TECHNICAL SUPPLEMENT

TO

ASYMMETRIC STORE POSITIONING AND PROMOTIONAL ADVERTISING STRATEGIES: THEORY AND EVIDENCE

Surendra Rajiv
Graduate School of Business
University of Chicago

Shantanu Dutta
Marshall School of Business
University of Southern California

Sanjay K. Dhar
Graduate School of Business
University of Chicago
PART I – THEORETICAL ANALYSIS

1. NOTATIONS

<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>EXPLANATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v &gt; 0$</td>
<td>Willingness-to-pay of $LV$ consumers while buying at store $L$.</td>
</tr>
<tr>
<td>$\theta &gt; 1$</td>
<td>Measure of asymmetry in the quality positioning of the competing stores (extent of vertical differentiation). $\theta = 1$ implies that stores are identical on the quality dimension.</td>
</tr>
<tr>
<td>$\beta &gt; 1$</td>
<td>Measures the magnitude of the switching cost that consumers incur while switching from the store that they patronize to the competing store (extent of horizontal differentiation). $\beta = \infty$ corresponds to stores being localized monopolies.</td>
</tr>
<tr>
<td>$\alpha &gt; 1$</td>
<td>Measure of $HV$ consumers’ higher valuation for service (extent of vertical consumer heterogeneity).</td>
</tr>
<tr>
<td>$\tau &gt; 1$</td>
<td>Measure of $HV$ consumers’ higher shopping cost (extent of horizontal consumer heterogeneity).</td>
</tr>
<tr>
<td>$s \in [0, \bar{s}]$</td>
<td>Measure of variation in consumers’ shopping costs (another dimension of horizontal consumer heterogeneity). Higher $\bar{s} &gt; 0$ means greater variation in consumer shopping costs.</td>
</tr>
<tr>
<td>$\rho \in [0,1]$</td>
<td>Proportion of $HV$ consumers.</td>
</tr>
<tr>
<td>$p_j^r$</td>
<td>“Regular” (non-discounted) price of store $j, j \in {H,L}$.</td>
</tr>
<tr>
<td>$p_j^s$</td>
<td>“Sale” (discounted) price of store $j, j \in {H,L}$.</td>
</tr>
<tr>
<td>$f_j \in [0,1]$</td>
<td>Frequency of advertised sales of store $j, j \in {H,L}$.</td>
</tr>
<tr>
<td>$k \geq 0$</td>
<td>Unit cost of advertising sale price.</td>
</tr>
<tr>
<td>$\tilde{k}_j$</td>
<td>Threshold marginal advertising cost so that if $k &gt; \tilde{k}_j$, store $j$ would not have the incentive to engage in promotional advertising.</td>
</tr>
</tbody>
</table>

Note: $LV$ and $HV$ represents low-valuation and high-valuation consumers, respectively.

2. PRELIMINARIES

In this section, we derive in detail the demand function facing stores $H$ and $L$ when consumers are imperfectly informed about prevailing retailing prices.

Let $\delta_j, j \in \{H,L\}$, denote the advertising decision of store $j$ such that $\delta_j = 1$ if store $j$ advertises its price $p_j$ while $\delta_j = 0$ otherwise.

Demand at Store $H$:

In our model we assume that when store $H$ does not advertise, $H$-patrons are aware of its price while $L$-patrons are not. Similarly, $H$-patrons are not aware of price at store $L$ unless store $L$ advertises its price. When store $H$ (store $L$) does not advertise its price, we assume that $L$-patrons (region $H$) form rational expectations about the unadvertised price. Essentially, rational expectations require that while forming their beliefs about unadvertised prices, consumers correctly recognize the incentives facing the competing stores influencing their promotional advertising strategies. In the context of our model, this implies that $H$-patrons, on not observing a price advertisement from store $L$, correctly infer that its price is “high” i.e. $p_L^r \geq p_H - v(\theta - 1)$. This is so because if store $L$‘s price were “low”, its optimal strategy would be to advertise its price and attract additional customers from region $H$. A similar argument would suggest that
L-patrons, on not observing a price ad from the store H, correctly infer that store H’s price is “high” i.e. \( p_H^* \geq p_L + \alpha v(\theta - \lambda) \).

Consider the demand at store H when it offers a low price, \( p_H \leq p_L + v(\theta - \lambda) < p_L + \alpha v(\theta - \lambda) \). In this case, since \( p_L \geq p_H - \alpha v(\theta - \lambda) \), even if store L advertises its price, neither HV nor LV H-patrons will have incentive to buy from store L. Further, since L-patrons are not aware of \( p_H \) and expect that \( p_H^* \geq p_L + \alpha v(\theta - \lambda) \), no L-patrons will buy at store H. Thus, the demand at store H is entirely obtained from H-patrons, specifically LV store L-patrons with \( s \leq \theta v - p_H \) and store H-patrons with \( s \leq (\alpha \theta v - p_H)/\tau \).

However, when store H advertises its price, since \( p_H < p_L + \alpha v(\theta - \lambda) \), store H will also attract LV as well as HV store L-patrons (besides LV and HV store H-patrons). Specifically, the set of LV and HV store L-patrons who will buy from store H at the advertised price \( p_H < p_L + \alpha v(\theta - \lambda) \) are given by

\[
s \leq \frac{v(\theta - \lambda) + p_L - p_H}{\beta - 1} \quad \text{and} \quad s \leq \frac{\alpha v(\theta - \lambda) + p_L - p_H}{(\beta - 1)\tau},
\]

(A.1)

respectively. Thus, the demand function of store H, as a function of competitive prices and advertising decisions, \( \{p_H, p_L, \delta_H, \delta_L\} \), when \( p_H \leq p_L + v(\theta - \lambda) \) is given by

\[
D_H \left( p_H, p_L, \delta_H, \delta_L \right) = \frac{\theta v[\alpha p + \tau (l - \rho)] - p_H[\rho + \tau (l - \rho)]}{\tau s} \quad \text{if} \quad p_H < p_L + v(\theta - \lambda)
\]

\[+ \delta_H \left[ \frac{v(\theta - l)[\alpha p + \tau (l - \rho)] + (p_L - p_H)[\rho + \tau (l - \rho)]}{(\beta - 1)\tau s} \right],
\]

(A.2)

which gives us Equation (3) in the text.

Now, consider the demand at store H when offers a high price, \( p_H > p_L + \alpha v(\theta - \lambda) \).

In this case, store H will not attract any HV or LV store L-patrons even if it were to advertise its price \( p_H \). As long as store L does not advertise, i.e., \( \delta_L = 0 \), store H will attract LV store H-patrons with \( s \leq \theta v - p_H \) and HV store H-patrons with \( s \leq (\alpha \theta v - p_H)/\tau \). However, if store L advertises, i.e., \( \delta_L = 1 \), since \( p_L < p_H - \alpha v(\theta - \lambda) \), store H will lose both LV and HV store H-patrons. Specifically, the set of LV and HV store H-patrons who will switch to store L when store L advertises its price \( p_L < p_H - \alpha v(\theta - \lambda) \), are given by

\[
s \leq \frac{p_H - p_L - v(\theta - \lambda)}{\beta - 1} \quad \text{and} \quad s \leq \frac{p_H - p_L - \alpha v(\theta - \lambda)}{(\beta - 1)\tau},
\]

(A.3)

respectively. Thus, the demand function of store H, as a function of competitive prices and advertising decisions, \( \{p_H, p_L, \delta_H, \delta_L\} \), when \( p_H > p_L + \alpha v(\theta - \lambda) \) is given by

\[
D_H \left( p_H, p_L, \delta_H, \delta_L \right) = \frac{\theta v[\alpha p + \tau (l - \rho)] - p_H[\rho + \tau (l - \rho)]}{\tau s} \quad \text{if} \quad p_H > p_L + v(\theta - \lambda)
\]

\[-\delta_L \left[ \frac{(p_H - p_L)[\rho + \tau (l - \rho)] - v(\theta - l)[\alpha p + \tau (l - \rho)]}{(\beta - 1)\tau s} \right],
\]

(A.4)

\footnote{Note that it is not optimal for store H to offer a price in the “intermediate” range such that \( p_L + \alpha v(\theta - \lambda) > p_H > p_L + v(\theta - 1) \). The store can always increase its profit by reducing its price further (i.e., to “low” price range) and attract low-valuation L-patrons as well.}
which gives us Equation (2) in the text. Equations (A.2) and (A.4) completely characterizes the demand at store $H$.

