Web Appendix

Winners and Losers in a Major Price War

Harald J. van Heerde, Els Gijsbrechts, and Koen Pauwels

Web Appendix A: Markov Chain Monte Carlo Estimation of a Hierarchical Multivariate Type II Tobit Model with Heterogeneous Response Parameters

We stack (a) the dependent variables of equations (2) and (3) for all households h and time periods t so that the vector of ln expenditures is $\mathbf{y}_i = [y_{1i1}, y_{1i2}, \ldots, y_{HiT}]'$ and the vector of store incidence indicator variables is $\mathbf{z}_i^* = [z_{1i1}^*, z_{1i2}^*, \ldots, z_{HiT}^*]'$, (b) the predictor variables for the store incidence equation, $\mathbf{V}_i'$ = [$v_{1i1}, v_{1i2}, \ldots, v_{HiT}$]′ and the predictor variables for the ln expenditures equation, $\mathbf{X}_i'$ = [$x_{1i1}, x_{1i2}, \ldots, x_{HiT}$]′, and (c) the error terms of these two equations for all households h and time periods t so that $[\boldsymbol{\varepsilon}_{ht}', \mathbf{u}_{ht}']$'= [$\varepsilon_{h1t}, \varepsilon_{h2t}, \ldots, \varepsilon_{hSt}$, $u_{h1t}, u_{h2t}, \ldots, u_{hSt}$]′ follows a ($2S$)-variate normal distribution with zero mean and full covariance matrix

$$\Sigma = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{12}' & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

where $\boldsymbol{\Sigma}_{11} = E(\boldsymbol{\varepsilon}_{ht}\boldsymbol{\varepsilon}_{ht}')$, $\boldsymbol{\Sigma}_{12} = E(\boldsymbol{\varepsilon}_{ht}\mathbf{u}_{ht}')$, and $\boldsymbol{\Sigma}_{22} = E(\mathbf{u}_{ht}\mathbf{u}_{ht}')$, which has ones on the diagonal since each selectivity mechanism is binary.

Next we specify the hierarchies associated with the two shopping decisions. We stack (i) the intercept coefficients across stores and equations (4) and (5) such that $[\boldsymbol{\alpha}_h', \boldsymbol{\iota}_h']$' $=[\alpha_{h1}, \alpha_{h2}, \ldots, \alpha_{hS}, \iota_{h1}, \iota_{h2}, \ldots, \iota_{hS}]'$ and (ii) the error terms of the hierarchical equations for all households h, $[\boldsymbol{\xi}_h', \boldsymbol{\tau}_h']$' = [$\xi_{h1}, \xi_{h2}, \ldots, \xi_{hS}, \tau_{h1}, \tau_{h2}, \ldots, \tau_{hS}$]′. We model the intercepts as follows:

$$\begin{bmatrix} \boldsymbol{\alpha}_h \\ \boldsymbol{\iota}_h \end{bmatrix} = \begin{bmatrix} \mathbf{I}_S & \\ & \mathbf{I}_S \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta} \\ \boldsymbol{\psi} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\xi}_h \\ \boldsymbol{\tau}_h \end{bmatrix}, \quad \begin{bmatrix} \boldsymbol{\xi}_h \\ \boldsymbol{\tau}_h \end{bmatrix} \sim N_{2S}(0, \mathbf{V})$$

where $\mathbf{I}_S$ is an S by S identity matrix, $\boldsymbol{\delta}$ is an S by 1 vector with store intercepts for the ln expenditure equation, $\boldsymbol{\psi}$ is an S by 1 vector with store intercepts for the incidence equation, and $\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{12}' & \mathbf{V}_{22} \end{bmatrix}$ is a full covariance matrix where $\mathbf{V}_{11} = E(\boldsymbol{\xi}_h\boldsymbol{\xi}_h')$, $\mathbf{V}_{12} = E(\boldsymbol{\xi}_h\boldsymbol{\tau}_h')$, and $\mathbf{V}_{22} = E(\boldsymbol{\tau}_h\boldsymbol{\tau}'_h)$. We summarize the SUR equations above as $\boldsymbol{\alpha} = \mathbf{W}\boldsymbol{\delta} + \boldsymbol{\xi}$ and $\boldsymbol{\iota} = \mathbf{W}\boldsymbol{\psi} + \boldsymbol{\tau}$, where $\mathbf{W} = \mathbf{I}_{2S} \otimes \mathbf{1}_H$ and $\mathbf{1}_H$ is an H by 1 vector of ones.

Finally, unobserved heterogeneity in the response parameters $\boldsymbol{\omega}_h$ and $\boldsymbol{\zeta}_h$ is modeled as follows:

$$\begin{bmatrix} \boldsymbol{\omega}_h \\ \boldsymbol{\zeta}_h \end{bmatrix} \sim N_{|\bar{\varpi}|+|\bar{\xi}|}\left(\begin{bmatrix} \bar{\boldsymbol{\omega}} \\ \bar{\boldsymbol{\zeta}} \end{bmatrix}, \boldsymbol{\Omega}\right)$$, where $\boldsymbol{\Omega}$ is a full covariance matrix.

