Technical Appendix for

Consumer Learning and Evolution of Consumer Brand Preferences

Appendix I. Consumer Learning

In this appendix, we provide the technical details for the equations in the learning model, discussed in Section II in the main text of the paper. For the ease of cross-reference, we use the same equation numbers in both the main document and this appendix.

We assume for each household $h \in \mathrm{H}$, brand $j \in \mathrm{J}$ and size $k \in \mathrm{K}$ that the initial perception errors υ_{h0jk} are correlated across sizes and their correlation matrix is given by:

$$R = \begin{bmatrix} 1 & \rho_{12} & 0 & . & . & . & . & 0 \\ . & 1 & \rho_{23} & . & . & . & . & 0 \\ . & . & 1 & \rho_{34} & . & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & 1 & \rho_{K-2,K-1} & 0 \\ . & . & . & . & . & . & 1 & \rho_{K-1,K} \\ . & . & . & . & . & . & . & 1 \end{bmatrix}$$

where $\rho_{h0kl} = \rho_{kl}$ are the initial correlations between sizes *k* and *l*. This specification indicates that only the adjacent sizes are correlated and that the initial size correlations are uniform across brands and households, by assumption. We denote by ρ_{htkl} the time $t \geq 1$ correlation coefficients between the sizes of the brands, which are updated over time as we describe below.

Since υ_{htjk} are correlated across sizes, we have the following relationships across different sizes of the brands for any $t \geq 1$:

$$Q_{Ehtjk} = \kappa_{htjk|l} Q_{Ehtjl} + \eta_{htjk|l} 1_{\{l \neq k\}}, \quad \eta_{htjk|l} \sim N\left(0, \sigma^2_{\eta_{htjk|l}}\right) \tag{1}$$

In mathematical terms, this is the linear projection of one vector on another, and κ can be thought of the ordinary least square coefficient for Q_{Ehtjl}. More specifically,

$$\kappa_{htjk|l} = \frac{\operatorname{cov}\left(Q_{Ehtjk}, Q_{Ehtjl}\right)}{\operatorname{var}\left(Q_{Ehtjl}\right)}$$

Obviously, when $k=l$, $\kappa_{htjk|l}=1$;when $k\neq l$,

$$\begin{aligned}
\operatorname{var}\left(Q_{Ehtjk}\right) &= \sigma^2_{\upsilon_{h,t-1,jl}} + \sigma^2_{\xi} 1_{\{l\neq k\}} \\
\operatorname{cov}\left(Q_{Ehtjk}, Q_{Ehtjl}\right) &= \operatorname{cov}\left(\upsilon_{h,t-1,jk}, \upsilon_{h,t-1,jl}\right) \quad \text{(since only } \upsilon\text{'s are correlated)} \\
&= \sigma_{\upsilon_{h,t-1,jk}} \sigma_{\upsilon_{h,t-1,jl}} \rho_{h,t-1,kl}
\end{aligned}$$

Therefore,

$$\kappa_{htjk|l} = \frac{\sigma_{\upsilon_{h,t-1,jk}} \sigma_{\upsilon_{h,t-1,jl}} \rho_{h,t-1,kl}}{\sigma^2_{\upsilon_{h,t-1,jl}} + \sigma^2_{\xi} 1_{\{l\neq k\}}} \tag{2}$$

To calculate the variance of $\eta_{htjk|l}$,

$$\begin{aligned}
\operatorname{var}\left(\eta_{htjk|l}\right) &= \operatorname{var}\left(Q_{Ehtjk} - \kappa_{htjk|l} Q_{Ehtjl}\right) \\
&= E\left[Q_{Ehtjk} - \kappa_{htjk|l} Q_{Ehtjl} - E\left(Q_{Ehtjk} - \kappa_{htjk|l} Q_{Ehtjl}\right)\right]^2 \\
&= E\left[\xi_{htjk} - \upsilon_{h,t-1,jk} - \kappa_{htjk|l}\left(\xi_{htjl} - \upsilon_{h,t-1,jl}\right)\right]^2
\end{aligned}$$

Through some algebra (and by substituting Equation (5) into the above equation), we have Equation (6):

$$\sigma^2_{\eta_{htjk|l}} = \frac{\left(\sigma^2_{\upsilon_{h,t-1,jk}} + \sigma^2_{\xi} 1_{\{l\neq k\}}\right)\left(\sigma^2_{\upsilon_{h,t-1,jl}} + \sigma^2_{\xi} 1_{\{l\neq k\}}\right) - \sigma^2_{\upsilon_{h,t-1,jk}} \sigma^2_{\upsilon_{h,t-1,jl}} \rho^2_{h,t-1,kl}}{\sigma^2_{\upsilon_{h,t-1,jl}} + \sigma_{\xi}{}^2 1_{\{l\neq k\}}} \tag{3}$$

whereas $1_{\{A\}}$ is the indicator function that returns 1 if the statement A is true and 0 otherwise.

It is clear from equations (4) and (5) that as long as ρ_{htkl} are non-zero, use-experience in a particular size of a brand provides information for other sizes of the same brand as well. [1] Our approach could be employed to study the spill-over effects of signals from one domain to another in the context of forward-looking consumers. As previously

[1] In mathematical terms, this is the linear projection of one vector on another.

indicated, all previous learning models with spill-over effects have assumed myopic consumers, with the exception of a working paper by Dickenson (2011), whose method requires risk-neutrality. Our model allows for risk-aversion as well, which has been shown to hold empirically in previous learning papers. Furthermore, it is clear from equation (4) also that $0<\left|\kappa_{htjk|l}\right|<1$ when $k \neq l$ whereas $\kappa_{htjk|k}=1$ otherwise. This implies that the noisy information provided by the use experience in a different size of a brand is less than the information provided by the use experience in the current size.