Demand at Store $L$:

Similarly, it can be shown the demand at store $L$ is given by

\[
D_L\left(\frac{p_H, p_L}{\delta_H, \delta_L}\right) = \frac{v[a\rho + \tau(l - \rho)] - p_L[p + \tau(l - \rho)]}{\tau s} \quad \text{if} \quad p_L \leq p_H - \alpha \nu(\theta - 1)
\]

\[+ \delta_L \left[ \frac{(p_H - p_L)[p + \tau(l - \rho)] - \nu(\theta - 1)[a\rho + \tau(l - \rho)]}{(\beta - 1)s} \right], \tag{A.5}\]

which gives us Equation (5) in the text.

\[
\frac{v[a\rho + \tau(l - \rho)] - p_L[p + \tau(l - \rho)]}{\tau s} \quad \text{if} \quad p_L > p_H - \alpha \nu(\theta - 1)
\]

\[\left[ \frac{\nu(\theta - 1)[a\rho + \tau(l - \rho)] + (p_L - p_H)[p + \tau(l - \rho)]}{(\beta - 1)s} \right]. \tag{A.6}\]

which gives us Equation (4) in the text. Equations (A.5) and (A.6) completely characterizes the demand at store $L$.

3. PROOFS OF PROPOSITIONS & LEMMAS:

Proof of Lemma 5 and Proposition 1:

In the absence of any promotional advertising, each store serves its own patrons only. The demand faced by stores $H$ and $L$ when they serve their own patrons is given by

\[
D_H(p_H) = \rho \left[ \frac{a\theta v - p_H}{\tau s} \right] + (l - \rho) \left[ \frac{\theta v - p_H}{s} \right] \quad \text{and} \quad D_L(p_L) = \rho \left[ \frac{a\nu p_L}{\tau s} \right] + (l - \rho) \left[ \frac{v - p_L}{s} \right]. \tag{A.7}\]

Therefore, the optimal monopoly prices of stores $H$ and $L$ are given by

\[
p_H^m = \frac{\theta v[a\rho + \tau(l - \rho)]}{2[p + \tau(l - \rho)]} \quad \text{and} \quad p_L^m = \frac{v[a\rho + \tau(l - \rho)]}{2[p + \tau(l - \rho)]}, \tag{A.8}\]

while the profits of stores $H$ and $L$ under monopoly pricing are given by

\[
\Pi_H^m = \frac{\theta^2 v^2[a\rho + \tau(l - \rho)]^2}{4\tau s[p + \tau(l - \rho)]} \quad \text{and} \quad \Pi_L^m = \frac{v^2[a\rho + \tau(l - \rho)]^2}{4\tau s[p + \tau(l - \rho)]}. \tag{A.9}\]

The equations (A.7)-(A.9) correspond to the equations in Table 2 in the text under No Advertising scenario.

To show that \(p_H = p_H^m, f_H = 0\) and \(p_L = p_L^m, f_L = 0\) are indeed the equilibrium strategies, we need to verify that neither store $H$ nor store $L$ has any incentive to unilaterally deviate from the equilibrium.
Deviation by Store $H$:

Suppose store $H$ unilaterally deviates by setting a price $p_H < p_L^m + \nu(\theta - l)$ and $\delta_H = 1$ so as to attract $L$-patrons. Then store $H$’s profit is given by

$$\Pi_H(p_H, p_L^m, \delta_H = 1) = p_H \left[ \frac{\nu(\theta - l)[\alpha \rho + \tau(l - \rho)] + (p_L^m + \nu - \beta)\rho + \tau(l - \rho)]}{(\beta - l)k_s} \right] - k,$$

(A.10)

and substituting for $p_L^m$ from equation (A.8), the optimality condition for store $H$ is

$$\frac{\partial \Pi_H}{\partial p_H} = \nu(\theta - l)[\alpha \rho + \tau(l - \rho)] + (\nu/2)[\alpha \rho + \tau(l - \rho)] - 2\beta p_H[\rho + \tau(l - \rho)] = 0,$$

(A.11)

$$\Rightarrow \bar{p}_H = \frac{\nu(2\theta - l)[\alpha \rho + \tau(l - \rho)]}{4\beta[\rho + \tau(l - \rho)]},$$

(A.12)

$$\Rightarrow \bar{\Pi}_H \left( p_H = \bar{p}_H, \delta_H = 1, p_L = p_L^m, \delta_L = 0 \right) = \frac{\nu^2(2\theta - l)^2[\alpha \rho + \tau(l - \rho)]^2}{16\beta(\beta - l)k_s[\rho + \tau(l - \rho)]} - k.$$

(A.13)

Note, equations (A.12)-(A.13) correspond to the equations in Table 2 in the text under the Advertising scenario.

Thus, for store $H$ to not unilaterally deviate from the equilibrium $\left( p_H^m, p_L^m, f_H = 0, f_L = 0 \right)$, it must be that

$$k > \bar{k}_H \equiv \bar{\Pi}_H - \Pi_H^m = \frac{\nu^2(2\theta - l)^2[\alpha \rho + \tau(l - \rho)]^2}{16\beta(\beta - l)k_s[\rho + \tau(l - \rho)]}.$$

(A.14)

Deviation by Store $L$:

Similarly, when store $L$ deviates unilaterally by setting a price $p_L < p_H^m - \alpha \nu(\theta - l)$ and $\delta_L = 1$, following the earlier logic it can be shown that

$$\bar{p}_L = \frac{\nu(2\theta - l)[\alpha \rho + \tau(l - \rho)]}{4\beta[\rho + \tau(l - \rho)]},$$

(A.16)

Note, equation (A.16) corresponds to the equation in Table 2 in the text under the Advertising scenario.

So for store $L$ not to deviate we must have

$$\bar{\Pi}_L = \frac{\nu^2(2\theta - l)^2[\alpha \rho + \tau(l - \rho)]^2}{16\beta(\beta - l)k_s[\rho + \tau(l - \rho)]} - k < \Pi_L^m = \frac{\nu^2[\alpha \rho + \tau(l - \rho)]^2}{4\tau s[\rho + \tau(l - \rho)]},$$

$$\Rightarrow k > \bar{k}_L \equiv \bar{\Pi}_L - \Pi_L^m = \frac{\nu^2(2\theta - l)^2[\alpha \rho + \tau(l - \rho)]^2}{16\beta(\beta - l)k_s[\rho + \tau(l - \rho)]}.$$

(A.17)

As shown below in Lemma 1, $\bar{k}_H > \bar{k}_L$. Thus, when $k < \bar{k}_L < \bar{k}_H$, both stores $H$ and $L$ have an incentive to unilaterally deviate from monopoly pricing, $p_j^m$, and advertise their discounted price, $\bar{p}_j$. 

\[\blacksquare\]
Claim 1: Gains From promotional advertising is higher for store H relative to store L.

We need to show that

\[ \tilde{k}_H > \tilde{k}_L \quad \text{(A.18)} \]

\[ 4\theta^2 \beta - 4\theta^2 + 1 > 4\theta^2 + 4\theta^2 \beta \]

\[ 4\beta > 1 \]

which holds since \( \beta > 1 \). This implies that store H has a higher incentive to engage in promotional advertising and it takes a higher marginal cost of advertising to deter it from following a policy of advertised sales. This corresponds to the directionality posited in equation 8 in the text.

Proof of Propositions 2 and 3:

A brief sketch is given in the Appendix to the manuscript. We provide additional details here.

As described in the Appendix to the manuscript, we follow a 3-step procedure:

Step 1: Derivation of Support Points of the Mixing Distributions \( \left\{ \hat{p}^s_H, \hat{p}^r_H \right\} \) and \( \left\{ \hat{p}^s_L, \hat{p}^r_L \right\} \):

Let \( \left( f_H, f_L \right) \) be any pair of frequency of advertised sales for stores H and L. Then, for \( \left( \hat{p}^s_j, \hat{p}^r_j \right) \) to be store j’s NE pricing strategy it must be that

\[ \Pi_j \left( \hat{p}^s_j, \hat{p}^r_j, \hat{p}^s_j | f_j, f_{-j} \right) \geq \Pi_j \left( p^s_j, p^r_j, \hat{p}^s_j | f_j, f_{-j} \right) \quad \text{for } \forall p^s_j, \forall p^r_j \]

where \(-j\) refers to the competing store.