We use an MCMC approach to estimate the marginal distributions of the latent dependent variables, parameters and covariances. The MCMC algorithm involves sampling sequentially from the relevant conditional distributions over a large number of iterations. These draws can be shown to converge to the marginal posterior distributions. Our implementation of the MCMC algorithm has 9 steps that are described below.

Conditional distributions

The first implementation step requires that we specify conditional distributions of the relevant variables. The solutions of these distributions follow from the normality assumption of the disturbances terms. We employ natural conjugate priors. Specifications of the conditional distributions are as follows:

1. y^*_{hit} is y_{hit} if $z_{hit}=1$, otherwise y^*_{hit} is drawn from a normal distribution:

$$y^*_{hit} \left| \mathbf{y}^*_{h,j\neq i,t}, \mathbf{z}^*_{ht}, \boldsymbol{\omega}_h, \alpha_{hi}, \iota_{hi} \Sigma \sim \begin{cases} y_{hit} \left| z_{hit} = 1 \right. \\ N(\mathbf{v}'_{hit}\boldsymbol{\omega}_h + \alpha_{hi} + \boldsymbol{\sigma}_{ij} \tilde{\Sigma}^{-1}_{jj} \begin{pmatrix} \mathbf{y}^*_{h,j\neq i,t} - \mathrm{E}(\mathbf{y}^*_{h,j\neq i,t}) \\ \mathbf{z}^*_{h,t} - \mathrm{E}(\mathbf{z}^*_{ht}) \end{pmatrix}, \sigma^{(11)}_{ii} - \boldsymbol{\sigma}_{ij} \tilde{\Sigma}^{-1}_{jj} \boldsymbol{\sigma}_{ji}) \left| z_{hit} = 0 \right. \end{cases} \right.$$

$$\text{where } \mathbf{y}^*_{ht} = \begin{bmatrix} y^*_{hit} \\ -- \\ \mathbf{y}^*_{h,j\neq i,t} \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \sigma^{(11)}_{ii} & | & \boldsymbol{\sigma}_{ij} \\ -- & & -- \\ \boldsymbol{\sigma}_{ji} & | & \tilde{\boldsymbol{\Sigma}}_{jj} \end{bmatrix}$$

As the notation suggests, the $\mathbf{y}^*_{ht}$ vector and $\boldsymbol{\Sigma}$ matrix are partitioned between the store chain of interest, i, and all other store chains, $j \neq i$ (the entries in $\boldsymbol{\Sigma}$ corresponding to $\mathbf{z}$ are not shuffled). Without loss of generality, we have assumed the store chain of interest to be the first. Each chain is then drawn in succession for household h, conditioning on $\mathbf{y}^*_{h,i\neq j,t}$, $\mathbf{z}^*_{ht}$, and Σ.

2. We next draw the latent dependent variable values for the probit component of the model. If the indicator variable z_{hit} = 1, then z^*_{hit} is drawn from a normal distribution, truncated below at 0. Otherwise, z^*_{hit} is drawn from a normal distribution, truncated above at 0.

$$\mathbf{z}^*_{hit}\left|\mathbf{z}^*_{h,j\neq i,t},\mathbf{y}^*_{ht},\boldsymbol{\zeta}_h,\alpha_{hi},\iota_{hi},\boldsymbol{\Sigma}\right.\sim N_T\left(\mathbf{x}'_{hit}\boldsymbol{\zeta}_h+\boldsymbol{\iota}_{hi}+\sigma_{ij}\tilde{\boldsymbol{\Sigma}}^{-1}_{jj}\begin{pmatrix}\mathbf{z}^*_{h,j\neq i,t}-\mathrm{E}(\mathbf{z}^*_{h,j\neq i,t})\\ \mathbf{y}^*_{h,t}-\mathrm{E}(\mathbf{y}^*_{ht})\end{pmatrix},\sigma_{ii}-\boldsymbol{\sigma}_{ij}\tilde{\boldsymbol{\Sigma}}^{-1}_{jj}\boldsymbol{\sigma}_{ji}\right)$$

where: $z^*_{ht}=\begin{bmatrix}z^*_{hit}\\ --\\ \mathbf{z}^*_{h,j\neq i,t}\end{bmatrix}$ and $\Sigma=\begin{bmatrix}\sigma^{(22)}_{ii} & | & \boldsymbol{\sigma}_{ij}\\ -- & & --\\ \boldsymbol{\sigma}_{ji} & | & \tilde{\boldsymbol{\Sigma}}_{jj}\end{bmatrix}$

The latent probit dependent variables are drawn using the inverse cdf method.

