

Optimal Marketing Strategies for a Customer Data Intermediary

JOSEPH PANCRAS and K. SUDHIR

Web Appendix

A. The Pricing Equations

Retailer

From (5), the retailer's optimization problem is as follows.

$$\max_{r_t^{xy}, \dots, r_{jt}^{xy}} \Pi_{Rt}^{xy} = \sum_{j=1}^J \sum_{i=1}^{N_t} [r_{jt}^{xy} - w_{jt}^{xy}] * \left[\sum_{k=1}^K f^k S_{ijt}^k (r_{jt}^{xy} - D_{ijt}^{xy}) \right] * M_t \quad (A1)$$

For the purposes of the derivation, we drop the superscripts x and y indicating whether a manufacturer bought targeting services and the subscript t that indexes time-period for clarity. These can be included appropriately into the final wholesale and retail margins. Hence the retailer objective is:

$$\max_{r_1, \dots, r_J} \Pi_R = \sum_{j=1}^J \sum_{i=1}^N [r_j - w_j] * \left[\sum_{k=1}^K f^k S_{ij}^k (r_j - D_{ij}) \right] * M_t \quad (A2)$$

Taking the derivative of the objective function with respect to the retail prices, the following first order condition for each product j is:

$$\sum_{i=1}^N \left[\sum_{m=1}^J [r_j - w_j] \frac{\partial S_{ij}^k (r_m - D_{ij})}{\partial r_j} + \sum_{k=1}^K f^k S_{ij}^k (r_m - D_{ij}) \right] = 0 \quad (A3)$$

where w_j is the wholesale price charged by manufacturer to the retailer for brand j .

Define Θ_R^i as the first derivatives of all the (individual consumers') shares with respect to all retail prices (retail prices are common across consumers), with element

$(j,m) = \frac{\partial S_{im}(r_m)}{\partial r_j}$. The retailer first order conditions can then be written in matrix form as:

$$\sum_{i=1}^N \left[\Theta_R^i [R - W] + S^i \right] = 0 \quad (A4)$$

Where R is the vector of retail prices and W is the vector of wholesale prices (which are common across all consumers) and S^i is the vector of shares for each consumer ‘ i ’ over all the brands:

$$R \equiv \begin{bmatrix} r_1 \\ \vdots \\ r_J \end{bmatrix}_{J \times 1}, \quad W \equiv \begin{bmatrix} w_1 \\ \vdots \\ w_J \end{bmatrix}_{J \times 1}, \quad S^i \equiv \begin{bmatrix} S_1^i \\ \vdots \\ S_J^i \end{bmatrix}_{J \times 1} \quad (\text{A5})$$

The vector of retail margins $[R - W]$ is obtained by inverting the above matrix equation:

$$R - W = - \underbrace{\left[\sum_{i=1}^N \Theta_R^i \right]^{-1}}_{\text{Retail Margin}} * \left[\sum_{i=1}^N S_i \right] \quad (\text{A6})$$

where the shares are:

$$S_i \equiv \sum_{k=1}^K f^k S_i^k \equiv \begin{bmatrix} \sum_{k=1}^K f^k S_{i1}^k \\ \vdots \\ \sum_{k=1}^K f^k S_{iJ}^k \end{bmatrix}_{J \times 1} \quad (\text{A7})$$

and the individual specific share derivative matrix with respect to retailer prices is:

$$\Theta_R^i = \begin{bmatrix} \frac{\partial S_{i1}}{\partial p_1} & \frac{\partial S_{i2}}{\partial p_1} & \dots & \frac{\partial S_{iJ}}{\partial p_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial S_{i1}}{\partial p_J} & \frac{\partial S_{i2}}{\partial p_J} & \dots & \frac{\partial S_{iJ}}{\partial p_J} \end{bmatrix}_{J \times J} \quad (\text{A8})$$

$$= \begin{bmatrix} \sum_{k=1}^K f^k [\alpha S_{i1}^k [1 - S_{i1}^k]] & \sum_{k=1}^K f^k [-\alpha S_{i1}^k S_{i2}^k] & \dots & \sum_{k=1}^K f^k [-\alpha S_{i1}^k S_{iJ}^k] \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^K f^k [-\alpha S_{iJ}^k S_{i1}^k] & \sum_{k=1}^K f^k [-\alpha S_{iJ}^k S_{i2}^k] & \dots & \sum_{k=1}^K f^k [\alpha S_{iJ}^k [1 - S_{iJ}^k]] \end{bmatrix}_{J \times J}$$