Consider first store H’s choice of optimal pricing strategy – i.e. \( \left( \hat{p}^s_H, \hat{p}^r_H \right) \) – given that store L follows its NE strategy \( \left( \hat{p}^s_L, \hat{p}^r_L, f_L \right) \). First, note that for \( \hat{p}^s_H \) to be store H’s optimal choice of “sale” price, we must have that

\[ \hat{p}^s_H \in \arg\max_{p} \Psi^s_H \left( p_H \mid \hat{p}^s_L, \hat{p}^r_L, f_L \right) \]

where

\[ \Psi^s_H \left( p_H \mid \hat{p}^s_L, \hat{p}^r_L, f_L \right) = p_H \left[ f_L D_H \left( p_H, \hat{p}^s_L, \hat{p}^r_L, f_L \right) + \left( 1 - f_L \right) D_H \left( p_H, \hat{p}^s_L, \hat{p}^r_L, 1, 0 \right) \right]. \]  

(A.19)

In equation (A.19), \( D_H \left( p_H, \hat{p}^s_L, \hat{p}^r_L, 1, 0 \right) \) refers to store H’s demand when both stores advertise their sale prices so that consumers are aware of competitive prices. From equation (A.2), we have

\[ D_H \left( p_H, \hat{p}^s_L, \hat{p}^r_L, 1, 0 \right) = \frac{v(\theta \beta - \alpha) \left( \alpha \rho + \tau (1 - \rho) \right) + \left( \hat{p}^s_H - \beta p^r_H \right) \left( \rho + \tau (1 - \rho) \right)}{(\beta - \rho) s}, \]  

(A.20)

Similarly, \( D_H \left( p^s_H, \hat{p}^r_L, 1, 0 \right) \) refers to store H’s demand when store L charges its regular price (and does not advertise) while store H posts a sale price (and advertises). From equation (A.2), we have

\[ D_H \left( p^s_H, \hat{p}^r_L, 1, 0 \right) = \frac{v(\theta \beta - \alpha) \left( \alpha \rho + \tau (1 - \rho) \right) + \left( \hat{p}^s_H - \beta p^r_H \right) \left( \rho + \tau (1 - \rho) \right)}{(\beta - \rho) s}, \]  

(A.21)

\[ \Rightarrow \frac{\partial \Psi^s_H}{\partial p^s_H} = v(\theta \beta - \alpha) \left( \alpha \rho + \tau (1 - \rho) \right) - \left[ 2 \beta \hat{p}^s_H - f_L p^r_L \right] \left( \rho + \tau (1 - \rho) \right) = 0. \]  

(A.22)

Similarly, for \( \hat{p}^r_H \) to be store H’s optimal choice of “regular” price, we must have that

\[ \hat{p}^r_H \in \arg\max_{p} \Psi^r_H \left( p_H \right) = p_H \left[ f_L D_H \left( p_H, \hat{p}^s_L, 0, 0 \right) + \left( 1 - f_L \right) D_H \left( p_H, \hat{p}^s_L, 0, 0 \right) \right]. \]  

(A.23)
From equation (A.4), we have
\[
D_H(p_H^*, \hat{p}_L, 0, 1) = \frac{v(\theta - 1)(\alpha p + \tau(l - \rho)) + \left(\hat{p}_L^* - \beta p_H^* \right) \rho + \tau(l - \rho)}{\tau s},
\] (A.24)
\[
D_H(p_H^*, \hat{p}_L^*, 0, 0) = \frac{0v(\alpha p + \tau(l - \rho)) - p_H^* \rho + \tau(l - \rho)}{\tau s},
\] (A.25)
\[
\Rightarrow \frac{\partial \Psi_H^I}{\partial p^*_H} = v[f'_L(0 - l) + \theta(\beta - l)][(\alpha p + \tau(l - \rho)) + f_L p^*_L \rho + \tau(l - \rho)];
\] (A.26)
\[
\Rightarrow -2 \hat{p}_H^* (\beta - l + f_H) \rho + \tau(l - \rho) = 0.
\] Similarly, considering the choice of store \(L\), we have for store \(L\)'s "sale" price
\[
\hat{p}_L^* \in \arg\max_{p^*_L} \Psi^*_L(\cdot) = p^*_L \left[ f_H D_L(\hat{p}_H^*, p^*_L, 1, 1) + (l - f_H) D_L(\hat{p}_H^*, p^*_L, 0, 0) \right].
\] (A.27)
From equation (A.5), we have
\[
D_L(\hat{p}_H^*, p^*_L, 1, 1) = \frac{v(\beta - 0)(\alpha p + \tau(l - \rho)) + \left(\hat{p}_H^* - \beta p^*_L \right) \rho + \tau(l - \rho)}{\tau s},
\] (A.28)
\[
D_L(\hat{p}_H^*, p^*_L, 0, 0) = \frac{v(\beta - 0)(\alpha p + \tau(l - \rho)) + \left(\hat{p}_H^* - \beta p^*_L \right) \rho + \tau(l - \rho)}{\tau s},
\] (A.29)
\[
\Rightarrow \frac{\partial \Psi_L^I}{\partial p^*_L} = v(\beta - 0)(\alpha p + \tau(l - \rho)) - 2\beta p^*_L - f_H p^*_L - (l - f_H) p_H^* \rho + \tau(l - \rho) = 0.
\] (A.30)
Similarly, for \(\hat{p}_L^*\) to be store \(L\)'s optimal choice of "regular" price, we must have that
\[
\hat{p}_L^* \in \arg\max_{p^*_L} \Psi^*_L(\cdot) = p^*_L \left[ f_H D_L(\hat{p}_H^*, p^*_L, 1, 1) + (l - f_H) D_L(\hat{p}_H^*, p^*_L, 0, 0) \right].
\] (A.31)
From equation (A.6), we have
\[
D_L(\hat{p}_H^*, p^*_L, 1, 1) = \frac{v(\beta - 0)(\alpha p + \tau(l - \rho)) + \left(\hat{p}_H^* - \beta p^*_L \right) \rho + \tau(l - \rho)}{\tau s},
\] (A.32)
\[
D_L(\hat{p}_H^*, p^*_L, 0, 0) = \frac{v(\beta - 0)(\alpha p + \tau(l - \rho)) - p^*_L \rho + \tau(l - \rho)}{\tau s},
\] (A.33)
\[
\Rightarrow \frac{\partial \Psi_L^I}{\partial p^*_L} = v[\beta - l - f_H(0 - l)][\alpha p + \tau(l - \rho)] + f_H p^*_L [\rho + \tau(l - \rho)];
\] (A.34)
\[
\Rightarrow -2 \hat{p}_L^* (\beta - l + f_H) \rho + \tau(l - \rho) = 0.
\] The support-points of store \(H\) and \(L\)'s mixed strategies \(\left\{ \hat{p}_H^*, \hat{p}_H^*, \hat{p}_L, \hat{p}_L^* \right\} \) are obtained by simultaneously solving equations (A.22), (A.26), (A.30) and (A.34), and is as follows:
\[
\hat{p}_H^* = \frac{v\left[ 2\beta(\beta - 1) + 2(\beta - 1)(\beta + f_H) + f_H(4\beta + 3f_H - 1) + f_L(\theta - 1) \left[ 4\beta^3 - (1 - f_H) \right] \right]}{4\beta(\beta + f_L - 1) - f_L(1 - f_H) \left( 4\beta + f_H - 1 \right) - 4f_H f_L(\beta + f_L - 1) \left( \beta + f_H - 1 \right)}
\] (A.35)
\[
\hat{p}_L^* = \frac{v\left[ (\beta - 1) + 2(\beta + f_H)(\beta + f_H - 1) + f_H(1 - f_H) \right]}{4\beta(\beta + f_H - 1) - f_H(1 - f_H) \left( 4\beta + f_H - 1 \right) - 4f_H f_L(\beta + f_H - 1) \left( \beta + f_H - 1 \right)}
\]
\[ \hat{p}_L^r = \frac{\eta \{ 2\beta (\beta - 1)[2(\beta - 1)(\beta + f_H) + f_L(4\beta + 3f_H - 1)] - f_H(1 - f_L)[4\beta^2 - 4f_H(1 - f_H)] \} }{4\beta(\beta + f_L - 1) - f_L(1 - f_H)[4\beta(\beta + f_H - 1) - f_H(1 - f_L)]} \]

(A.36)

\[ \hat{p}_H^r = \frac{\eta \{ (\beta - 1)[2(\beta + f_H)(\beta + f_H - 1)(\beta - f_H) + \phi(1 - f_H)[4\beta(\beta + f_H - 1) - f_H(1 - f_L)] - (\theta - 1) \} }{4\beta(\beta + f_L - 1) - f_L(1 - f_H)[4\beta(\beta + f_H - 1) - f_H(1 - f_L)]} \]

(A.37)

\[ \hat{p}_L^s = \frac{\eta \{ (\beta - 1)[4(\beta + f_H)(\beta + f_H - 1)(\beta - f_H) + \phi(1 - f_H)[4\beta(\beta + f_H - 1) - f_H(1 - f_L)] - (\theta - 1) \} }{4\beta(\beta + f_L - 1) - f_L(1 - f_H)[4\beta(\beta + f_H - 1) - f_H(1 - f_L)]} \]

(A.38)

where \( \eta = [\alpha \rho + \tau (1 - \rho)] / [\rho + \tau (1 - \rho)] \). Equations (A.35)-(A.38) characterize the support points of the mixing distribution and are reported in the Appendix of the text after (A.4).

