Appendix

MCMC Algorithm

For estimation we use a Markov Chain Monte Carlo algorithm of the Gibbs sampler form. Below we describe each conditional distribution making up the Gibbs sampler. We ran the Markov Chain for a total of 110,000 iterations, dropping the first 10,000 to allow for burn-in and then retaining every 10th of the next 100,000 iterations for analysis. The resulting 10,000 draws passed all standard convergence checks.

Sampling U^O_{ict}, U^K_{ict}

U^O_{ict} and U^K_{ict} are sampled according to the purchase behavior of customer i in period t in category c. There are three possible outcomes:

1. **No order**: Here we only need to sample U^P_{ict}:

$$U^O_{ict} \sim \text{TN}_{(-\infty,0)}(-\bar{\omega}(R_i, \sigma_{\psi c}) + \mu_{ict}, 1),$$

   where $\text{TN}_{a,b}(\mu, \sigma^2)$ denotes the normal distribution with mean μ and σ^2 truncated to the region (a, b).

2. **Order and return:** In this case we sample the utilities as

$$U^O_{ict}|U^K_{ict} \sim \text{TN}_{(0,\infty)}\Big(-\bar{\omega}(R_i, \sigma_{\psi c}) + \mu_{ict} + \frac{1}{1+\sigma^2_{\psi c}}(U^K_{ict} - \mu_{ict} - R_i), \frac{\sigma^2_{\psi c}}{1+\sigma^2_{\psi c}}\Big),$$
$$U^K_{ict}|U^O_{ict} \sim \text{TN}_{(-\infty,0)}\Big(R_i + \bar{\omega}(R_i, \sigma_{\psi c}) + U^O_{ict}, \sigma^2_{\psi c}\Big).$$

3. **Order and keep:** In this case we sample the utilities as

$$U^O_{ict}|U^K_{ict} \sim \text{TN}_{(0,\infty)}\Big(-\bar{\omega}(R_i, \sigma_{\psi c}) + \mu_{ict} + \frac{1}{1+\sigma^2_{\psi c}}(U^K_{ict} - \mu_{ict} - R_i), \frac{\sigma^2_{\psi c}}{1+\sigma^2_{\psi c}}\Big),$$
$$U^K_{ict}|U^O_{ict} \sim \text{TN}_{(0,\infty)}\Big(R_i + \bar{\omega}(R_i, \sigma_{\psi c}) + U^O_{ict}, \sigma^2_{\psi c}\Big).$$

Sampling $\theta_i = (\beta_{i1}, \beta_{i2}, \beta_{i3}, \log R_i)$

Recall from the text that $\mu_{ict} \equiv \beta'_{ic}X_{ict}$. We sample θ_i in blocks by sampling $\beta_{ic}|\theta_{-\beta_{ic},i}$, $c = 1, 2, 3$ and $\log R_i|\theta_{-\log R_i,i}$. Since

$$U^O_{ict} + \bar{\omega}(R_i, \sigma_{\psi c}) = \beta'_{ic}X_{ict} + \epsilon^0_{ict},$$

the conditional for β_{ic} is normal with precision and mean

$$p_{ic} = \big(\sum X_{ict}X'_{ict} + \Lambda_{\beta_{ic}}\big)^{-1}$$
$$\mu_{ic} = p_{ic}\Big(\big(\sum X'_{ict}[U^O_{it} + \bar{\bar{\omega}}(R_i, \sigma_{\psi c})]\big) + \Lambda_{\beta_{ic}}\mu_{\beta_{ic}}\Big),$$

where $\Lambda_{\beta_{ic}}$ and $\mu_{\beta_{ic}}$ is the precision and mean of the prior distribution of β_{ic} conditional on $\beta_{ic'}$, $c' \neq c$ and $\log R_i$.

