0.38 (0.07)	1	
Wii Wheel	-0.14 (0.07)	0.21 (0.08)	-0.02 (0.09)	0.29 (0.08)	0.54 (0.06)	0.04 (0.08)	-0.36 (0.07)	0.01 (0.08)	-0.08 (0.08)	0.10 (0.07)	0.46 (0.07)	0.43 (0.07)	1
Wii Motion	-0.18 (0.07)	0.16 (0.07)	-0.06 (0.09)	0.23 (0.09)	0.35 (0.08)	0.01 (0.08)	-0.32 (0.07)	-0.01 (0.07)	-0.08 (0.07)	-0.06 (0.08)	0.61 (0.05)	0.51 (0.06)	0.69 (0.04)

Table 10: Correlation Matrix of Individual Preferences (posterior SD) - IndepCPE

	Price	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel
Xbox 360	-0.60 (0.05)	1											
Xbox One	-0.63 (0.04)	0.92 (0.03)	1										
Xbox Kinect	-0.61 (0.05)	0.90 (0.04)	0.93 (0.03)	1									
Xbox Wheel	-0.58 (0.05)	0.88 (0.06)	0.90 (0.05)	0.88 (0.05)	1								
PS3	-0.63 (0.04)	0.93 (0.03)	0.96 (0.02)	0.93 (0.03)	0.91 (0.05)	1							
PS4	-0.63 (0.04)	0.92 (0.05)	0.95 (0.03)	0.91 (0.06)	0.90 (0.05)	0.96 (0.02)	1						
PS Eye	-0.61 (0.04)	0.89 (0.05)	0.93 (0.03)	0.90 (0.04)	0.88 (0.05)	0.94 (0.02)	0.94 (0.02)	1					
PS Move	-0.63 (0.04)	0.92 (0.04)	0.95 (0.03)	0.92 (0.05)	0.91 (0.05)	0.96 (0.02)	0.97 (0.01)	0.94 (0.02)	1				
PS Wheel	-0.62 (0.04)	0.93 (0.04)	0.95 (0.03)	0.92 (0.04)	0.91 (0.05)	0.96 (0.02)	0.97 (0.01)	0.94 (0.02)	0.97 (0.01)	1			
Wii	-0.63 (0.04)	0.92 (0.04)	0.96 (0.02)	0.93 (0.03)	0.90 (0.05)	0.96 (0.02)	0.96 (0.04)	0.93 (0.03)	0.95 (0.03)	0.95 (0.03)	1		
Wii U	-0.63 (0.04)	0.93 (0.03)	0.96 (0.02)	0.93 (0.03)	0.90 (0.05)	0.96 (0.01)	0.96 (0.03)	0.93 (0.02)	0.96 (0.02)	0.96 (0.02)	0.96 (0.02)	1	
Wii Wheel	-0.61 (0.05)	0.90 (0.06)	0.92 (0.04)	0.89 (0.06)	0.89 (0.06)	0.93 (0.03)	0.95 (0.02)	0.91 (0.03)	0.94 (0.02)	0.94 (0.02)	0.92 (0.05)	0.93 (0.04)	1
Wii Motion	-0.61 (0.04)	0.91 (0.04)	0.94 (0.02)	0.91 (0.04)	0.89 (0.05)	0.94 (0.03)	0.95 (0.03)	0.92 (0.03)	0.95 (0.02)	0.95 (0.02)	0.94 (0.02)	0.95 (0.02)	0.92 (0.03)