Therefore the retail price is given by

$$R = W - \left[\sum_{i=1}^N \Theta_{iR} \right]^{-1} * \left[\sum_{i=1}^N S_i \right] \quad (\text{A9})$$

Manufacturer

A manufacturer ‘ m ’ offering a subset \mathfrak{N}_m of brands in the market sets the wholesale price w_{jt}^{xy} (where $j \in \mathfrak{N}_m$) and the coupon face values to individual households (D_{ijt}^{xy}) so as to maximize the manufacturer’s profits. A manufacturer who has not been sold the 1:1 marketing service will have coupon face values set to zero. The manufacturer takes into account the knowledge that retailer prices (r_{jt}^{xy}) will be set taking into account the wholesale prices and the coupon face values that have been issued to individual households.

$$\Pi_{mt}^{xy} = \sum_{j \in \mathfrak{N}_m} \sum_{i=1}^{N_i} [w_{jt}^{xy} - D_{ijt}^{xy} - c_{jt}] * [S_{ijt}^{xy}(r_{jt}^{xy}(w_{jt}^{xy}, D_{ijt}^{xy}) - D_{ijt}^{xy})] * M_t \quad (\text{A10})$$

where c_{jt} is the marginal cost of the manufacturer for brand j in period t , and $S_{ijt}^{xy}(r_{jt}^{xy}(w_{jt}^{xy}, D_{ijt}^{xy}) - D_{ijt}^{xy})$ is the probability of household i , buying brand j in period t given the decisions of manufacturers 1 (denoted by x) and 2 (denoted by y) to purchase the purchase history data, and M_t is the total size of the market in period t . We present the first order conditions for the manufacturer dropping the x, y superscripts and the ‘ t ’ subscript and writing retail price as r_j (not as $r_j(w_j, D_{ij})$) for clarity.

We write $w_{ij} = w_j - D_{ij}$ since the manufacturer sets both the wholesale price and the individual coupon face values to maximize profit. As discussed earlier, even though the manufacturer sets the wholesale price and Catalina sets the coupon face value, analytically it does not matter whether we make this distinction. The first order condition with respect to w_{ij} is:

$$\sum_{i=1}^N \left[\sum_{j \in \mathfrak{N}_m} \left\{ [w_{ij} - c_j] * \frac{dS_{ij}(r_j - D_{ij})}{dw_{ij}} + S_{ij}(r_j - D_{ij}) \right\} \right] = 0 \quad (\text{A11})$$

Define Θ_W^i for each individual consumer such that it contains the first derivatives of all the (individual consumers’) shares with respect to all wholesale prices (wholesale

prices are common across consumers), with element $(j,m) = \frac{\partial S_{im}(r_m - D_{im})}{\partial w_{ij}}$. To account

for the set of brands owned by the same manufacturer, define the manufacturer's ownership matrix O_W such that element (j,m) is equal to one if the manufacturer who sells brand j also sells brand m , and zero otherwise. The manufacturer's first order condition can then be written in matrix form as:

$$\sum_{i=1}^N \left[[O_W \bullet \Theta_W^i] [W_i - C] + S^i \right] = 0 \quad (\text{A12})$$

where $[O_W \bullet \Theta_W^i]$ is the element by element multiplication of the two matrices, W_i is the vector of wholesale prices less the individual coupon values, C is the vector of marginal costs of the manufacturer (C is common across all consumers), and S^i is the vector of shares for each consumer i :

$$W_i = \begin{bmatrix} w_1 - D_{i1} \\ \vdots \\ w_J - D_{iJ} \end{bmatrix}_{J \times 1}, \quad C = \begin{bmatrix} c_1 \\ \vdots \\ c_J \end{bmatrix}_{J \times 1}, \quad S_i = \begin{bmatrix} S_{i1} \\ \vdots \\ S_{iJ} \end{bmatrix}_{J \times 1} \quad (\text{A13})$$