The conditional for $\tilde{R}_i \equiv \log R_i$ is

$$p(\tilde{R}_i|.) \propto \prod_{c=1}^{3} \Big\{ \prod_{t=1}^{T_{ic}} \phi\big(U^O_{ict}| - \bar{\omega}(\exp\{\tilde{R}_i\}, \sigma_{\psi c}) + \mu_{it}, 1\big) \times \prod_{t:D^O_{ict}=1} \phi\big(U^K_{ict}| \exp\{\tilde{R}_i\} + \bar{\omega}(\exp\{\tilde{R}_i\}, \sigma_{\psi c}) + U^O_{ict}, \sigma^2_{\psi c}\big)\Big\} \times \phi\big(\tilde{R}_i|\mu_{\tilde{R}_i}, \Lambda_{\tilde{R}_i}\big), \quad (1)$$

where $D^O_{ict} = 1$ indicates that an order was placed. $\Lambda_{\tilde{R}_i}$ and $\mu_{\tilde{R}_i}$ is the precision and mean of the prior distribution of $\log R_i$ conditional on β_{ic}, $c = 1, 2, 3$. Here we use a Metropolis-Hastings step centered at the previous draw of $\tilde{R}_h$ with a normally distributed step size, tuned to get a reasonable accept rate.

<u>SAMPLING $\sigma_{\psi c}$</u>

The conditional for $\tau_{\psi c} = 1/\sigma^2_{\psi c}$ is

$$p(\tau_{\psi c}|.) \propto p(\tau_{\psi c}) \times \prod_{i=1}^{N} L_i(\tau_{\psi c}), \quad (2)$$

where

$$L_i(\tau_{\psi c}) = \prod_{t:D^O_{ict}=1} \phi\big(U^K_{ict}|R_i + \bar{\omega}(R_i, \tau_{\psi c}) + U^O_{ict}, \tau^{-1}_{\psi c}\big) \times \prod_{t=1}^{T_{ic}} \phi\big(U^O_{ict}| - \bar{\omega}(R_i, \tau_{\psi c}) + \mu_{it}, 1\big) \quad (3)$$

and $p(\tau_\psi)$ is the prior distribution of τ_ψ. We use a gamma distribution $G(\alpha = 2, \beta = 1)$ as a prior for τ_ψ. We change the parametrization to the log scale, $\log \tau_\psi$, and sample $\log \tau_\psi$ using a random walk MH step centered on the previous value of $\log \tau_\psi$ with a normally distributed step size, tuned to get a reasonable accept rate.

<u>SAMPLING $\bar{\beta}, \bar{R}, \Omega$</u>

This is simply multivariate regression and follows standard Gibbs updates for multivariate regression, see Rossi, Allenby and McCulloch (2005). We use a Wishart prior for Ω^{-1}, $W(\nu, S)$, with parameters $\nu = 9$, $S = (1/\nu)I_7$. We use a normal prior for $(\bar{\beta}, \bar{R})$ with mean vector zero and variance $I_7 * 100$.

Price elasticities

The order-and-keep probability is

$$\Pr(U_{it}^O > 0, U_{it}^K > 0 | R_i, \mu_{it}) = \int_{-\frac{(-\bar{\omega}(R_i,\sigma_\psi)+\mu_{it})}{\sigma_\varepsilon}}^{\infty} \Phi\Big(\frac{R_i + \mu_{it} + \sigma_\varepsilon \varepsilon}{\sigma_\psi}\Big)\phi(\varepsilon)d\varepsilon$$

Using Leibnitz's rule we get

$$\begin{aligned}\frac{\partial \Pr(U_{it}^O > 0, U_{it}^K > 0 | R_i, \mu_{it})}{\partial p} &= \Phi\Big(\frac{R+\bar{\omega}}{\sigma_\psi}\Big)\phi\Big(\frac{-\bar{\omega}(R_i,\sigma_\psi)+\mu_{it}}{\sigma_\varepsilon}\Big)\frac{\beta_i^p}{\sigma_\varepsilon} + \\ &\int_{-\frac{(-\bar{\omega}(R_i,\sigma_\psi)+\mu_{it})}{\sigma_\varepsilon}}^{\infty} \phi\Big(\frac{R_i + \mu_{it} + \sigma_\varepsilon \varepsilon}{\sigma_\psi}\Big)\phi(\varepsilon)\frac{\beta_i^p}{\sigma_\psi}d\varepsilon\end{aligned} \tag{4}$$