\[\text{Step 2: Derivation of Stores' Reaction Functions } f_H(f_L) \text{ and } f_L(f_H) :\]

First consider store \( H \)'s choice of \( f_H \), given that store \( L \)'s frequency of advertised sales is \( f_L \). Substituting for \( \hat{p}_H^r, \hat{p}_H^s, \hat{p}_L^r \), and \( \hat{p}_L^s \) from equations (A.35)-(A.38) in equations (A.19), (A.23), (A.27) and (A.31), we obtain expressions for \( \Psi_H^r(f_H,f_L), \Psi_H^s(f_H,f_L), \Psi_L^r(f_H,f_L), \Psi_L^s(f_H,f_L) \):

For store \( H \) to randomize between the pure strategies of posting (and not advertising) its “regular” price, \( \hat{p}_H^r \), and posting (and advertising) its “sale” price, \( \hat{p}_H^s \), its expected profits must be the same under the two strategies i.e.

\[ \Psi_H^r(f_H,f_L) = \Psi_H^s(f_H,f_L) - k. \]

which yields the following optimality condition for the frequency of advertised sale for store \( H \):

\[ F_1(\hat{f}_H, \hat{f}_L, \alpha, \beta, \theta, \nu, \tau, \rho) \equiv (\beta - 1)\hat{f}_H \bar{k} \epsilon \{ \rho + \tau (1 - \rho) \} \times \left\{ 4\beta(\beta + f_H - 1) - f_H(1 - f_L) \right\} \times \]

\[ \left\{ 4\beta(\beta + f_L - 1) - f_L(1 - f_H) \right\} - 4f_Hf_L(\beta - f_H) - \left\{ (\beta - 1) + f_L \right\} \times \]

\[ \left\{ f_Hf_L(1 - f_H)(1 - f_L) - 4\beta(\beta + f_H - 1) \right\} - 4\beta(\beta + f_L - 1) - 4\beta(\beta + f_H - 1) - 4\beta(\beta + f_L - 1) \times \]

\[ - 2\beta f_H f_L(1 - f_H) - \beta f_H f_L(1 - f_H) - 2\beta f_H f_L(1 - f_H) - 2\beta f_H f_L(1 - f_H) \times \]

\[ + f_H(2\beta - 1) + f_L(2\beta - 1) + f_H(2\beta - 1) + f_L(2\beta - 1) + f_H(2\beta - 1) \times \]

\[ + f_L(2\beta - 1) + f_H(2\beta - 1) - 4\beta f_H f_L(1 - f_L) - f_H f_L(1 - f_L) - 5\beta f_H f_L(1 - f_L) - 5\beta f_H f_L(1 - f_L) \times \]

\[ - 2\beta f_H f_L(1 - f_H) \times \}

\[ \{ \alpha \rho + \tau (1 - \rho) \} \times \nu^2 = 0. \]

Similarly, for store \( L \) to randomize between posting (and not advertising) its “regular” price, \( \hat{p}_L^r \), and posting (and advertising) its “sale” price, \( \hat{p}_L^s \), it must be that

\[ \Psi_L^r(f_H,f_L) = \Psi_L^s(f_H,f_L) - k. \]

which gives the following optimality condition for the frequency of advertised sale for store \( L \):

\[ F_2(\hat{f}_H, \hat{f}_L, \alpha, \beta, \theta, \nu, \tau, \rho) \equiv (\beta - 1)\hat{f}_H \bar{k} \epsilon \{ \rho + \tau (1 - \rho) \} \times \left\{ 4\beta(\beta + f_H - 1) - f_H(1 - f_L) \right\} \times \]

\[ \left\{ 4\beta(\beta + f_L - 1) - f_L(1 - f_H) \right\} - 4f_Hf_L(\beta - f_H) - \left\{ (\beta - 1) + f_L \right\} \times \]

\[ \left\{ f_Hf_L(1 - f_H)(1 - f_L) - 4\beta(\beta + f_H - 1) \right\} - 4\beta(\beta + f_L - 1) - 4\beta(\beta + f_H - 1) - 4\beta(\beta + f_L - 1) \times \]

\[ - 2\beta f_H f_L(1 - f_H) - \beta f_H f_L(1 - f_H) - 2\beta f_H f_L(1 - f_H) - 2\beta f_H f_L(1 - f_H) \times \]

\[ + f_H(2\beta - 1) + f_L(2\beta - 1) + f_H(2\beta - 1) + f_L(2\beta - 1) + f_H(2\beta - 1) \times \]

\[ + f_L(2\beta - 1) + f_H(2\beta - 1) - 4\beta f_H f_L(1 - f_L) - f_H f_L(1 - f_L) - 5\beta f_H f_L(1 - f_L) - 5\beta f_H f_L(1 - f_L) \times \]

\[ - 2\beta f_H f_L(1 - f_H) \times \}

\[ \{ \alpha \rho + \tau (1 - \rho) \} \times \nu^2 = 0. \]
Equations (A.40) and (A.42) implicitly define the reaction functions $f_H(f_L)$ and $f_L(f_H)$.

Step 3: Derivation of Promotional Advertising Nash Equilibrium $\{f_H^*, f_L^*\}$:

The NE is obtained by simultaneously solving the 2 (implicit) reaction functions $F_H(f_H, f_L) = 0$ and $F_L(f_H, f_L) = 0$.

Observe that (A.40) and (A.42) are a fourth-order polynomial in $f_H$ and $f_L$ respectively. As such, we could not obtain closed form expressions for the NE frequencies $\{f_H^*, f_L^*\}$ (from simultaneous solution of equation (A.40) and equation (A.42).

As suggested by the Editor, in order to prove erstwhile Proposition 2 and Proposition 3, we ran extensive simulations over a range of parameter values for which the support points were positive (i.e. $\hat{p}_j^* \geq 0$, $\hat{p}_j^* \geq 0$ for $j \in \{H, L\}$). We did this by varying each parameter at a time over the feasible range. This simulation exercise verifies that for the entire range of simulation, the NE frequencies $\{f_H^*, f_L^*\}$ existed in the $[0,1]$ interval.

The simulation results are summarized below:

| Parameter Values: $\nu = 1; \theta = 1.1; \tau = 1.5; \beta = 1.25; \bar{s} = 1; \rho = 0.4; \kappa = 0.05$ |
|-------------------------------------------------|-----------------|-----------------|
| \(\alpha\)                                      | \(f_H^*\)       | \(f_L^*\)       |
| 1.1                                             | 0.3263           | 0.1270           |
| 1.2                                             | 0.3401           | 0.1352           |
| 1.3                                             | 0.3530           | 0.1424           |
| 1.4                                             | 0.3662           | 0.1481           |
| 1.5                                             | 0.3802           | 0.1551           |
| 1.6                                             | 0.3924           | 0.1612           |
| 1.7                                             | 0.4052           | 0.1683           |
| 1.8                                             | 0.4170           | 0.1740           |
| 1.9                                             | 0.4291           | 0.1794           |
| 2.0                                             | 0.4403           | 0.1852           |
TABLE 2: Comparative Statics for $f^*_H$ and $f^*_L$ with respect to $\beta$
Parameter Values: $v = 1; \theta = 1.1; \alpha = 1.5; \tau = 1.5; \tilde{s} = 1; \rho = 0.4; \kappa = 0.05$

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$f^*_H$</th>
<th>$f^*_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>0.4142</td>
<td>0.1762</td>
</tr>
<tr>
<td>1.2</td>
<td>0.3940</td>
<td>0.1650</td>
</tr>
<tr>
<td>1.3</td>
<td>0.3634</td>
<td>0.1371</td>
</tr>
<tr>
<td>1.4</td>
<td>0.3281</td>
<td>0.1023</td>
</tr>
<tr>
<td>1.5</td>
<td>0.2892</td>
<td>0.0672</td>
</tr>
<tr>
<td>1.6</td>
<td>0.2643</td>
<td>0.0514</td>
</tr>
<tr>
<td>1.7</td>
<td>0.2410</td>
<td>0.0511</td>
</tr>
<tr>
<td>1.8</td>
<td>0.2194</td>
<td>0.0507</td>
</tr>
<tr>
<td>1.9</td>
<td>0.1971</td>
<td>0.0503</td>
</tr>
<tr>
<td>2.0</td>
<td>0.1760</td>
<td>0.0501</td>
</tr>
</tbody>
</table>

TABLE 3: Comparative Statics for $f^*_H$ and $f^*_L$ with respect to $\kappa$
Parameter Values: $v = 1; \theta = 1.1; \alpha = 1.5; \tau = 1.5; \beta = 1.25; \tilde{s} = 1; \rho = 0.4$