From the manufacturer first order conditions, we can write the manufacturer margin from a particular household i $[W_i - C]$ as follows:

$$[W_i - C] = [O_W \bullet \Theta_W^i]^{-1} * [-S^i] \quad (\text{A14})$$

The share derivatives with respect to wholesale matrix Θ_W^i need to be calculated. As mentioned earlier, the manufacturer response matrix has the elements $(j,m) = \frac{\partial S_{im}(r_m - D_{im})}{\partial w_{ij}}$. Define the Jacobian matrix of derivatives of all retail prices to all

wholesale prices (for consumer 'i') as Δ_{rw}^i , with the element $(j,x) = \frac{dr_x(W)}{dw_j}$.¹ Then Θ_W^i

can be re-written as:

$$\Theta_W^i = \Delta_{rw}^i \Theta_R^i \quad (\text{A15})$$

¹ Berto Villas Boas and Hellerstein (2006) discuss two conditions to assure the invertibility of the Jacobian matrix: (1) additive separability of costs across products for retailer and manufacturer and (2) no interaction between manufacturer and retailer costs. Since these assumptions are maintained in our paper, the Jacobian is invertible.

In the Manufacturer Stackelberg game, manufacturers anticipate how the retailer will respond to changes in wholesale prices and use these reactions when setting wholesale prices. We can solve for the retail reactions $\frac{dr_x(W)}{dw_j}$ by taking the total derivative of the retailer's first order condition with respect to the retail price r_j and the wholesale price w_j :

$$\Psi_W^i * \begin{bmatrix} \frac{dr_1}{dw_1} & \frac{dr_1}{dw_2} & \dots & \frac{dr_1}{dw_J} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{dr_J}{dw_1} & \frac{dr_J}{dw_2} & \dots & \frac{dr_J}{dw_J} \end{bmatrix}_{J \times J} = \begin{bmatrix} \frac{\partial S_1^i}{\partial r_1} & \frac{\partial S_2^i}{\partial r_1} & \dots & \frac{\partial S_J^i}{\partial r_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial S_1^i}{\partial r_J} & \frac{\partial S_2^i}{\partial r_J} & \dots & \frac{\partial S_J^i}{\partial r_J} \end{bmatrix}_{J \times J} \quad (\text{A16})$$

where

$$\Psi_W^i = \begin{bmatrix} 2 \frac{\partial S_1^i}{\partial r_1} + \sum_{m=1}^J [r_m - w_m] \frac{\partial^2 S_m^i}{\partial r_1^2} & \dots & \frac{\partial S_J^i}{\partial r_1} + \frac{\partial S_1^i}{\partial r_J} + \sum_{m=1}^J [r_m - w_m] \frac{\partial^2 S_m^i}{\partial r_J \partial r_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial S_1^i}{\partial r_J} + \frac{\partial S_J^i}{\partial r_1} + \sum_{m=1}^J [r_m - w_m] \frac{\partial^2 S_m^i}{\partial r_1 \partial r_J} & \dots & 2 \frac{\partial S_J^i}{\partial r_J} + \sum_{m=1}^J [r_m - w_m] \frac{\partial^2 S_m^i}{\partial r_J^2} \end{bmatrix}_{J \times J} \quad (\text{A17})$$

The second derivatives are obtained for these relationships of a,b,c (where there is an equality sign, the index a will be preferred to c or b if a is in the equality, and b will be preferred to c if b is in the equality). :

$$\frac{\partial^2 S_a^i}{\partial r_c \partial r_b} = \begin{cases} \alpha^2 * S_a^i * [1 - S_a^i] * [1 - 2 * S_a^i] & a = c = b \\ 2 * \alpha^2 * S_a^i * S_b^i * S_c^i & a \neq c \neq b \\ \alpha^2 * S_a^i * S_b^i * [2 * S_a^i - 1] & a = c \neq b \\ \alpha^2 * S_a^i * S_c^i * [2 * S_a^i - 1] & a = b \neq c \\ \alpha^2 * S_a^i * S_b^i * [2 * S_b^i - 1] & a \neq c = b \end{cases} \quad (\text{A18})$$