Table 11: Posterior Means of Individual Θ (posterior SD) - MCMhet

	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel
Xbox One	-4.91 (0.18)	0										
Xbox Kinect	1.15 (0.12)	1.04 (0.1)	0									
Xbox Wheel	1.23 (0.10)	1.14 (0.09)	-1.44 (0.13)	0								
PS3	-4.82 (0.14)	-4.93 (0.17)	-1.45 (0.12)	-1.52 (0.13)	0							
PS4	-4.86 (0.15)	-4.96 (0.16)	-1.43 (0.12)	-1.41 (0.13)	-5.08 (0.13)	0						
PS Eye	-1.49 (0.15)	-1.50 (0.15)	-1.46 (0.13)	-1.41 (0.15)	0.95 (0.12)	1.15 (0.11)	0					
PS Move	-1.45 (0.13)	-1.55 (0.15)	-1.49 (0.14)	-1.43 (0.14)	1.02 (0.09)	1.26 (0.09)	-1.63 (0.13)	0				
PS Wheel	-1.53 (0.12)	-1.48 (0.13)	-1.51 (0.13)	-1.47 (0.12)	1.06 (0.09)	1.20 (0.11)	-1.38 (0.14)	-1.39 (0.12)	0			
Wii	-4.79 (0.16)	-4.91 (0.17)	-1.43 (0.13)	-1.49 (0.13)	-4.83 (0.15)	-4.79 (0.16)	-1.38 (0.12)	-1.52 (0.13)	-1.45 (0.13)	0		
Wii U	-4.89 (0.17)	-4.87 (0.15)	-1.46 (0.14)	-1.31 (0.12)	-4.97 (0.13)	-4.81 (0.16)	-1.47 (0.12)	-1.46 (0.13)	-1.44 (0.16)	-4.99 (0.15)	0	
Wii Wheel	-1.59 (0.13)	-1.34 (0.12)	-1.47 (0.13)	-1.49 (0.12)	-1.50 (0.16)	-1.44 (0.14)	-1.48 (0.13)	-1.45 (0.14)	-1.53 (0.12)	1.02 (0.10)	1.11 (0.10)	0
Wii Motion	-1.45 (0.14)	-1.50 (0.15)	-1.58 (0.15)	-1.45 (0.13)	-1.46 (0.11)	-1.36 (0.12)	-1.45 (0.12)	-1.43 (0.12)	-1.40 (0.12)	1.18 (0.10)	1.25 (0.09)	-1.36 (0.12)

Table 12: Posterior Means of Individual Θ (posterior SD) - MCMhom

	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel
Xbox One	-2.74 (0.13)	0										
Xbox Kinect	1.57 (0.07)	0.87 (0.10)	0									
Xbox Wheel	2.32 (0.05)	1.84 (0.05)	-0.06 (0.13)	0								
PS3	-1.10 (0.11)	-0.97 (0.05)	-1.33 (0.05)	-0.25 (0.05)	0							
PS4	-1.08 (0.1)	-1.72 (0.04)	-0.97 (0.05)	-0.13 (0.04)	-3.03 (0.05)	0						
PS Eye	-0.70 (0.04)	-0.10 (0.07)	-0.28 (0.09)	0.22 (0.1)	0.43 (0.08)	1.41 (0.08)	0					
PS Move	-1.28 (0.06)	-0.73 (0.16)	-0.67 (0.10)	-0.29 (0.06)	0.89 (0.06)	1.73 (0.09)	-0.9 (0.07)	0				
PS Wheel	-0.47 (0.08)	0.16 (0.06)	-0.17 (0.09)	-0.09 (0.08)	1.19 (0.05)	1.69 (0.05)	0.84 (0.13)	0.44 (0.09)	0			
Wii	-0.94 (0.04)	-0.95 (0.08)	-0.08 (0.09)	-0.38 (0.05)	-1.37 (0.05)	-0.54 (0.05)	-0.99 (0.08)	-0.48 (0.09)	-0.39 (0.04)	0		
Wii U	-1.23 (0.13)	-0.85 (0.08)	-0.74 (0.14)	0.29 (0.08)	-1.66 (0.13)	-0.91 (0.06)	-0.62 (0.05)	-0.03 (0.09)	0.28 (0.08)	-2.18 (0.11)	0	
Wii Wheel	-0.63 (0.11)	-0.49 (0.04)	-0.32 (0.07)	-0.18 (0.05)	-0.40 (0.04)	-0.64 (0.04)	0.47 (0.10)	-0.73 (0.06)	-0.64 (0.12)	0.57 (0.09)	0.94 (0.13)	0
Wii Motion	-0.06 (0.08)	-0.53 (0.04)	-1.05 (0.03)	-0.11 (0.05)	-0.70 (0.04)	-0.43 (0.07)	-0.27 (0.05)	-0.81 (0.10)	-0.12 (0.08)	1.22 (0.05)	2.01 (0.07)	0.53 (0.05)