3. The parameters in $\boldsymbol{\omega}_h$ and $\boldsymbol{\zeta}_h$ are drawn from a SUR model with variance/covariance matrix of disturbances Σ:

$$\begin{bmatrix}\boldsymbol{\omega}_h^{(t)}\\ \boldsymbol{\zeta}_h^{(t)}\end{bmatrix}\left|\mathbf{y}^{*(t)},\mathbf{z}^{*(t)},\boldsymbol{\Sigma}^{(t-1)}\right.\sim N\left(\mathbf{O}_h\left(\begin{bmatrix}\mathbf{V}'_h & 0\\ 0 & \mathbf{X}'_h\end{bmatrix}\left(\boldsymbol{\Sigma}^{-1(t-1)}\otimes\mathbf{I}_T\right)\begin{bmatrix}\mathbf{y}_h^{*(t)}\\ \mathbf{z}_h^{*(t)}\end{bmatrix}+\boldsymbol{\Omega}^{-1}\begin{bmatrix}\overline{\boldsymbol{\omega}}\\ \overline{\boldsymbol{\zeta}}\end{bmatrix}\right),\mathbf{O}_h\right),\text{ where}$$

$$\mathbf{O}_h=\left(\begin{bmatrix}\mathbf{V}'_h & 0\\ 0 & \mathbf{Z}'_h\end{bmatrix}\left(\boldsymbol{\Sigma}^{-1(t-1)}\otimes\mathbf{I}_T\right)\begin{bmatrix}\mathbf{V}'_h & 0\\ 0 & \mathbf{Z}'_h\end{bmatrix}+\boldsymbol{\Omega}^{-1}\right)^{-1}$$

4. The vector of household intercepts $[\boldsymbol{\alpha}'_h,\boldsymbol{\iota}'_h]'$ is drawn from a SUR model with variance/covariance matrix of disturbances $\boldsymbol{\Sigma}$:

$$\begin{bmatrix}\boldsymbol{\alpha}_h^{(t)}\\ \boldsymbol{\iota}_h^{(t)}\end{bmatrix}\left|\mathbf{y}^{*(t)},\mathbf{z}^{*(t)},\boldsymbol{\omega}_h^{(t)},\boldsymbol{\zeta}_h^{(t)},\boldsymbol{\Sigma}^{(t-1)},\mathbf{V}^{(t-1)},\boldsymbol{\delta}^{(t-1)},\boldsymbol{\psi}^{(t-1)}\right.\sim N(\mathbf{Q}(\mathbf{U}'\ (\boldsymbol{\Sigma}^{-1(t-1)}\otimes\mathbf{I}_T)\begin{bmatrix}\mathbf{r}_h^{\alpha}\\ \mathbf{r}_h^{\iota}\end{bmatrix}+\mathbf{V}^{-1(t-1)}\begin{bmatrix}\boldsymbol{\delta}^{(t-1)}\\ \boldsymbol{\psi}^{(t-1)}\end{bmatrix}),\mathbf{Q})$$

where $\mathbf{Q}=\left(\mathbf{U}'(\boldsymbol{\Sigma}^{-1(t-1)}\otimes\mathbf{I}_T)\mathbf{U}+\mathbf{V}^{(t-1)-1}\right)^{-1}$, $\mathbf{U}=\mathbf{I}_{2S}\otimes\mathbf{1}_T$,

$$\mathbf{r}_h^{\alpha}=\begin{bmatrix}\mathbf{y}_{h1}^{*(t)}\\ \mathbf{y}_{h2}^{*(t)}\\ \vdots\\ \mathbf{y}_{hS}^{*(t)}\end{bmatrix}-\begin{bmatrix}\mathbf{V}'_{h1}\\ \mathbf{V}'_{h2}\\ \\ \mathbf{V}'_{hS}\end{bmatrix}\boldsymbol{\omega}_h^{(t)},\text{ and }\mathbf{r}_h^{\iota}=\begin{bmatrix}\mathbf{z}_{h1}^{*(t)}\\ \mathbf{z}_{h2}^{*(t)}\\ \vdots\\ \mathbf{z}_{hS}^{*(t)}\end{bmatrix}-\begin{bmatrix}\mathbf{X}'_{h1}\\ \mathbf{X}'_{h2}\\ \\ \mathbf{X}'_{hS}\end{bmatrix}\boldsymbol{\zeta}_h^{(t)}.$$

6. The vector of hyper-parameters, $[\boldsymbol{\delta}',\boldsymbol{\psi}']'$ is drawn from a SUR model with variance/covariance matrix of disturbances $\mathbf{V}$:

$$\begin{bmatrix}\boldsymbol{\delta}^{(t)}\\ \boldsymbol{\psi}^{(t)}\end{bmatrix}\left|\boldsymbol{\alpha}^{(t)},\boldsymbol{\iota}^{(t)},\mathbf{V}^{(t-1)},\mathbf{V}_{\delta\psi},\overline{\boldsymbol{\delta}},\overline{\boldsymbol{\psi}}\right.\sim N(\mathbf{F}(\mathbf{W}'(\mathbf{V}^{-1(t-1)}\otimes\mathbf{I}_H)\begin{bmatrix}\boldsymbol{\alpha}^{(t)}\\ \boldsymbol{\iota}^{(t)}\end{bmatrix}+\mathbf{V}_{\delta\psi}^{-1}\begin{bmatrix}\overline{\boldsymbol{\delta}}\\ \overline{\boldsymbol{\psi}}\end{bmatrix}),\mathbf{F}),\text{ where}$$

$$\mathbf{F}=\left(\mathbf{W}'(\mathbf{V}^{-1(t-1)}\otimes\mathbf{I}_H)\mathbf{W}+\mathbf{V}_{\delta\psi}^{-1}\right)^{-1}.$$