<table>
<thead>
<tr>
<th>$\kappa$</th>
<th>$f^*_H$</th>
<th>$f^*_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.7501</td>
<td>0.3108</td>
</tr>
<tr>
<td>0.02</td>
<td>0.6018</td>
<td>0.2567</td>
</tr>
<tr>
<td>0.03</td>
<td>0.5044</td>
<td>0.2149</td>
</tr>
<tr>
<td>0.04</td>
<td>0.4337</td>
<td>0.1818</td>
</tr>
<tr>
<td>0.05</td>
<td>0.3803</td>
<td>0.1548</td>
</tr>
<tr>
<td>0.06</td>
<td>0.3356</td>
<td>0.1332</td>
</tr>
<tr>
<td>0.07</td>
<td>0.3002</td>
<td>0.1141</td>
</tr>
<tr>
<td>0.08</td>
<td>0.2704</td>
<td>0.0968</td>
</tr>
<tr>
<td>0.09</td>
<td>0.2438</td>
<td>0.0826</td>
</tr>
<tr>
<td>0.10</td>
<td>0.2209</td>
<td>0.0702</td>
</tr>
</tbody>
</table>

TABLE 4: Comparative Statics for $f^*_H$ and $f^*_L$ with respect to $\rho$
Parameter Values: $v = 1; \theta = 1.1; \alpha = 1.5; \tau = 1.5; \beta = 1.25; \tilde{s} = 1; \kappa = 0.05$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$f^*_H$</th>
<th>$f^*_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.3665</td>
<td>0.1497</td>
</tr>
<tr>
<td>0.30</td>
<td>0.3701</td>
<td>0.1503</td>
</tr>
<tr>
<td>0.35</td>
<td>0.3748</td>
<td>0.1527</td>
</tr>
<tr>
<td>0.40</td>
<td>0.3802</td>
<td>0.1551</td>
</tr>
<tr>
<td>0.45</td>
<td>0.3838</td>
<td>0.1573</td>
</tr>
<tr>
<td>0.50</td>
<td>0.3887</td>
<td>0.1598</td>
</tr>
<tr>
<td>0.55</td>
<td>0.3942</td>
<td>0.1618</td>
</tr>
<tr>
<td>0.60</td>
<td>0.3988</td>
<td>0.1647</td>
</tr>
<tr>
<td>0.65</td>
<td>0.4041</td>
<td>0.1669</td>
</tr>
<tr>
<td>0.70</td>
<td>0.4094</td>
<td>0.1702</td>
</tr>
</tbody>
</table>
TABLE 5: Comparative Statics for $f^*_H$ and $f^*_L$ with respect to $\tilde{s}$
Parameter Values: $v = 1; \theta = 1.1; \alpha = 1.5; \tau = 1.5; \beta = 1.25; \rho = 0.4; \kappa = 0.05$

<table>
<thead>
<tr>
<th>$\tilde{s}$</th>
<th>$f^*_H$</th>
<th>$f^*_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5491</td>
<td>0.2342</td>
</tr>
<tr>
<td>0.6</td>
<td>0.5044</td>
<td>0.2148</td>
</tr>
<tr>
<td>0.7</td>
<td>0.4662</td>
<td>0.1970</td>
</tr>
<tr>
<td>0.8</td>
<td>0.4340</td>
<td>0.1819</td>
</tr>
<tr>
<td>0.9</td>
<td>0.4051</td>
<td>0.1681</td>
</tr>
<tr>
<td>1.0</td>
<td>0.3800</td>
<td>0.1553</td>
</tr>
<tr>
<td>1.1</td>
<td>0.3573</td>
<td>0.1428</td>
</tr>
<tr>
<td>1.2</td>
<td>0.3359</td>
<td>0.1327</td>
</tr>
<tr>
<td>1.3</td>
<td>0.3181</td>
<td>0.1230</td>
</tr>
<tr>
<td>1.4</td>
<td>0.3002</td>
<td>0.1142</td>
</tr>
</tbody>
</table>

TABLE 6: Comparative Statics for $f^*_H$ and $f^*_L$ with respect to $\theta$
Parameter Values: $v = 1; \alpha = 1.5; \tau = 1.5; \beta = 1.25; \tilde{s} = 1; \rho = 0.4; \kappa = 0.05$

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$f^*_H$</th>
<th>$f^*_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.025</td>
<td>0.2738</td>
<td>0.2190</td>
</tr>
<tr>
<td>1.050</td>
<td>0.3081</td>
<td>0.1971</td>
</tr>
<tr>
<td>1.075</td>
<td>0.3430</td>
<td>0.1758</td>
</tr>
<tr>
<td>1.100</td>
<td>0.3798</td>
<td>0.1552</td>
</tr>
<tr>
<td>1.125</td>
<td>0.4232</td>
<td>0.1504</td>
</tr>
<tr>
<td>1.150</td>
<td>0.4729</td>
<td>0.1502</td>
</tr>
<tr>
<td>1.175</td>
<td>0.5258</td>
<td>0.1501</td>
</tr>
<tr>
<td>1.200</td>
<td>0.5842</td>
<td>0.1500</td>
</tr>
<tr>
<td>1.225</td>
<td>0.6477</td>
<td>0.1498</td>
</tr>
<tr>
<td>1.250</td>
<td>0.7186</td>
<td>0.1496</td>
</tr>
</tbody>
</table>

TABLE 7: Comparative Statics for $f^*_H$ and $f^*_L$ with respect to $\tau$
Parameter Values: $v = 1; \theta = 1.1; \alpha = 1.5; \beta = 1.25; \tilde{s} = 1; \rho = 0.4; \kappa = 0.05$

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$f^*_H$</th>
<th>$f^*_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>0.4196</td>
<td>0.1753</td>
</tr>
<tr>
<td>1.2</td>
<td>0.4077</td>
<td>0.1689</td>
</tr>
<tr>
<td>1.3</td>
<td>0.3970</td>
<td>0.1640</td>
</tr>
<tr>
<td>1.4</td>
<td>0.3881</td>
<td>0.1591</td>
</tr>
<tr>
<td>1.5</td>
<td>0.3803</td>
<td>0.1548</td>
</tr>
<tr>
<td>1.6</td>
<td>0.3718</td>
<td>0.1510</td>
</tr>
<tr>
<td>1.7</td>
<td>0.3657</td>
<td>0.1506</td>
</tr>
<tr>
<td>1.8</td>
<td>0.3608</td>
<td>0.1503</td>
</tr>
<tr>
<td>1.9</td>
<td>0.3568</td>
<td>0.1499</td>
</tr>
<tr>
<td>2.0</td>
<td>0.3517</td>
<td>0.1497</td>
</tr>
</tbody>
</table>
Frequency of Promotion: First, note that we have already analytically showed that \( \tilde{k}_H > \tilde{k}_L \) i.e. the incentive to advertise sales is higher for the high-quality store than the low quality store. This itself suggests that while randomizing between “regular” and “sale” prices, store H will put a higher point mass on the “regular” price. Further, the numerical results above also clearly establish that \( f_H^* > f_L^* \) for the range of parametric values considered.

Depth of Discount: We prove the result regarding the % depth of discount in 2 steps:

- In Claim 2, we first show that, if \( f_H > f_L \) as shown above, the “regular” price of store \( H \) is greater than the “regular” price of store \( L \) i.e. \( \hat{p}_H^* > \hat{p}_L^* \);
- In Claim 3, we then show that, if \( f_H > f_L \) as shown above, the “depth of discount” for store \( L \) is higher than that for store \( H \) i.e. \( \Delta \equiv (\hat{p}_L^* - \hat{p}_L^*) > \Delta_H \equiv (\hat{p}_H^* - \hat{p}_H^*) \).

Claims 2 & 3 together, then establish that the “% depth of discount” for store \( L \) is higher than that for store \( H \) i.e. \( \% \Delta_L \equiv \frac{\hat{p}_L^* - \hat{p}_L^*}{\hat{p}_L^*} > \% \Delta_H \equiv \frac{\hat{p}_H^* - \hat{p}_H^*}{\hat{p}_H^*} \).