Writing the total derivative of the retailer's first order condition in matrix form:

$$\Psi_W^i * [\Delta_{rW}^i]^T = \Theta_R^i \quad (\text{A19})$$

Where $[\Delta_{rW}^i]^T$ is the transpose of the matrix Δ_{rW}^i . Therefore Δ_{rW}^i is obtained as:

$$\Delta_{rW}^i = [[\Psi_W^i]^{-1} * \Theta_R^i]^T \quad (\text{A20})$$

The wholesale price to the retailer is given by $w_j = \max_i w_{ij}$ and the individual specific discount is given by $D_{ij} = w_j - w_{ij}$.

B. Endogeneity Correction

We correct for price endogeneity using the control function approach developed in Petrin and Train (2004). The control function approach has similarities to Rivers and Vuong (1988) and Villas Boas and Winer (1999). The ‘control function’ approach (Hausman 1978) uses extra variables to control for the part of the unobserved component of demand that is correlated with price. In principle, the control functions are constructed using as arguments the differences between observed prices and the predicted prices which are arrived at using all the relevant demand and supply variables observed by the econometrician.

Consider the utility equation:

$$u_{ijt} = X_{ijt}\beta - r_{jt}\alpha + \xi_{jt} + \varepsilon_{ijt} \quad (\text{B1})$$

and rewrite it incorporating the control function as:

$$u_{ijt} = X_{ijt}\beta - r_{jt}\alpha + f(\mu_{jt}; \omega) + [\xi_{jt} - f(\mu_{jt}; \omega)] + \varepsilon_{ijt} \quad (\text{B2})$$

where $f(\mu_{jt}; \omega)$ is the function that controls for the correlation of the unobserved component ξ_{jt} with the price r_{jt} , μ_{jt} are control variables used in such a correction, and ω are the coefficients for μ_{jt} . Let the redefined unobserved component be $\eta_{jt} = [\xi_{jt} - f(\mu_{jt}; \omega)]$. If the function $f(\mu_{jt}; \omega)$ could be constructed and added to the utility function, it is clear from equation (B2) that the resulting random component $\eta_{jt} + \varepsilon_{ijt}$ would no longer be correlated with price (by construction), and the estimates obtained would be corrected for price endogeneity. Petrin and Train (2004) show that (under a wide range of conditions) the control function $f(\mu_{jt}; \omega)$ is linear in the price residuals of a regression of price on its primitives. In our context, we estimate a regression of prices against factor costs as follows:

$$r = \kappa_j + \zeta^* B_t + \mu_{jt} \quad (\text{B3})$$

Where B_t are the factor prices, κ_j are brand specific intercepts and μ_{jt} are the residuals from this regression. Thus $f(\mu_{jt};\omega) = \omega\mu_{jt}$, and we write equation (B2) as :

$$u_{ijt} = X_{ijt}\beta - r_{jt}\alpha + \omega\mu_{jt} + \eta_{jt} + \varepsilon_{ijt} \quad (\text{B4})$$

This utility equation (B4) is used in estimating the latent class model rather than equation (1) of the text to perform the endogeneity correction. Different specifications can be used for ω (Petrin and Train 2004 pages 25-26), and we present the results where ω is segment-specific, i.e., $[\omega_k]_{k=1,\dots,K}$.

References

- Berto Villas-Boas, S. and Rebecca Hellerstein (2006), "Identification of Supply Models of Manufacturer and Retailer Oligopoly Pricing", *Economics Letters*, 90(1), 132-140.
- Hausman, J. A. (1978), "Specification Tests in Econometrics", *Econometrica*, Vol. 46, No. 6, Nov., 1978, pp. 1251-1271.
- Petrin, A. and K.E. Train (2004), "Omitted Product Attributes in Discrete Choice Models," NBER Working Paper No.W9452, National Bureau of Economic Research, Cambridge, MA.
- Rivers, D. and Q.H. Vuong (1988), "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models", *Journal of Econometrics*, 39(November), 347-366.
- Villas-Boas, M. and R. Winer (1999), "Endogeneity in Brand Choice Models," *Management Science*, 45(10), 1324-1338.