7. The vector of hyper-parameters, $\left[\overline{\boldsymbol{\omega}}', \overline{\zeta}'\right]'$, is drawn from a SUR model with variance/covariance matrix of disturbances

$$\boldsymbol{\Omega}: \begin{bmatrix} \overline{\boldsymbol{\omega}}^{(t)} \\ \overline{\zeta}^{(t)} \end{bmatrix} | \boldsymbol{\omega}^{(t)}, \zeta^{(t)}, \boldsymbol{\Omega} \sim N\left(\mathbf{J}\left(\mathbf{1}_H{}' \otimes \mathbf{I}_{|\overline{\omega}|+|\overline{\zeta}|})(\mathbf{I}_H \otimes \boldsymbol{\Omega}^{-1}) \begin{bmatrix} \boldsymbol{\omega}^{(t)} \\ \zeta^{(t)} \end{bmatrix} + \overline{\boldsymbol{\Omega}}^{-1} \begin{bmatrix} \overline{\overline{\boldsymbol{\omega}}} \\ \overline{\overline{\zeta}} \end{bmatrix}\right), \mathbf{J}\right)$$

where $\mathbf{J} = \left((\mathbf{1}_H' \otimes \mathbf{I}_{|\overline{\omega}|+|\overline{\psi}|})(\mathbf{I}_H \otimes \boldsymbol{\Omega}^{-1})(\mathbf{1}_H \otimes \mathbf{I}_{|\overline{\omega}|}) + \overline{\boldsymbol{\Omega}}^{-1}\right)^{-1}$.

8. $\boldsymbol{\Sigma}$ is drawn from an inverted Wishart distribution with $HT+\nu_\Sigma$ degrees of freedom:

$$\boldsymbol{\Sigma}^{-1(t)} | \boldsymbol{\omega}^{(t)}, \zeta^{(t)}, \mathbf{y}^{*(t)}, \mathbf{z}^{*(t)}, \boldsymbol{\delta}^{(t)}, \boldsymbol{\psi}^{(t)}, \mathbf{V}^{(t)}, \mathbf{V}_\Sigma, \nu_\Sigma \sim Wish\left(HT + \nu_\Sigma, \left(\mathbf{V}_\Sigma + \begin{bmatrix} \boldsymbol{\varepsilon} \\ \mathbf{u} \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}' & \mathbf{u}' \end{bmatrix}\right)^{-1}\right).$$

9. $\mathbf{V}$ is drawn from an inverted Wishart distribution with $H+\nu_\alpha$ degrees of freedom:

$$\mathbf{V}^{-1(t)} | \boldsymbol{\alpha}^{(t)}, \boldsymbol{\iota}^{(t)}, \boldsymbol{\delta}^{(t)}, \boldsymbol{\psi}^{(t)}, \overline{\mathbf{V}}, \nu_\alpha \sim Wish\left(H + \nu_\alpha, \left(\overline{\mathbf{V}} + \begin{bmatrix} \boldsymbol{\xi} \\ \boldsymbol{\tau} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}' & \boldsymbol{\tau}' \end{bmatrix}\right)^{-1}\right).$$

9. $\boldsymbol{\Omega}$ is drawn from an inverted Wishart distribution with $H+\nu_\omega$ degrees of freedom.

$$\boldsymbol{\Omega}^{-1(t)} | \boldsymbol{\omega}^{(t)}, \overline{\boldsymbol{\omega}}^{(t)}, \zeta^{(t)}, \overline{\zeta}^{(t)}, \mathbf{V}_\Omega, \nu_\Omega \sim Wish\left(H + \nu_\Omega, \left(\mathbf{V}_\Omega + \sum_{h=1}^{H} \left[\begin{pmatrix} \boldsymbol{\omega}_h \\ \zeta_h \end{pmatrix} - \begin{pmatrix} \overline{\boldsymbol{\omega}} \\ \overline{\zeta} \end{pmatrix}\right] \left[\begin{pmatrix} \boldsymbol{\omega}_h \\ \zeta_h \end{pmatrix} - \begin{pmatrix} \overline{\boldsymbol{\omega}} \\ \overline{\zeta} \end{pmatrix}\right]'\right)^{-1}\right)$$

.