Before formally stating and proving the claims, we first rewrite the expressions for \( \hat{p}_H^*, \hat{p}_H^*, \hat{p}_L^*, \hat{p}_L^* \) in a more compact notational form:

\[
\hat{p}_H^* = \frac{\eta N_1}{D}; \quad \hat{p}_H^* = \frac{\eta N_2}{D}; \quad \hat{p}_L^* = \frac{\eta N_3}{D}; \quad \text{and,} \quad \hat{p}_L^* = \frac{\eta N_4}{D} \tag{A.43}
\]

where \( \eta = [\alpha + \tau (l - \rho)]/\rho + \tau (l - \rho) \) and

\[
N_1 = 2\beta (l - \beta) \left[ 2 (l - \beta) + f_H (4 \beta + 3 f_L - l) - f_H (\theta - l) (4 \beta^3 - (l - f_H) (4 \beta^2 + f_H (l - f_L))] \right],
\]

\[
N_2 = (l - \beta) \left[ 4 (l - \beta) (f + f_L) (f + f_H - l) + f_H (\theta - l) (4 \beta (f + f_H - l) - \theta) (l - f_H)] + (l - \beta) (4 \beta^2 + f_H (l - f_L))] - (l - f_H) (4 \beta^2 + f_H (l - f_L))] - \theta \right) + \beta (f_H (f + f_H - l) - f_H (l - f_H)] + \beta (f_H (f + f_L - l) + f_H (l - f_H)] + \beta (f_H (f + f_H - l) + f_H (l - f_H)] - f_H (l - f_H)].
\]

\[
N_3 = 2\beta (l - \beta) [2 (l - \beta) (2 \beta + f_H) + f_L (4 \beta + 3 f_H - l)] - f_H (\theta - l) (4 \beta^3 - (l - f_H) (4 \beta^2 + f_H (l - f_L))] \right],
\]

\[
N_4 = (l - \beta) [4 \beta (f_H (f + f_H - l) - f_H (l - f_H)] + f_H (\theta - l) (4 \beta (f + f_H - l) - \theta) (l - f_H)] - \theta \right) + \beta (f_H (f + f_H - l) - f_H (l - f_H)] + \beta (f_H (f + f_L - l) + f_H (l - f_H)] + \beta (f_H (f + f_H - l) - f_H (l - f_H)] - f_H (l - f_H)].
\]

\[
D = [4 \beta (f + f_H - l) - f_H (l - f_H)] (4 \beta (f + f_H - l) - f_H (l - f_H)] - f_H (l - f_H)].
\]

We now formally state and prove the Claims 2 & 3.

**Claim 2:** If \( f_H > f_L \), then the regular price of store \( H \) is higher than that of store \( L \) i.e.

\[
\hat{p}_H^* > \hat{p}_L^*. \tag{A.44}
\]

**Proof of Claim 2:** In view of equation (A.43), to show that \( \hat{p}_H^* > \hat{p}_L^* \), we need to show that \( N_1 > N_3 \).

\[
\Rightarrow 2 \beta (l - \beta) [2 (l - \beta) (2 \beta + f_H) + f_H (4 \beta + 3 f_H - l)] + f_H (\theta - l) (4 \beta^3 - (l - f_H) (4 \beta^2 + f_H (l - f_L))] > 2 \beta (l - \beta) [2 (l - \beta) (2 \beta + f_H) + f_H (4 \beta + 3 f_H - l)] - f_H (\theta - l) (4 \beta^3 - (l - f_H) (4 \beta^2 + f_H (l - f_L))]
\]

\[
\Rightarrow 2 \beta (l - \beta) [2 (l - \beta) (2 \beta + f_H) + f_H (4 \beta + 3 f_H - l)] + f_H (\theta - l) (4 \beta^3 - (l - f_H) (4 \beta^2 + f_H (l - f_L))] > 2 \beta (l - \beta) [2 (l - \beta) (2 \beta + f_H) + f_H (4 \beta + 3 f_H - l)] - f_H (\theta - l) (4 \beta^3 - (l - f_H) (4 \beta^2 + f_H (l - f_L))]
\]

\[
\Rightarrow 2 \beta (l - \beta) [2 \beta (f_H - f_H) + f_H - f_H] + f_H (\theta - l) (4 \beta^3 - (l - f_H) (4 \beta^2 + f_H (l - f_L))] > 0
\]

which holds since \( f_H > f_L \).
Claim 3: If $f_H > f_L$, then the magnitude of the depth of discount for store L is higher than that for store H i.e.,

$$\Delta L = \hat{p}^*_L - \hat{p}^*_L > \Delta H = \hat{p}^*_H - \hat{p}^*_H.$$  \hspace{1cm} (A.45)

Proof of Claim 3: In view of equation (A.43), to show that $\Delta L > \Delta H$, we need to show that $N_3 - N_4 > N_1 - N_2$. After some algebraic manipulation, it can be shown that

$$N_3 - N_4 = (\beta - I)(I - f_H)\left[4\theta(\beta - I + f_L) + 2\beta f_L - 2 f_H (I - f_L) - f_H \theta (I - f_H)\right]; \text{ and,}$$

$$N_1 - N_2 = (\beta - I)(I - f_L)\left[4\theta(\beta - I + f_H) - f_L (I - f_H) + 2\theta \left(\beta f_H - f_L (I - f_H)\right)\right].$$

Thus, after some further algebraic manipulation it can be shown that

$$\Delta = \Delta L - \Delta H = (N_3 - N_4) - (N_1 - N_2) = 4\beta^2(\theta - I)(\beta - (I - f_H)(I - f_L))$$

$$+ 2\beta f_H f_L (\theta - I) + 2 (\beta - f_H)(1 - f_L)(1 - f_H)(\theta f_H - f_L) > 0,$$

since $\theta > I, \beta > I, f_H, f_L \in [0, I]$ and $f_H > f_L$.

From Claims 2 & 3, we conclude that $\%\Delta L < \%\Delta L$.

\[\text{\blacksquare}\]

Proof Of Proposition 3:

Recall from the discussion in Lemma 5 and Proposition 1 that $\tilde{k}_H$ and $\tilde{k}_L$ measure the incentives of stores H and L to engage in promotion advertising. Thus, given the cost of advertising ($k$), lower the value of $\tilde{k}_j$, higher the incentive to engage in promotional advertising (equivalently, the frequency of advertised sales) of store $j$.

From equation (A.15), we have

$$\tilde{k}_H = \frac{\nu^2(4\theta^2 \beta - 4\theta \beta + I)\left[\alpha \rho + \tau (I - \rho)\right]^2}{16\beta (\beta - I) \tau s \left[\rho + \tau (I - \rho)\right]}.$$  \hspace{1cm} (A.46)

Similarly, from equation (A.17), we have

$$\tilde{k}_L = \frac{\nu^2(\theta^2 - 4\theta \beta + 4\beta)\left[\alpha \rho + \tau (I - \rho)\right]^2}{16\beta (\beta - I) \tau s \left[\rho + \tau (I - \rho)\right]}.$$  \hspace{1cm} (A.47)

Further, recall that $\Delta k = (\tilde{k}_H - \tilde{k}_L)$ measures the differential across the high- and low-quality stores in their incentives to engage in promotion advertising. Thus, higher the values of $\Delta k$, greater the differential across the stores in their frequencies of advertised sales.

Now, from (A.46) and (A.47), we have

$$\Delta \tilde{k} = \frac{\nu^2(4\beta - 1)(\theta^2 - 1)\left[\alpha \rho + \tau (I - \rho)\right]^2}{16\beta (\beta - I) \tau s \left[\rho + \tau (I - \rho)\right]}.$$  \hspace{1cm} (A.48)

Thus, we have

$$\frac{\partial \Delta \tilde{k}}{\partial \alpha} = \frac{\nu^2 \rho (4\beta - 1)(\theta^2 - 1)\left[\alpha \rho + \tau (I - \rho)\right]}{8 \beta (\beta - I) \tau s \left[\rho + \tau (I - \rho)\right]} > 0$$

since $\beta > 1, \tau > 1$ and $\rho \in (0,1)$.

Further,

$$\frac{\partial \Delta \tilde{k}}{\partial \theta} = \frac{\nu^2 (4\beta - 1)(\theta^2 - 1)\left[\alpha \rho + \tau (I - \rho)\right]^2}{8 \beta (\beta - I) \tau s \left[\rho + \tau (I - \rho)\right]} > 0$$  \hspace{1cm} (A.49)

$$\frac{\partial \Delta \tilde{k}}{\partial \rho} = \frac{\nu^2 (4\beta - 1)(\theta^2 - 1)\left[\alpha \rho + \tau (I - \rho)\right] \left[\tau - 1\right] \left[\alpha \rho + \tau (I - \rho)\right] + 2 \tau (\alpha - 1)}{16 \beta (\beta - I) \tau s \left[\rho + \tau (I - \rho)\right]^2} > 0$$  \hspace{1cm} (A.50)
\[
\frac{\partial \Delta k}{\partial \beta} = - \frac{\nu^2(4\beta^2 - 2\beta + 1)(\theta^2 - 1)(\alpha + \tau(1 - \rho))^2}{16\beta^2(\beta - 1)^2 \tau \frac{1}{s}[\rho + \tau(1 - \rho)]} < 0 \quad (A.52)
\]
\[
\frac{\partial \Delta \tilde{k}}{\partial s} = - \frac{\nu^2(4\beta - 1)(\theta^2 - 1)(\alpha + \tau(1 - \rho))^2}{16\beta \frac{1}{s^2}(\beta - 1)^2 \tau \frac{1}{s}[\rho + \tau(1 - \rho)]} < 0 \quad (A.53)
\]
PART II – EMPIRICAL ANALYSIS

1. MODELING THE RELATIONSHIP BETWEEN TDAYS AND SERV:

Since the variable TDAYS can only take positive values, using a linear model is inappropriate and hence we use the proportional hazard function approach (Jain and Vilcassim, 1991) with a baseline exponential hazard. Section 1.1 below describes the basic ideas behind this econometric model.

