When Wal-Mart Enters: How Incumbent Retailers React and How This Affects Their Sales Outcomes

KUSUM L. AILAWADI, JIE ZHANG, ARADHNA KRISHNA, and MICHAEL W. KRUGER

WEB APPENDIX

ESTIMATION DETAILS OF THE THIRD STAGE ANALYSIS

Let i = category i, e = experiment store e, k = retail format k (supermarket, drug, and mass merchandiser), and τ = 1, 2 denoting the first and second six months after entry, respectively. The main model is specified as:

(A1) $S_{ie}^{\tau=2} = \alpha + X_{ie}^{\tau=2}\gamma_k + Z_{ie}\theta + \varepsilon_{ie}$,

where $S_{ie}^{\tau=2}$ is the impact on sales in category i and store e during the second six months after entry; $X_{ie}^{\tau=2}$ is a vector of the seven marketing mix reactions in category i and store e during the same time period; and Z_{ie} is a vector of the category and store characteristics. $S_{ie}^{\tau=2}$ and $X_{ie}^{\tau=2}$ are estimated as $\hat{\beta}_{3ie}^{V}$ in Equation (1). The parameter vectors γ_k contain format-specific coefficients of the marketing mix reactions, and θ is the parameter vector of Z_{ie}. The random term ε_{ie} is assumed to follow an i.i.d. normal distribution, $\varepsilon_{ie} \sim N(0, \sigma_{\varepsilon}^{2})$. As noted in the main text, sales outcomes and reactions in assortment size and regular price are scaled to make them comparable across categories and stores. The other five reaction variables are measured as percentages and thus do not need to be scaled.

We account for potential endogeneity in $X_{ie}^{\tau=2}$ by using $X_{ie}^{\tau=1}$ as instrumental variables. Following standard procedures, each endogenous explanatory variable is first predicted by the following model:

(A2) $X_{p,ie}^{\tau=2} = \eta_{p,k} + X_{p,ie}^{\tau=1}\lambda_{p,k} + Z_{ie}\mu_k + \xi_{p,ie}$

where p denotes the p-th marketing mix reaction variable in $X_{p,ie}^{\tau=2}$ and $X_{p,ie}^{\tau=1}$, the coefficients $\eta_{p,k}$ and $\lambda_{p,k}$ and coefficient vector μ_k are retail format specific, and the random term $\xi_{p,ie}$ follows a Normal distribution, $\xi_{p,ie} \sim N\left(0, \sigma_{p,\xi}^2\right)$.

To account for estimation uncertainty in $X_{p,ie}^{\tau=1}$ and $X_{p,ie}^{\tau=2}$, we use a simulated maximum likelihood method to estimate Equation (A2). The likelihood function for each marketing mix reaction variable in category i and store e is the probability density function (PDF) of $\xi_{p,ie}$:

(A3) $$L_{p,ie} = \frac{1}{\sqrt{2\pi\sigma_{p,\xi}^2}} exp\left\{-\frac{1}{2\sigma_{p,\xi}^2}[X_{p,ie}^{\tau=2} - (\eta_{p,k} + X_{p,ie}^{\tau=1}\lambda_{p,k} + Z_{ie}\mu_k)]^2\right\},$$

and the likelihood function of the entire sample is:

(A4) $$L_p = \prod_{\forall i,e} L_{p,ie},$$

where the true value of $X_{p,ie}^{\tau=1}$ and $X_{p,ie}^{\tau=2}$ each follows a Normal distribution, $X_{p,ie}^{\tau} \sim N\left(\overline{X}_{p,ie}^{\tau}, \sigma_{p,ie\tau}^2\right)$, τ = 1 or 2, and $\overline{X}_{ie}^{\tau}$ is the parameter estimate and $\sigma_{p,ie\tau}$ is the standard error for each marketing mix reaction variable, obtained from Equation (1). Direct computation of the likelihood function Equation (A4) is intractable because it involves taking integrals over the distributions of $X_{p,ie}^{\tau=1}$ and $X_{p,ie}^{\tau=2}$ for each category i and store e. Instead, we compute L_p based on an unbiased numerical simulator for $L_{p,ie}$ (Lee 1999):

(A5) $$\hat{L}_{p,ie} = \frac{1}{M}\sum_{m=1}^{M} L_{p,ie}^{m}(X_{p,iem}^{\tau=2}, X_{p,iem}^{\tau=1}, Z_{ie}, \eta_{p,k}, \lambda_{p,k}, \mu_k, \sigma_{p,\xi}^2),$$

where $X_{p,iem}^{\tau=2}$ and $X_{p,iem}^{\tau=1}$ are the m-th draw from the random distributions $X_{p,ie}^{\tau} \sim N\left(\overline{X}_{p,ie}^{\tau}, \sigma_{p,ie\tau}^2\right)$, τ = 1 or 2, and the computation of $L_{p,ie}^{m}$ follows Equation (A3). We generate M = 500 random draws from each of the distributions. Equation (A2) is estimated by maximizing the simulated log-likelihood function given by:

(A6) $$\hat{LL}_p = \sum_{i,e} log(\hat{L}_{p,ie}).$$

We then compute the predicted value of each marketing mix reaction, $\hat{X}^{\tau=2}_{p,ie}$, which is uncorrelated with the error term ε_{ie}. The final model to be estimated is:

(A7) $$S^{\tau=2}_{ie} = \alpha + \hat{X}^{\tau=2}_{ie}\gamma_k + Z_{ie}\theta + \varepsilon_{ie},$$

where all notations are defined as before. Since $S^{\tau=2}_{ie}$ is also estimated with uncertainty, we estimate Equation (A7) using a simulated maximum likelihood procedure similar to that described above. Through the procedures outlined above, we are able to address the problem of uncertainty in the estimates of the dependent and key independent variables, as well as potential endogeneity of marketing mix reactions, and are able to get consistent estimates of the parameters of interest.

Reference

Lee, Lung-Fei (1999), "Statistical Inference with Simulated Likelihood Functions," *Econometric Theory*, Vol. 15, No. 3 (June), 337-360.