For identification purposes we need ones on the diagonal of incidence equation error matrix $\boldsymbol{\Sigma}_{22} = E(\mathbf{u}_{ht}\mathbf{u}_{ht}')$ (while $\boldsymbol{\Sigma}$ is positive definite), and we need to rescale the parameters from the incidence equation **re**lative to $\boldsymbol{\Sigma}_{22}$. To achieve this, we follow the procedure proposed by Edwards and Allenby (2003) and Rossi, Allenby, and McCulloch (2005, p. 108). That is, we do not impose any restrictions when drawing $\boldsymbol{\Sigma}$. Instead, we postprocess the draws using two diagonal matrices $\mathbf{C}_1 = \begin{bmatrix} \mathbf{I}_S & 0 \\ 0 & \mathrm{diag}(\boldsymbol{\Sigma}_{22})^{-\frac{1}{2}} \end{bmatrix}$ and $\mathbf{C}_2 = \begin{bmatrix} \mathbf{I}_S & 0 \\ 0 & (1/\sqrt{\boldsymbol{\Sigma}_{22}^{(1,1)}}) \cdot \mathbf{I}_S \end{bmatrix}$, where $\Sigma_{22}^{(1,1)}$ is the upper left diagonal element of $\boldsymbol{\Sigma}_{22}$. After completing the Gibbs chain, we calculate $\boldsymbol{\Sigma}^* = \mathbf{C}_1\boldsymbol{\Sigma}\mathbf{C}_1'$ for each saved Gibbs draw, monitor its convergence, and use it for inference purposes. Analogously, for inferences we use $\mathbf{V}^* = \mathbf{C}_1\mathbf{V}\mathbf{C}_1'$, $(\boldsymbol{\iota}', \boldsymbol{\alpha}')^{*\prime} = \mathbf{C}_1 \cdot (\boldsymbol{\iota}', \boldsymbol{\alpha}')'$; $(\boldsymbol{\delta}', \boldsymbol{\psi}')^{*\prime} = \mathbf{C}_1 \cdot (\boldsymbol{\delta}', \boldsymbol{\psi}')'$; $(\boldsymbol{\omega}', \zeta')^{*\prime} = \mathbf{C}_2 \cdot (\boldsymbol{\omega}', \zeta')'$; $(\overline{\boldsymbol{\omega}}', \overline{\zeta}')^{*\prime} = \mathbf{C}_2 \cdot (\overline{\boldsymbol{\omega}}', \overline{\zeta}')'$, and $\boldsymbol{\Omega}^* = \mathbf{C}_2\boldsymbol{\Omega}\mathbf{C}_2'$.

Prior distributions

The second implementation step is to specify prior distributions for the parameters of interest. Note that the priors are set to be non-informative so that inferences are driven by the data.

The prior distribution of $[\boldsymbol{\delta}', \boldsymbol{\psi}]'$ is N($[\bar{\boldsymbol{\delta}}', \boldsymbol{\psi}']'$, $V_{\delta\psi}$), where $[\bar{\boldsymbol{\delta}}', \boldsymbol{\psi}']' = 0$ and $\mathbf{V}_{\delta\psi}$ = diag(10^3).

The prior distribution of $[\bar{\boldsymbol{\omega}}', \bar{\boldsymbol{\zeta}}']'$ is N($[\bar{\bar{\boldsymbol{\omega}}}', \bar{\bar{\boldsymbol{\zeta}}}']'$, $\bar{\boldsymbol{\Omega}}$), where $[\bar{\bar{\boldsymbol{\omega}}}', \bar{\bar{\boldsymbol{\zeta}}}']$=0 and $\bar{\boldsymbol{\Omega}}$ = diag(10^3).

The prior distribution of $\boldsymbol{\Sigma}^{-1}$ is Wishart: W(ν_Σ,$\mathbf{V}_\Sigma$), where $\nu_\Sigma = 2S+2$ and $\mathbf{V}_\Sigma$ = diag(10^{-3}).

The prior distribution of $\mathbf{V}^{-1}$ is Wishart: W(ν_α, $\bar{\mathbf{V}}$), where $\nu_\alpha = 2S+2$ and $\bar{\mathbf{V}}$ = diag(10^{-3}).

The prior distribution of $\boldsymbol{\Omega}^{-1}$ is Wishart: W(ν_ω, $\mathbf{V}_\Omega$), where $\nu_\omega = |\bar{\boldsymbol{\omega}}|+|\bar{\boldsymbol{\zeta}}|+2$ and $\mathbf{V}_\Omega$ = diag(10^{-3}).

Initial values

The third implementation step is to set initial values for the parameters of the marginal distributions. The starting values for $\boldsymbol{\omega}$ and $\boldsymbol{\delta}$ are computed by OLS, using ln(y_{bit}) as the dependent variable of the regression. The covariance matrix, $\boldsymbol{\Sigma}_{11}$, is initiated by computing the sample covariances of this regression's residuals. In a similar fashion, the starting values for the patronage equation parameters, $\boldsymbol{\zeta}$, are computed by OLS, using z_{hit} as the dependent variable, and the residuals from this regression, u_{bit}, are used to compute the sample correlations, which serve as the initial value for $\boldsymbol{\Sigma}_{22}$.

The final step is to generate N_1+N_2 random draws from the conditional distributions. We use a "burn in" of N_1=10,000 iterations, after which the draws have clearly converged based on visual checks. To reduce autocorrelation in the MCMC draws, we "thin the line," using every 50th draw in the final N_2 = 10,000 draws for our estimation. In this way, 200 draws are used to estimate marginal posterior distributions of the parameters of interest. Test runs of our Gauss implementation of the MCMC draws show that we can retrieve parameters used to simulate artificial data.

References

Edwards, Yancy D. and Greg M. Allenby (2003), "Multivariate Analysis of Multiple Response Data," *Journal of Marketing Research*, 40 (August), 321–34.