Further, since we are also pooling data across stores and across product groups, we need to control for store- and product group-specific unobserved heterogeneity. Failure to do so may result in serial correlation thereby making the parameter estimates inconsistent (Heckman and Singer, 1984). To account for unobserved heterogeneity on two dimensions – viz., store and product group – we use a random effects model with discrete heterogeneity distribution – the latent class random effects specification (Kamakura and Russell, 1989). Section 1.2 details the specifics of our approach.

1.1 Proportional Hazard Function Model:

Let the random variable \( T_{ij} \) (short form for the variable TDAYS\( _{ij} \)) denote the inter-sale time (i.e. the time elapsed between successive advertised sales events) for store \( i \) in product category \( j \). We define the probability that store \( i \) advertises a sale in product category \( j \) given that it has not done so at time \( t \) since the last advertised sales during the interval \((t, t+dt)\) as \( \Pr(t \leq T_{ij} \leq t+dt | T_{ij} \geq t) \). Dividing this probability by \( dt \) gives the instantaneous rate of promotional advertising (over the interval \( dt \)) by store \( i \) in product category \( j \) after \( t \) days have elapsed since the last sale event in this product category. The hazard function, then, is defined as the limit of this rate of probability change as the interval converges to zero. Formally,

\[
h_{ij}(T_{ij} = t) \equiv \lim_{dt \to 0} \frac{\Pr(t \leq T_{ij} \leq t+dt | T_{ij} \geq t)}{dt}.
\]  

(B.1)

Let \( F_{ij}(t) = \Pr(T_{ij} \leq t), t \geq 0 \) be the distribution function and \( f_{ij}(t) \) be the probability density function of inter-sale time at time \( t \). The hazard function, the density function and distribution function can be shown to be related as follows (Kalbfleisch and Prentice, 1980; Lancaster, 1990):

\[
h_{ij}(t) = \frac{f_{ij}(t)}{1 - F_{ij}(t)}.
\]  

(B.2)

From equation (B.2), we can readily see that the hazard rate is given by the density function of inter-sales time divided by the probability that store \( i \) has not advertised a sale in product category \( j \) at time \( t \) since the last sale event. The probability density function of the inter-sales time can be shown to be given by (Kalbfleisch and Prentice, 1980)

\[
f_{ij}(t) = h_{ij}(t) \exp\left[-\int_0^t h_{ij}(u) \, du\right].
\]  

(B.3)

In our analysis, we use the exponential specification for the hazard function. The rationale is as follows. The conceptualization of sales promotions as mixed strategies (Narasimhan 1988, Raju, Srinivasan and Lal 1990) implies that different observations from a store correspond to the store independently drawing from its equilibrium promotional strategies. In particular, it implies lack of duration dependence for inter-sales time. The exponential hazard function has the form

\[
h_{ij}(t) = \lambda_{ij},
\]  

(B.4)

where \( \lambda_{ij} \) are the parameters that determine the shape of the hazard function. With the exponential hazard specification, the density function is

\[
f_{ij}(t) = \lambda_{ij} \exp(-\lambda_{ij} t).
\]  

(B.5)

To incorporate the impact of SERV and the control variables NAME and SALE, we use the proportional hazard specification and reparametrize the hazard rate as follows:

\[
h_{ij}(T_{ij} = T_{ijk}) = \lambda_{ij} \times g(X_{ijk}) \equiv \lambda_{ij} \times \exp\left(\alpha_1 \times SERV_{ijk} + \alpha_2 \times NAME_{ijk} + \alpha_3 \times SALE_{ijk}\right),
\]  

(B.6)
so that the density function of inter-sale time is given by
\[ f_{ij}(T_{ij} = T_{ijk}) = \lambda_{ij} \times g(X_{ijk}) \times \exp[\lambda_{ij} \times g(X_{ijk}) \times T_{ijk}]. \] (B.7)

where $T_{ijk}$ stands for the inter-sale time for the $k$th advertised sale by store $i$ in product category $j$.

The data set used in the analysis has the following panel structure. We have a time series observation $(k = 1, \ldots, K_{ij})$ on advertised sales by store $i$, $i = 1, \ldots, N$, in product groups $j$, $j = 1, \ldots, J$. Then, the sample likelihood for the hazard specification (B.6) is given by:

\[ L = \prod_{i=1}^{N} \prod_{j=1}^{J} \prod_{k=1}^{K_{ij}} \lambda_{ijk} \times g(X_{ijk}) \times \exp[-\lambda_{ijk} \times g(X_{ijk}) \times T_{ijk}]. \] (B.8)

In our data set, there are 6 stores and 14 product groups i.e., $N = 6$ and $J = 14$.

1.2 Incorporating Unobserved Store- and Product-Specific Heterogeneity:

As mentioned earlier, since we are also pooling data across stores and across product groups, we need to control for store- and product group-specific unobserved heterogeneity. We capture unobserved heterogeneity through variations on the baseline hazard rate $\lambda_{ij}$. In the heterogeneity literature (e.g. Chintagunta, Jain and Vilcassim, 1993; Gonul and Srinivasan, 1995), two approaches for incorporating unobserved heterogeneity have been proposed – fixed effects and random effects specifications. In the fixed effects specification, a separate hazard rate, $\lambda_{ij}$, is computed for each store-product group pair. Given our data set with 6 store and 14 product groups, that would entail estimating 84 hazard rate parameters, besides parameters for $SERV$, $NAME$ and $SALE$. Clearly, this is cumbersome and will result in inefficient parameter estimates. In the random effects specification, the hazard rate, $\lambda_{ij}$, is assumed to be distributed across stores and product group according to some discrete distribution – the semi-parametric random effects model (also called the latent class model; Kamakura and Russell, 1989) – or according to some (pre-specified) continuous density function – the parametric random effects model. In our analysis, we adopt the latent class approach.

To allow for both store- and product-specific effects, we parametrize the hazard rate, $\lambda_{ij}$, as

\[ \lambda_{ij} = \eta_i + \zeta_m, \] (B.9)

where $\eta_i$ and $\zeta_m$ denote the store-specific and product group-specific random components, respectively. Let set $S = \{\eta_1, \ldots, \eta_S\}$ denote the set of support points of the (discrete) distribution of the store-specific random component and let $\{\gamma_1, \ldots, \gamma_S\}$ be the associated probability masses with $\sum_{l=1}^{l=S} \gamma_l = 1$. Similarly, let set $P = \{\zeta_1, \ldots, \zeta_P\}$ denote the set of support points of the (discrete) distribution of the product group-specific random component and let $\{\omega_1, \ldots, \omega_P\}$ be the associated probability masses with $\sum_{m=1}^{m=P} \omega_m = 1$.

Then, the sample likelihood is given by

\[ L = \prod_{i=1}^{N} \left\{ \sum_{\eta_i \in S} \sum_{\zeta_m \in P} \prod_{j=1}^{J} \prod_{k=1}^{K_{ij}} (\eta_i + \zeta_m) g(X_{ijk}) \exp\left[-(\eta_i + \zeta_m) g(X_{ijk}) T_{DAYS_{ijk}}\right] \omega_m \gamma_l \right\}. \] (B.10)

1.3 Interpretation of Parameters:

Note that for the exponential hazard specification, the expected inter-sale time between the $k$-1th and the $k$th advertised sale for store $i$ in product category $j$ is given by

\[ E(T_{ijk}) = \frac{1}{\lambda_{ij} \times \exp(\alpha_1 \times SERV_{ijk} + \alpha_2 \times NAME_{ijk} + \alpha_3 \times SALE_{ijk})}. \] (B.11)

Thus, if $\alpha_1 > 0$, expected interval between advertised sales decreases with $SERV$. Note that the variable $SERV$ is reverse coded so that $SERV = 1$ means “excellent sales help” while $SERV = 5$ means “poor sales
2. MODELING THE RELATIONSHIP BETWEEN DISCOUNT AND SERV:

Since the variable DISCOUNT can only take values between 0 and 1, using a linear model is inappropriate and hence our use of the two-limit probit approach (e.g. Nakamura and Nakamura, 1983; Datar, Jordan, Kekre, Rajiv and Srinivasan, 1997). This approach explicitly recognizes this fact and incorporates the implied truncation on the random error component. Note that the two-limit probit model represents a generalization of the Tobit model (Heckmann, 1979).