Let us now denote by J_{ht} and K_{ht} the respective brand and size purchases of household *h* on purchase occasion *t*. With this definition and the above relationships between the sizes *k* and *l*, we are now ready to describe how the consumers update their perceived qualities in our model. We assume after buying size *l* of brand *j* on purchase occasion *t*, household *h* updates the priors about the mean quality of size *k* of brand *j* using Bayesian updating rules (see, for example, DeGroot 1970). We can describe the updating process in three steps:

Step (1): when $J_{ht} \neq j$, then nothing is learned for brand *j*, both perceived quality and perception variability remain the same as those from the last purchase;

Step (2): when $J_{ht}=j$ and $K_{ht}=k=n$, we have the following updating equations for $t>1$:

$$Q_{htjk} = Q_{h,t-1,jk} + \beta\left(Q_{Ehtjk} - Q_{h,t-1,jk}\right)$$
$$\upsilon_{htjk} = \upsilon_{h,t-1,jk} + \beta\left(\xi_{htjk} - \upsilon_{h,t-1,jk}\right)$$
$$\upsilon_{htjk} - \upsilon_{h,t-1,jk} = \beta\left(\xi_{htjk} - \upsilon_{h,t-1,jk}\right)$$

Now we multiply both sides by $\xi_{htjk} - \upsilon_{h,t-1,jk}$, and take the expectation, we get,

$$\beta = \frac{\sigma^2_{\upsilon_{h,t-1,jk}}}{\sigma^2_{\upsilon_{h,t-1,jk}} 1_{\{J_{h,t-1} \neq j\}} + \sigma^2_{\xi}}$$

And therefore,

$$\upsilon_{htjk} = \upsilon_{h,t-1,jk} + \frac{\sigma^2_{\upsilon_{h,t-1,jk}}}{\sigma^2_{\upsilon_{h,t-1,jk}} + \sigma^2_{\xi}} \cdot \left\{\xi_{htjk} - \upsilon_{h,t-1,jk}\right\}$$

$$\sigma^2_{\upsilon_{htjk}} = \frac{\sigma^2_{\xi}\sigma^2_{\upsilon_{h,t-1,jk}}}{\sigma^2_{\upsilon_{h,t-1,jk}} + \sigma^2_{\xi}}$$

Step (3): when $J_{ht} = j$ and $K_{ht} = k \neq n$, since $Q_{Ehtjn} = \kappa_{htjn|k} Q_{Ehtjk} + \eta_{htjn|k} 1_{\{n \neq k\}}$, we can show that $Q_{htjn} = \kappa_{htjn|k} Q_{htjk}$. Insert this into the quality updating equation, and using the same trick as in step (2), we can get,

$$\beta = \frac{\kappa^2_{htjk|K_{h,t-1}} \sigma^2_{\upsilon_{h,t-1,jk}}}{\kappa^2_{htjk|K_{h,t-1}} \left(\sigma^2_{\upsilon_{h,t-1,jk}} 1_{\{J_{h,t-1} \neq j\}} + \sigma^2_{\xi}\right) + \sigma^2_{\eta_{htjk|K_{h,t-1}}} 1_{\{K_{h,t-1} \neq k\}}}$$

And therefore we have the following updating equations for $t > 1$:

$$\upsilon_{htjk} = \kappa_{htjk|K_{h,t-1}} \upsilon_{h,t-1,jk} + \frac{\kappa^2_{htjk|K_{h,t-1}} \sigma^2_{\upsilon_{h,t-1,jk}}}{\kappa^2_{htjk|K_{h,t-1}} \left(\sigma^2_{\upsilon_{h,t-1,jk}} 1_{\{J_{h,t-1} \neq j\}} + \sigma^2_{\xi}\right) + \sigma^2_{\eta_{htjk|K_{h,t-1}}} 1_{\{K_{h,t-1} \neq k\}}} \cdot \left\{\kappa_{htjk|K_{h,t-1}} \left(\xi_{htjk} - \upsilon_{h,t-1,jk}\right) + \eta_{h,t-1,jk|K_{h,t-1}} 1_{\{K_{h,t-1} \neq k\}}\right\} 1_{\{J_{h,t-1} \neq j\}} \quad (4)$$

$$\sigma^2_{\upsilon_{htjk}} = \frac{\kappa^2_{htjk|K_{h,t-1}} \left(\kappa^2_{htjk|K_{h,t-1}} \sigma^2_{\xi} + \sigma^2_{\eta_{htjk|K_{h,t-1}}} 1_{\{K_{h,t-1} \neq k\}}\right) \sigma^2_{\upsilon_{h,t-1,jk}}}{\kappa^2_{htjk|K_{h,t-1}} \left(\sigma^2_{\upsilon_{h,t-1,jk}} 1_{\{J_{h,t-1} \neq j\}} + \sigma^2_{\xi}\right) + \sigma^2_{\eta_{htjk|K_{h,t-1}}} 1_{\{K_{h,t-1} \neq k\}}} \quad (5)$$

Note that Equations (7) and (8) in Step (3) also nest the equations in Steps (1) and (2).