Rossi, Peter E., Greg M. Allenby, and Robert McCulloch (2005), *Bayesian Statistics and Marketing*. Chichester, West Sussex: John Wiley & Sons.

Web Appendix B: Time-Series Analyses of Aggregated Data

This appendix summarizes additional time-series analyses of the aggregated data series of weekly price and the stock prices of publicly traded retailers.

1. Competitive reaction functions from weekly price analyses

Following Leeflang and Wittink (1996), we regress the log of basket price on the log of prices at all other retailers, lags of all log prices, and exogenous variables (intercept and dummy variables for the year-end and Easter weeks, to control for seasonality patterns in prices). We estimate these reaction functions on two data periods: before and after the start of the price war. The results (competitive interactions significant at the 10% level) are displayed in table B1-B2.

Table B1: Competitive interactions before start price war (10% significance level)

	Impact on (Dependent Variables):					
Impact of (explanatory variables):	$logP_t$ (AH)	$logP_t$ (Aldi)	$logP_t$ (C1000)	$logP_t$ (Edah)	$logP_t$ (Lidl)	$logP_t$ (SdB)
$logP_t$ (AH)	X		.16			
$logP_t$ (Aldi)		X		.18		
$logP_t$ (C1000)	.25		X			.55
$logP_t$ (Edah)		.20		X		
$logP_t$ (Lidl)					X	
$logP_t$ (SdB)			.31			X
$logP_{t-1}$(AH)				.22		
$logP_{t-1}$(Aldi)		.21				
$logP_{t-1}$(C1000)			.19			
$logP_{t-1}$(Edah)	.22			.32		
$logP_{t-1}$(Lidl)		.19			.19	
$logP_{t-1}$(SdB)						
Constant					-.34	
Week1			-.06	-.05	.10	
Week51	.05	.06				
Week 52						
Easter	.03	.04	.02			
R^2	*.66*	*.40*	*.74*	*.53*	*.28*	*.60*

Note first that competitive interactions are indeed observed among retail chains, and that the estimated magnitudes (all below .60) have face validity. Importantly, we observe more (significant) interactions after the start of the price war for every retailer. For Albert Heijn, for instance, the number of significant competitive interactions increases from 2 before the price war to 4 after the start of the price war. This increased competitive pattern is consistent with the definitional conditions of a price war, as opposed to a heavy-promotion period.

Table B2: Competitive interactions after start price war (10% significance level)

	$\log P_t$ (AH)	$\log P_t$ (Aldi)	$\log P_t$ (C1000)	$\log P_t$ (Edah)	$\log P_t$ (Lidl)	$\log P_t$ (SdB)
$\log P_t$ (AH)	X	.24	.14	.19		.14
$\log P_t$ (Aldi)	.43	X			.24	.19
$\log P_t$ (C1000)	.19		X			.28
$\log P_t$ (Edah)				X		.15
$\log P_t$ (Lidl)		.11			X	
$\log P_t$ (SdB)	.24	.18	.35	.49		X
$\log P_{t-1}$(AH)	.21					
$\log P_{t-1}$(Aldi)		.38				
$\log P_{t-1}$(C1000)				.26		
$\log P_{t-1}$(Edah)	.14					
$\log P_{t-1}$(Lidl)			.17		.21	
$\log P_{t-1}$(SdB)						.30
Constant	.20	-.36		-.40	-.23	.17
Week1	-.07					
Week51	.05	.03		.05		
Week 52	.04		.03	.05	.04	
Easter	.05					
R^2	*.70*	*.56*	*.47*	*.32*	*.33*	*.63*

2. Structural break analysis of retailer stock prices

We obtained information on the financial performance of the companies owning the retailers in our study. Such data is only available for the Dutch firms: Schuitema (owning C1000), Laurus (owning Edah and Super de Boer) and Koninklijke Ahold/Albert Heijn Arena (owning Albert Heijn). For the German hard discounters Aldi en Lidl, there is hardly any financial information available: they are not listed on any stock exchange, and annual reports are aggregated far beyond the Dutch performance level. In what follows, we consider stock option prices, combined with the (limited) financial performance information available from annual reports

We consider weekly “adjusted stock prices” (a technical correction of opening prices) for Koninklijke Ahold, Schuitema, and Laurus, on the Amsterdam Stock exchange (AEX). To control for overall market influences, we divide each company’s stock price by the overall stock market index (AEX). Figures B1 displays these stock price indices over time. We note a clear trend in each index, as well as a clear dip for Ahold in Februari 2003 – this dip coincides with the accounting scandal.

We regress each relative stock price index on (1) a trend, (2) the pulse variable representing each Price War Round and (3) the step variable representing each price war round (accumulative for each consecutive round). The following findings emerge:

(1) While the trend for Schuitema is positive, it is negative for Laurus and Ahold.

(2) Once the trend is accounted for, the price war does not significantly influence the relative stock prices of Laurus or Schuitema[1]. We do, however, find a significant effect for Ahold, which is negative but insignificant (p=0.15) for the pulse price round variable, but positive (p=0.001) for the cumulative ‘long-term’ step price round variable.