2.1 Two-Limit Probit Model:

Let \( y_{ijk}^* \) be the (latent) depth of discount in product group \( j \) advertised by store \( i \) in the \( k \)th promotional advertisement. Further, let \( X_{ijk} \) be a vector of explanatory variables (\( SERV_{ijk}, NAME_{ijk}, SALE_{ijk} \)) impacting \( y_{ijk}^* \), and \( \beta \) be the vector of coefficients with \( \alpha \) being the intercept. Then, the (latent) depth of discount for promotional advertisement \( k \) by store \( i \) in product group \( j \) is given by

\[
y_{ijk}^* = \alpha + X_{ijk} \beta + \epsilon_{ijk} = \alpha + \beta_1 \times SERV_{ijk} + \beta_2 \times NAME_{ijk} + \beta_3 \times SALE_{ijk} + \epsilon_{ijk},
\]

where the random disturbance \( \epsilon_{ijk} \sim N(0,\sigma^2) \). However, note that the (latent) depth of discount \( y_{ijk}^* \) is not observed unless \( y_{ijk}^* \in (0,1) \). Denoting \( y_{ijk} \) to be the observed depth of discount for the \( k \)th promotional advertisement by store \( i \) in product group \( j \), we have

\[
y_{ijk} = 0 \quad \text{if} \quad y_{ijk}^* \leq 0; \quad y_{ijk} = y_{ijk}^* \quad \text{if} \quad 0 < y_{ijk}^* < 1; \quad \text{and,} \quad y_{ijk} = 1 \quad \text{if} \quad y_{ijk}^* \geq 1,
\]

which implies that the relationship between observed depth of discount and the explanatory variables is

\[
y_{ijk} = \alpha + X_{ijk} \beta + \psi_{ijk} + \nu_{ijk},
\]

where

\[
\psi_{ijk} = \frac{\phi(\epsilon_{ijk}) - \phi(f_{ijk})}{\Phi(f_{ijk}) - \Phi(\epsilon_{ijk})}; \quad \epsilon_{ijk} = \frac{-\alpha - X_{ijk} \beta}{\sigma}; \quad f_{ijk} = \frac{1 - \alpha - X_{ijk} \beta}{\sigma},
\]

with \( \phi(\cdot) \) and \( \Phi(\cdot) \) being the standard normal density and distribution functions, respectively. Further, it can be shown (Nakamura and Nakamura, 1983) that the mean and variance of the random disturbance \( \nu_{ijk} \) are 0 and \( \sigma^2 \mu_{ijk}^2 \) respectively where \( \mu_{ijk} = 1 + \kappa_{ijk} - \psi_{ijk}^2 \), with \( \psi_{ijk} \) as defined in equation (B.15) and

\[
\kappa_{ijk} = \frac{\epsilon_{ijk} \Phi(\epsilon_{ijk}) - f_{ijk} \Phi(f_{ijk})}{\Phi(f_{ijk}) - \Phi(\epsilon_{ijk})}.
\]

Observe that the disturbance \( \nu_{ijk} \) in the regression model (A.14) is heteroskedastic (Heckman, 1979). Hence, ordinary least square (OLS) estimates of the coefficients will be inefficient and the standard t-tests for assessing the significance of the parameters will be inappropriate because the standard errors will be incorrect. Hence, we use the Maximum Likelihood Estimation procedure to estimate the model parameters.

The data set used in the analysis has the following panel structure. We have a time series observation \((k = 1, \ldots, K)\) on the depth of discount for \( K \) advertised sales by store \( i = 1, \ldots, N \), in product groups \( j = 1, \ldots, J \). Then, the sample likelihood for the two-limit probit specification (B.14) is given by:
\[ L = \prod_{i=1}^{i=N} \prod_{j=1}^{j=J} \prod_{k=1}^{k=K_{ij}} \frac{1}{\sqrt{2\pi \sigma^2 \mu_{ijk}^2}} \times \exp \left[ -\frac{\left( y_{ijk} - \alpha - X'_{ijk} \beta - \sigma \psi_{ijk} \right)^2}{2\sigma^2 \mu_{ijk}^2} \right] \]  
(B.17)

\[ = \prod_{i=1}^{i=N} \prod_{j=1}^{j=J} \prod_{k=1}^{k=K_{ij}} \frac{1}{\sigma \mu_{ijk}} \Phi \left[ \frac{y_{ijk} - \alpha - X'_{ijk} \beta - \sigma \psi_{ijk}}{\sigma \mu_{ijk}} \right]. \]  
(B.18)

where \( \Phi(\cdot) \) is the standard normal density. In our data set, there are 6 stores and 14 product groups i.e., \( N = 6 \) and \( J = 14 \).

### 2.2 Incorporating Unobserved Store- and Product-Specific Heterogeneity:

As in the case of the hazard function model (described in section 1.2), we use the latent class approach to incorporate unobserved heterogeneity in the two-limit probit model, equation (B.14). The unobserved store- and product-specific heterogeneity is captured by allowing the intercept term, \( \alpha_{ij} \), in equation (B.12) to randomly vary across stores and product groups. Specifically, we parametrize the intercept term, \( \alpha_{ij} \), as

\[ \alpha_{ij} = \eta_l + \zeta_m, \]  
(B.19)

where \( \eta_l \) and \( \zeta_m \) denote the store-specific and product group-specific random components, respectively. Let set \( S = \{\eta_1, \ldots, \eta_S\} \) denote the set of support points of the (discrete) distribution of the store-specific random component and let \( \{\gamma_1, \ldots, \gamma_S\} \) be the associated probability masses with \( \sum_{l=1}^{l=S} \gamma_l = 1 \). Similarly, let set \( P = \{\zeta_1, \ldots, \zeta_P\} \) denote the set of support points of the (discrete) distribution of the product group-specific random component and let \( \{\omega_1, \ldots, \omega_P\} \) be the associated probability masses with \( \sum_{m=1}^{m=P} \omega_m = 1 \).

Thus, the relationship between observed depth of discount and the explanatory variables is

\[ y_{ijk} = \eta_l + \zeta_m + X'_{ijk} \beta + \sigma \psi_{ijk} + v_{ijk}^{lm}, \]  
(B.20)

where

\[ \psi_{ijk} = \frac{\Phi(e_{ijk}^{lm}) - \Phi(f_{ijk}^{lm})}{\Phi(f_{ijk}^{lm}) - \Phi(e_{ijk}^{lm})}; \quad e_{ijk}^{lm} = \frac{-\eta_l - \zeta_m - X'_{ijk} \beta}{\sigma}; \quad f_{ijk}^{lm} = \frac{1 - \eta_l - \zeta_m - X'_{ijk} \beta}{\sigma}, \]  
(B.21)

and \( \Phi(\cdot) \) and \( \Phi(\cdot) \) are the standard normal density and distribution functions, respectively. Further, it can be shown (Nakamura and Nakamura, 1983) that the mean and variance of the random disturbance \( v_{ijk} \) are 0 and \( \sigma^2 \mu_{ijk}^{lm} \) respectively where \( \mu_{ijk}^{lm} = 1 + \kappa_{ijk}^{lm} - \psi_{ijk}^{lm} \) and \( \psi_{ijk}^{lm} \) as defined in equation (B.21) and

\[ \kappa_{ijk}^{lm} = \frac{e_{ijk}^{lm} \Phi(e_{ijk}^{lm}) - f_{ijk}^{lm} \Phi(f_{ijk}^{lm})}{\Phi(f_{ijk}^{lm}) - \Phi(e_{ijk}^{lm})}. \]  
(B.22)

Then, the sample likelihood is given by

\[ L = \prod_{i=1}^{i=N} \left\{ \sum_{\eta_l \in S} \sum_{\zeta_m \in P} \left( \prod_{k=1}^{k=K_{ij}} \prod_{l=1}^{l=S} \omega_m \right) \phi \left( y_{ijk} - \eta_l - \zeta_m - X'_{ijk} \beta - \sigma \psi_{ijk}^{lm} \right) \right\} \]  
(B.22)

where \( \phi(\cdot) \) is the standard normal density.

### 2.3 Interpretation of Parameters:
The interpretation of the parameters, $\beta$'s, is exactly the same as in the case of a linear model. Thus, if $\beta_1 > 0$, expected depth of discount for an advertised sales increases with $SERV$. Again, recall that the variable $SERV$ is reverse coded so that $SERV=1$ means “excellent sales help” while $SERV=5$ means “poor sales help”. Now, Hypothesis 2 states that the depth of discount decreases with a store’s quality/service positioning. Thus, (given reverse coding for $SERV$), based on Hypothesis 2 we expect $\beta_1 > 0$. Similarly, based on our discussion in Section 4.4.1 in the paper, we expect $\beta_2 < 0$ and $\beta_3 > 0$.

**REFERENCES**