Table 13: Correlation Matrix of Error Terms (posterior SD) - MvP

	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel
Xbox One	-0.03 (0.05)	1										
Xbox Kinect	0.51 (0.05)	0.38 (0.05)	1									
Xbox Wheel	0.45 (0.06)	0.47 (0.06)	0.24 (0.07)	1								
PS3	-0.17 (0.07)	-0.15 (0.08)	-0.18 (0.07)	-0.19 (0.06)	1							
PS4	-0.27 (0.06)	-0.34 (0.08)	-0.33 (0.04)	-0.13 (0.06)	-0.36 (0.05)	1						
PS Eye	-0.30 (0.05)	-0.29 (0.06)	-0.18 (0.09)	-0.23 (0.07)	0.09 (0.06)	0.40 (0.05)	1					
PS Move	-0.37 (0.05)	-0.21 (0.09)	-0.27 (0.05)	-0.12 (0.08)	0.21 (0.07)	0.53 (0.04)	-0.04 (0.05)	1				
PS Wheel	-0.24 (0.06)	-0.14 (0.08)	-0.32 (0.08)	-0.08 (0.08)	0.20 (0.06)	0.50 (0.05)	0.33 (0.04)	0.35 (0.04)	1			
Wii	-0.02 (0.05)	-0.05 (0.04)	-0.05 (0.07)	-0.25 (0.06)	-0.24 (0.06)	-0.12 (0.08)	-0.33 (0.06)	-0.19 (0.08)	-0.29 (0.07)	1		
Wii U	-0.09 (0.06)	-0.06 (0.08)	-0.15 (0.09)	-0.09 (0.07)	-0.36 (0.08)	0.00 (0.07)	-0.10 (0.06)	-0.14 (0.06)	-0.18 (0.10)	-0.15 (0.05)	1	
Wii Wheel	-0.11 (0.08)	0.06 (0.07)	-0.07 (0.08)	-0.10 (0.07)	-0.11 (0.06)	-0.27 (0.07)	-0.14 (0.07)	-0.28 (0.06)	-0.17 (0.07)	0.39 (0.05)	0.39 (0.05)	1
Wii Motion	-0.22 (0.06)	0.05 (0.07)	-0.26 (0.08)	-0.20 (0.08)	-0.27 (0.06)	-0.19 (0.05)	-0.30 (0.05)	-0.19 (0.05)	-0.13 (0.07)	0.48 (0.04)	0.51 (0.04)	0.46 (0.06)