(3) Including a separate ‘scandal’ variable representing the accounting scandal does not change this result, except that it makes the pulse variable somewhat more significant (p=.12).

Adding the interaction of the trend with the price war yields results consistent with the annual revenue changes obtained from the retailers’ annual reports. For Laurus, the start of the price war marks the start of a negative trend, while Schuitema’s stock price shows no significant effect. For Albert Heijn, we find that the price war stops the negative trend in the stock price[2]. In sum, it appears that price war may have allowed AH to restore confidence in the future of the company, and thus reduce the decline in stock market prices.

[1] A unit root analysis revealed that we could reject the hypothesis of a unit root for the stock price indices of Ahold and Laurus, but not for Schuitema. For Schuitema, we therefore use the differenced series as the dependent variable.

[2] Including an additional autoregressive component in the model did not change this result.

Figure B1: Stock Price Indices

Reference

Leeflang, Peter S.H. and Dick R. Wittink (1996), "Competitive Reaction Versus Consumer Response: Do Managers Overreact?," International Journal of Research in Marketing, 13 (2), 103-20.

Appendix C: Decomposing Revenue Changes

Expected revenues of household *h* in chain *i* in week *t* are given by:

(C1) $E(R_{hit}) = Pr(z^*_{hit} = 1)E(y^*_{hit})$ where

(C2) $\Pr(z^*_{hit} = 1) = \Phi(\iota_{hi} + \mathbf{x}'_{hit}\boldsymbol{\zeta}_{ht})$ and

(C3) $E(y^*_{hit}) = \exp(\alpha_{hi} + \mathbf{v}'_{hit}\boldsymbol{\omega}_{ht} + \frac{1}{2}\Sigma_{ii})$,

where the term $\frac{1}{2}\Sigma_{ii}$ is the Goldberger correction (see e.g., Hanssens, Parsons and Schultz (2001, p. 395) and Σ_{ii} is the element (i,i) of the error covariance matrix Σ.

If we define $t = 0$ for the quarter preceding the price war, and $t = 1$ for a quarter after the price war, we can write the change in revenues as:

$$\begin{aligned}
&\Delta E(R_{hi}) = E(R_{hi1}) - E(R_{hi0}) \\
\text{(C4)}\quad &= Pr(z^*_{hi0} = 1) \cdot \left(E(y^*_{hi1}) - E(y^*_{hi0})\right) + \left(Pr(z^*_{hi1} = 1) - Pr(z^*_{hi0} = 1)\right) \cdot E(y^*_{hi1}) \\
&\equiv Pr(z^*_{hi0} = 1) \cdot \Delta E(y^*_{hi}) + \Delta Pr(z^*_{hi} = 1) \cdot E(y^*_{hi1})
\end{aligned}$$

Hence the right-hand side has two components: the original incidence probability times the change in expenditures , plus a change in incidence probability times (new) expenditures. We can rewrite the change in incidence probability as:

$$\begin{aligned}
&\Pr(z^*_{hi1} = 1) - \Pr(z^*_{hi0} = 1) \\
&= \Phi(\iota_{hi} + \mathbf{x}'_{hi1}\boldsymbol{\zeta}_{h1}) - \Phi(\iota_{hi} + \mathbf{x}'_{hi0}\boldsymbol{\zeta}_{h0}) \\
&= \int_{-\infty}^{\iota_{hi}+\mathbf{x}'_{hi1}\boldsymbol{\zeta}_{h1}} e^{-\frac{1}{2}t^2}\,dt - \int_{-\infty}^{\iota_{hi}+\mathbf{x}'_{hi0}\boldsymbol{\zeta}_{h0}} e^{-\frac{1}{2}t^2}\,dt \\
\text{(C5)}\quad &= \int_{\iota_{hi}+\mathbf{x}'_{hi0}\boldsymbol{\zeta}_{h0}}^{\iota_{hi}+\mathbf{x}'_{hi1}\boldsymbol{\zeta}_{h1}} e^{-\frac{1}{2}t^2}\,dt \\
&= \int_{\iota_{hi}+\mathbf{x}'_{hi0}\boldsymbol{\zeta}_{h0}}^{\iota_{hi}+\mathbf{x}'_{hi1}\boldsymbol{\zeta}_{h0}} e^{-\frac{1}{2}t^2}\,dt + \int_{\iota_{hi}+\mathbf{x}'_{hi1}\boldsymbol{\zeta}_{h0}}^{\iota_{hi}+\mathbf{x}'_{hi1}\boldsymbol{\zeta}_{h1}} e^{-\frac{1}{2}t^2}\,dt \\
&\equiv \Delta\Pr(z^*_{hi} = 1 \mid \Delta\mathbf{x}_{hi}, \boldsymbol{\zeta}_{h0}) + \Delta\Pr(z^*_{hi} = 1 \mid \mathbf{x}_{hi1}, \Delta\boldsymbol{\zeta}_{h})
\end{aligned}$$

Hence the right-hand side decomposes the change in incidence probability in two components: a part due to a change in the independent variables, and a part due to changes in the parameters.