The integral in the second part of this equation can be solved to give

$$\int_{-\frac{(-\bar{\omega}(R_i,\sigma_\psi)+\mu_{it})}{\sigma_\varepsilon}}^{\infty} \phi\Big(\frac{R_i + \mu_{it} + \sigma_\varepsilon \varepsilon}{\sigma_\psi}\Big)\phi(\varepsilon)\frac{\beta_i^p}{\sigma_\psi}d\varepsilon = \frac{\beta_i^p}{\bar{\sigma}}\phi\Big(\frac{R+\mu_{it}}{\bar{\sigma}}\Big)\Phi\Big(\frac{(-\bar{\omega}+\mu_{it})/\sigma_\varepsilon + \mu_\varepsilon}{\sigma_\psi/\bar{\sigma}}\Big), \tag{5}$$

where

$$\begin{aligned}\bar{\sigma} &= \sqrt{\sigma_\psi^2 + \sigma_\varepsilon^2}, \\ \mu_\varepsilon &= -\frac{\sigma_\varepsilon}{\bar{\sigma}}(R_i + \mu_{it}).\end{aligned}$$

Collecting terms we find the marginal effect of a price change on the order-and-keep probability to be

$$\begin{aligned}\frac{\partial \Pr(U_{it}^O > 0, U_{it}^K > 0 | R_i, \mu_{it})}{\partial p} =& \beta_i \Big\{ \Phi\Big(\frac{R+\bar{\omega}}{\sigma_\psi}\Big)\phi\Big(\frac{-\bar{\omega}+\mu_{it}}{\sigma_\varepsilon}\Big)\frac{1}{\sigma_\varepsilon} + \\ &\frac{1}{\bar{\sigma}}\phi\Big(\frac{R+\mu_{it}}{\bar{\sigma}}\Big)\Phi\Big(\frac{(-\bar{\omega}+\mu_{it})/\sigma_\varepsilon + \mu_\varepsilon}{\sigma_\psi/\bar{\sigma}}\Big)\Big\}\end{aligned} \tag{6}$$

The marginal price effect on the order probability is

$$\frac{\partial \Pr(U_{it}^O > 0 | R_i, \mu_{it})}{\partial p} = \beta_i^p \phi\Big(\frac{-\bar{\omega}+\mu_{it}}{\sigma_\varepsilon}\Big)\frac{1}{\sigma_\varepsilon} \tag{7}$$

Finally, since

$$1 = \Pr((U_{it}^O > 0, U_{it}^K > 0 | R_i, \mu_{it}) + \Pr((U_{it}^O > 0, U_{it}^K < 0 | R_i, \mu_{it}) + 1 - \Pr(U_{it}^O > 0 | R_i, \mu_{it})$$

we have that the marginal effect of a price change on the order-and-return probability is

$$\frac{\partial \Pr(U_{it}^O > 0, U_{it}^K < 0 | R_i, \mu_{it})}{\partial p} = \frac{\partial \Pr(U_{it}^O > 0 | R_i, \mu_{it})}{\partial p} - \frac{\partial \Pr(U_{it}^O > 0, U_{it}^K > 0 | R_i, \mu_{it})}{\partial p}. \tag{8}$$

Having computed the marginal effects it is now straightforward to compute elasticities by appropriate scaling by the base probability and price.

Adding Competition

In this section we show how the model in the text is modified when customers consider purchases from two retailers. Assume that items can be ordered from two competing retailers, A and B. Let net utility from keeping an ordered item from retailer j be

$$U_{ijt}^K = \mu_{ijt} + \psi_{it} + R_{ij} + \varepsilon_{ijt}, \qquad j = A, B. \tag{9}$$

Since the item is assumed to be identical for the two retailers ψ_{it} has no j-subscript. Retailer/item preference is captured by μ_{ijt} and ε_{ijt} while R_{ij} is return cost for retailer j.