Table 14: Posterior Means of Cross-Price Effects (posterior SD) - IndepCPE

	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel	Wii Motion
Xbox 360 →	-	-	-1.03 (0.07)	-0.57 (0.11)	-	-	-	-	-	-	-	-	-
Xbox One →	-	-	-	-0.66 (0.05)	-	-0.14 (0.02)	-	-	-	-	-	-	-
Xbox Kinect →	-0.57 (0.09)	-	-	-	-	-	-	-	-	-	-	-	-
Xbox Wheel →	-0.46 (0.24)	-	-	-	-	-	-	-	-	-	-	-	-
PS3 →	-0.41 (0.15)	-	-	-	-	-0.17 (0.03)	-0.41 (0.08)	-0.42 (0.09)	-0.30 (0.10)	-	-	-	-
PS4 →	-	-0.45 (0.04)	-	-	-0.34 (0.04)	-	-0.62 (0.05)	-0.66 (0.05)	-0.77 (0.09)	-	-0.37 (0.06)	-	-
PS Eye →	-	-	-	-	-	-	-	-0.17 (0.16)	-	-	-	-	-
PS Move →	-	-	-	-	-	-	-	-	-	-	-	-	-
PS Wheel →	-	-	-	-	-0.55 (0.19)	-	-	-0.31 (0.19)	-	-	-	-	-
Wii →	-0.45 (0.07)	-	-	-	-0.26 (0.08)	-	-	-	-	-	-0.50 (0.10)	-0.35 (0.06)	-0.32 (0.07)
Wii U →	-	-	-	-	-	-	-	-	-	-0.57 (0.03)	-	-0.62 (0.04)	-0.53 (0.04)
Wii Wheel →	-	-	-	-	-	-	-	-	-	-	-	-	-
Wii Motion →	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 15: Heterogeneity of Individual Θ (posterior SD) - MCMhet

	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel
Xbox One	2.91 (0.09)	0	-	-	-	-	-	-	-	-	-	-
Xbox Kinect	2.26 (0.09)	2.18 (0.09)	0	-	-	-	-	-	-	-	-	-
Xbox Wheel	2.31 (0.08)	2.24 (0.08)	2.88 (0.11)	0	-	-	-	-	-	-	-	-
PS3	2.95 (0.08)	2.88 (0.08)	2.85 (0.09)	2.90 (0.10)	0	-	-	-	-	-	-	-
PS4	2.92 (0.09)	2.88 (0.09)	2.80 (0.11)	2.84 (0.09)	2.82 (0.08)	0	-	-	-	-	-	-
PS Eye	2.84 (0.11)	2.88 (0.12)	2.87 (0.10)	2.80 (0.12)	2.08 (0.10)	2.24 (0.09)	0	-	-	-	-	-
PS Move	2.85 (0.10)	2.87 (0.11)	2.85 (0.10)	2.86 (0.10)	2.15 (0.09)	2.33 (0.08)	2.93 (0.09)	0	-	-	-	-
PS Wheel	2.89 (0.09)	2.84 (0.10)	2.86 (0.10)	2.91 (0.09)	2.18 (0.09)	2.29 (0.09)	2.92 (0.10)	2.87 (0.10)	0	-	-	-
Wii	2.96 (0.08)	2.90 (0.10)	2.87 (0.11)	2.89 (0.10)	2.94 (0.07)	2.96 (0.08)	2.81 (0.11)	2.86 (0.09)	2.88 (0.11)	0	-	-
Wii U	2.91 (0.08)	2.92 (0.08)	2.88 (0.10)	2.79 (0.11)	2.87 (0.08)	2.95 (0.08)	2.85 (0.10)	2.88 (0.10)	2.86 (0.10)	2.86 (0.08)	0	-
Wii Wheel	2.91 (0.09)	2.80 (0.10)	2.88 (0.10)	2.89 (0.09)	2.87 (0.11)	2.83 (0.10)	2.91 (0.10)	2.83 (0.12)	2.90 (0.09)	2.15 (0.09)	2.23 (0.09)	0
Wii Motion	2.85 (0.10)	2.89 (0.10)	2.91 (0.10)	2.83 (0.10)	2.85 (0.10)	2.79 (0.10)	2.84 (0.10)	2.82 (0.09)	2.86 (0.09)	2.27 (0.08)	2.33 (0.07)	2.88 (0.10)