We can rewrite the change in conditional expenditures as:

(C6) $E(y^*_{hi1}) - E(y^*_{hi0}) = \exp(\alpha_{hi} + \mathbf{v}'_{hi1}\boldsymbol{\omega}_{h1} + \frac{1}{2}\Sigma_{ii}) - \exp(\alpha_{hi} + \mathbf{v}'_{hi0}\boldsymbol{\omega}_{h0} + \frac{1}{2}\Sigma_{ii})$

If we write the joint effect of the independent variables as $\mathbf{v}'_{hi}\boldsymbol{\omega}_{hi} = \mathbf{v}'_{hi}\tilde{\boldsymbol{\omega}}_h + \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h - \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h$, where $\overline{v}'_{hi}$ and $\overline{\omega}_h$ are averages across the pre- and post price war periods, we can rewrite (C6) as:

$$= \exp(\alpha_{hi} + \mathbf{v}'_{hi1}\boldsymbol{\omega}_{h1} + \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h - \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h + \frac{1}{2}\Sigma_{ii}) - \exp(\alpha_{hi} + \mathbf{v}'_{hi0}\boldsymbol{\omega}_{h0} + \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h - \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h + \frac{1}{2}\Sigma_{ii})$$

$$= \exp(\alpha_{hi} + \frac{1}{2}\Sigma_{ii} + \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h) * \left(\exp(\tilde{\mathbf{v}}'_{hi1}\tilde{\boldsymbol{\omega}}_{h1} - \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h) - \exp(\tilde{\mathbf{v}}'_{hi0}\tilde{\boldsymbol{\omega}}_{h0} - \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h)\right)$$

Using a linear Taylor series approximation $\exp(x) = 1 + x + o(n^2)$ we can rewrite (C6) as:

(C7) $E(y^*_{hi1}) - E(y^*_{hi0})$

$$\approx \exp(\alpha_{hi} + \frac{1}{2}\Sigma_{ii} + \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h) * \left(1 + \mathbf{v}'_{hi1}\boldsymbol{\omega}_{h1} - \overline{v}'_{hi}\overline{\omega}_h - 1 - \mathbf{v}'_{hi0}\boldsymbol{\omega}_{h0} + \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h + o(n^2)\right)$$

$$= \exp(\alpha_{hi} + \frac{1}{2}\Sigma_{ii} + \overline{\mathbf{v}}'_{hi}\overline{\boldsymbol{\omega}}_h) * \left(\mathbf{v}'_{hi1}\boldsymbol{\omega}_{h1} - \mathbf{v}'_{hi0}\boldsymbol{\omega}_{h0} + o(n^2)\right)$$

$$= \exp(\alpha_{hi} + \frac{1}{2}\Sigma_{ii} + \overline{\mathbf{v}}'_{hi1}\overline{\boldsymbol{\omega}}_{h1}) * \left((\mathbf{v}'_{hi1} - \mathbf{v}'_{hi0})\boldsymbol{\omega}_{h0} + \mathbf{v}'_{hi1}(\boldsymbol{\omega}_{h1} - \boldsymbol{\omega}_{h0}) + o(n^2)\right)$$

$$\equiv \Delta E(y^*_{hi} \mid \Delta\mathbf{v}_{hi}, \boldsymbol{\omega}_{h0}) + \Delta E(y^*_{hi} \mid \Delta\boldsymbol{\omega}_h, \mathbf{v}_{hi1}) + o(n^2)$$

By substituting (C5) and (C7) into (C4), we can write the change in expected store expenditures as:

(C8)

$$\Delta E(R_{hi})$$

$$= \underbrace{\Pr(z^*_{hi0}=1)\Delta E(y^*_{hi} \mid \Delta\mathbf{v}_{hi}, \boldsymbol{\omega}_{h0})}_{\text{(a)Expenditure change due to changed independent variables}} + \underbrace{\Pr(z^*_{hi0}=1)\Delta E(y^*_{hi} \mid \Delta\boldsymbol{\omega}_h, \mathbf{v}_{hi1})}_{\text{(b) Expenditure change due to changed coefficients}} + \underbrace{\Pr(z^*_{hi0}=1)o(n^2)}_{\text{(c) Expenditure approximation error}}$$

$$+ \underbrace{\Delta\Pr(z^*_{hi}=1 \mid \Delta\mathbf{x}_{hi}, \boldsymbol{\zeta}_{h0})E(y^*_{hi1})}_{\text{(d) Incidence change due to change independent variables}} + \underbrace{\Delta\Pr(z^*_{hi}=1 \mid \mathbf{x}_{hi1}, \Delta\boldsymbol{\zeta}_h)E(y^*_{hi1})}_{\text{(e) Incidence change dueto changed coefficients}}$$

We use equilibrium estimates for conditional expenditures and incidence by repeated substitution. In particular, in the expenditures model we use four lags of the predicted unconditional expenditures (product of incidence probability and conditional expenditures) on the right hand side to predict for period t, t+1, etcetera, until convergence. In the incidence probability model we repeatedly use lagged predicted incidence probabilities as independent variables, until convergence.