Following the derivation in the main text, we find that expected utility of ordering from retailer j is

$$\mathrm{E}[U_{ijt}(\text{order})] = H(\mu_{ijt} + \varepsilon_{ijt}, R_{ij}, \sigma_\psi), \qquad j = A, B, \tag{10}$$

where the H-function is as defined in the text. The three possible order outcomes are

$$\begin{aligned}
\mathrm{E}[U_{iAt}(\text{order})] > \mathrm{E}[U_{iBt}(\text{order})],\ \mathrm{E}[U_{iAt}(\text{order})] > 0 &\implies \text{order from } A,\\
\mathrm{E}[U_{iBt}(\text{order})] > \mathrm{E}[U_{iAt}(\text{order})],\ \mathrm{E}[U_{iBt}(\text{order})] > 0 &\implies \text{order from } B,\\
\mathrm{E}[U_{iAt}(\text{order})] < 0, \mathrm{E}[U_{iBt}(\text{order})] < 0 &\implies \text{no order.}
\end{aligned}$$

Note that in contrast to the monopolist retailer model we can no longer transform the model into a standard linear latent utility model. For example, the event "order from A and return item" corresponds to the event

$$\begin{aligned}
H(\mu_{iAt} + \varepsilon_{iAt}, R_{iA}, \sigma_\psi) &> H(\mu_{iBt} + \varepsilon_{iBt}, R_{iB}, \sigma_\psi),\\
H(\mu_{iAt} + \varepsilon_{iAt}, R_{iA}, \sigma_\psi) &> 0,\\
U_{iAt}^K = \mu_{iAt} + \psi_{it} + R_{iA} + \varepsilon_{iAt} &< 0.
\end{aligned}$$

It is not possible to transform this into a model comparable to the one in the text. This does not mean that this model cannot be estimated – it can (although the added complexity does make estimation more cumbersome).

A more fundamental problem is that the above model requires purchase and return data for two competing retailers. We do not have this data. This raises the question whether a partial equilibrium analysis can be done. We address this in the following section.

Partial Equilibrium Analysis

If we only observe purchase and return data for retailer A, then utility derived from orders (and returns) from retailer B will be in the "no-purchase" option for A. In other words, the impact of changes in retailer B's return policy will affect the no purchase option.

Let the utility of ordering and keeping an item from retailer A be

$$U_{it}^{A} = \mu_{it}^{A} + \psi_{it} + \varepsilon_{it}^{A}, \tag{11}$$

and let utility from the outside option be

$$U_{it}^{out} = \lambda_{0i} + \lambda_i' Z_{it} + \varepsilon_{it}^{out}. \tag{12}$$

The outside option captures utility from not buying from A, i.e., either not buying at all in that time period, or buying from B. We let Z_{it} capture changes in B's return policy. The net utility of ordering and keeping an item from A is then

$$U_{it}^{K} \equiv U_{it}^{A} - U_{it}^{out} = \mu_{it}^{A} - \lambda_i' Z_{it} + \psi_{it} + \varepsilon_{it}^{A} - \varepsilon_{it}^{out}, \tag{13}$$

where we have included the intercept λ_{0i} into the intercept in μ_{it}^{A}. Now, if we observe changes in B's return policy (or other competitors return policy), we can estimate the impact by using a dummy in Z_{it} to capture the return policy change. This would show up as a intercept change in the net utility of ordering and keeping an item from A.

For our data we do know the identity of the main competitor to our focal retailer. However, neither the focal retailer (for which we have data) nor the main competitor changed return policies in the observation period. This means that Z_{it} was constant throughout the observation, and this makes it impossible to separately identify the competitive effect λ_i. These parameters must be swept into the intercept of μ_{it}^{A} leading to the specification used in the paper:

$$U_{it}^{K} = \mu_{it} + \psi_{it} + \varepsilon_{it}. \tag{14}$$

Note that return policies in a market are likely to change much less frequently than prices. This makes it very difficult to identify competitive effects of return policies. How to do this is an open research question.