Table 16: Heterogeneity in Cross-Price Effects (posterior SD) - IndepCPE

	Xbox 360	Xbox One	Xbox Kinect	Xbox Wheel	PS3	PS4	PS Eye	PS Move	PS Wheel	Wii	Wii U	Wii Wheel	Wii Motion
Xbox 360 →	-	-	0.81 (0.04)	0.70 (0.05)	-	-	-	-	-	-	-	-	-
Xbox One →	-	-	-	0.57 (0.03)	-	0.55 (0.02)	-	-	-	-	-	-	-
Xbox Kinect →	0.79 (0.05)	-	-	-	-	-	-	-	-	-	-	-	-
Xbox Wheel →	1.25 (0.12)	-	-	-	-	-	-	-	-	-	-	-	-
PS3 →	0.78 (0.07)	-	-	-	-	0.65 (0.04)	0.67 (0.04)	0.81 (0.06)	0.69 (0.05)	-	-	-	-
PS4 →	-	0.64 (0.02)	-	-	0.59 (0.03)	-	0.61 (0.04)	0.68 (0.04)	0.68 (0.06)	-	0.61 (0.05)	-	-
PS Eye →	-	-	-	-	-	-	-	1.19 (0.12)	-	-	-	-	-
PS Move →	-	-	-	-	-	-	-	-	-	-	-	-	-
PS Wheel →	-	-	-	-	1.36 (0.14)	-	-	1.34 (0.14)	-	-	-	-	-
Wii →	0.75 (0.04)	-	-	-	0.73 (0.07)	-	-	-	-	-	0.74 (0.06)	0.64 (0.04)	0.61 (0.04)
Wii U →	-	-	-	-	-	-	-	-	-	0.66 (0.03)	-	0.62 (0.03)	0.60 (0.03)
Wii Wheel →	-	-	-	-	-	-	-	-	-	-	-	-	-
Wii Motion →	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 17: Demand Interactions between Game Consoles and Accessories of the Same Brand (posterior SD)

	MCMhet	MCMhom	MvP	IndepCPE	
				→	←
Xbox 360-Xbox Kinect	1.15 (0.12)	1.57 (0.07)	0.51 (0.05)	-1.03 (0.07)	-0.57 (0.09)
Xbox 360-Xbox Wheel	1.23 (0.10)	2.32 (0.05)	0.45 (0.06)	-0.57 (0.11)	-0.46 (0.24)
Xbox One-Xbox Kinect	1.04 (0.10)	0.87 (0.10)	0.38 (0.05)	-	-
Xbox One-Xbox Wheel	1.14 (0.09)	1.84 (0.05)	0.47 (0.06)	-0.66 (0.05)	-
PS3-PS Eye	0.95 (0.12)	0.43 (0.08)	0.09 (0.06)	-0.41 (0.08)	-
PS3-PS Move	1.02 (0.09)	0.89 (0.06)	0.21 (0.07)	-0.42 (0.09)	-
PS3-PS Wheel	1.06 (0.09)	1.19 (0.05)	0.20 (0.06)	-0.30 (0.10)	-0.55 (0.19)
PS4-PS Eye	1.15 (0.11)	1.41 (0.08)	0.40 (0.05)	-0.62 (0.05)	-
PS4-PS Move	1.26 (0.09)	1.73 (0.09)	0.53 (0.04)	-0.66 (0.05)	-
PS4-PS Wheel	1.20 (0.11)	1.69 (0.05)	0.50 (0.05)	-0.77 (0.09)	-
Wii-Wii Wheel	1.02 (0.1)	0.57 (0.09)	0.39 (0.05)	-0.35 (0.06)	-
Wii-Wii Motion	1.18 (0.10)	1.22 (0.05)	0.48 (0.04)	-0.32 (0.07)	-
Wii U-Wii Wheel	1.11 (0.10)	0.94 (0.13)	0.39 (0.05)	-0.62 (0.04)	-
Wii U-Wii Motion	1.25 (0.09)	2.01 (0.07)	0.51 (0.04)	-0.53 (0.04)	-

References

Propp, J. G. and Wilson, D. B. Exact sampling with coupled markov chains and applications to statistical mechanics. Random Structures & Algorithms, 9(1-2):223–252, 1996.