Technical Appendix

for Estimating Price Elasticities with Theory-Based Priors

To make the procedure as general as possible we rewrite our model in SUR form:

$$y_s \sim N(X_s\beta_s,\ \Sigma_s \otimes I_T),\quad \Sigma_s^{-1} \sim W(\nu_\Sigma,\ \bar{V}_\Sigma^{-1})$$

In this case the *s* subscript denotes an individual store, and the dimension of the $\boldsymbol{y}_s$ vector is *M* brands by *T* weeks. In rewriting the model we have implicitly stacked the vector of observations for each brand on top of one another in the following manner:

$$\boldsymbol{y_s} = \begin{bmatrix} \boldsymbol{q_{1s}} \\ \vdots \\ \boldsymbol{q_{Ms}} \end{bmatrix},\quad \boldsymbol{q_{is}} = \begin{bmatrix} q_{i1s} \\ \vdots \\ q_{iTs} \end{bmatrix},\quad X_s = \begin{bmatrix} X_{1s} & & \\ & \ddots & \\ & & X_{Ms} \end{bmatrix},$$

$$X_{is} = \begin{bmatrix} 1 & \ln(x_{1s}/P_{1s}) & p_{11s} & \cdots & p_{M1s} & f_{i1s} & d_{i1s} \\ 1 & \ln(x_{2s}/P_{2s}) & p_{12s} & \cdots & p_{M2s} & f_{12s} & d_{i2s} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 1 & \ln(x_{Ts}/P_{ts}) & p_{1Ts} & \cdots & p_{MTs} & f_{1Ts} & d_{iTs} \end{bmatrix}$$

The second stage of our hierarchical model refers to the hyper-distribution from which the vector of parameters for each store is drawn:

$$vec(\mathrm{E}_s) \mid \mu_s, \phi_s \sim N(vec(\bar{\mathrm{E}}_s),\ \Delta) \text{ for } s=1,\cdots,S \quad , \quad \Delta^{-1} \sim W(\nu_\Delta,\ V_\Delta^{-1})$$

where the expected price elasticity matrix is the restricted one implied by an additive utility model:

$$\bar{\mathrm{E}}_s = \phi_s\, diag(\mu_s) - \phi_s \mu_s (\mu_s \odot \boldsymbol{w}_s)'$$

The remaining parameters are drawn from:

$$\beta_s \sim N(\bar{\beta}_s,\ \Lambda) \text{ for } s=1,\cdots,S \quad , \quad \Lambda^{-1} \sim W(\nu_\Lambda,\ V_\Lambda^{-1})$$

The third stage of our model expresses the prior on the hyper-distribution:

$$\bar{\beta} \sim N(\theta,\ V_\theta)$$

Estimation Using the Gibbs Sampler

Our purpose is to compute the posterior distribution of the model parameters. The posterior distribution contains all the information from our sample given our distributional assumptions. From the posterior distribution we can compute the means, which are commonly used as point estimates, along with any other measures of the distribution that are of interest. The following data and parameters are supplied by the analyst:

$$X \,,\, Y \quad ; \quad \theta \,,\, V_\theta \,,\, \nu_\Lambda \,,\, V_\Lambda \,,\, \nu_\Delta \,,\, V_\Delta \,,\, \nu_\Sigma \,,\, V_\Sigma \,,\, \nu_\phi \,,\, V_\Phi \,,\, \bar{\bar{\phi}} \,,\, V_{\bar{\phi}}$$

The general procedure for finding the marginal posterior distribution is to compute the joint posterior and then integrate out all parameters except those of interest. In this case the joint distribution of our model can be written as:

$$p(\beta_1,\cdots,\beta_S,E_1,\cdots,E_s,\phi_1,\cdots,\phi_S,\Sigma_1,\cdots,\Sigma_S,\bar{\beta},\Lambda,\Delta,\bar{\phi},\lambda_\phi \mid \text{data, priors}) \propto$$

$$\prod_{s=1}^{S} like(\beta_s,E_s,\Sigma_s \mid \Delta,\Lambda)\; p(\bar{\beta} \mid \beta_1,\cdots,\beta_S,\Lambda)\; p(\Lambda)\; p(\Delta)\; p(\theta)\; p(\bar{\phi})\; p(\lambda_\phi)$$

If we wanted to find the marginal posterior distribution of θ we would need to solve:

$$\begin{aligned} & p(\bar{\beta} \mid \theta,V_\theta,\nu_\Lambda,V_\Lambda,\nu_\Delta,V_\Delta,\nu_\Sigma,V_\Sigma,\nu_\phi,V_\phi,\bar{\bar{\phi}},V_{\bar{\phi}}) \\ &= \int p(\beta,\epsilon,\Sigma,\phi,\bar{\beta},\Lambda,\Delta,\bar{\phi},\lambda_\phi)\, d\beta\, dE\, d\Sigma\, d\phi\, d\bar{\phi}\, d\lambda_\phi \end{aligned}$$

The analytic solution to this integral is not known even with natural conjugate priors. To understand the difficulty in solving this integral, we refer the reader to the simpler case of trying to solve a single stage SUR model (Zellner 1971, pp. 240-6) for which the analytic solution is not known either. Therefore we will have to rely upon numerical procedures to find the solution. Unfortunately the high dimension of the integral makes it difficult to find a numerical solution using conventional numerical integration techniques.

An alternate method is through the use the Gibbs sampler. The Gibbs sampler requires the solution of the conditional distributions, which can be easily derived due to the hierarchical structure of the model. For a good introduction to the Gibbs sampler see Casella and George (1992). We do not advocate the use of Gibbs sampler based on computational efficiency, instead we advocate its use because of its ease of implementation. The most desirable solution would be an analytical one, but given that this solution does not exist in closed form we satisfy ourselves with a numerical solution.

The Gibbs sampler employed in this paper requires sequentially randomly sampling from each of the conditional distributions. It has been shown by Gelfand and Smith (1990) and Gelfand et al (1990) that this draws converge to the posterior marginal distributions. The general outline of the procedure is:

1. Select starting values for the parameters of the marginal posterior distributions. In our practice the least squares estimates of these parameters provide good starting points.
2. Generate M_1+M_2 sets of random numbers with each set being drawn in the following manner:

$$\beta_s^{(k)} \veebar p(\beta_s \mid \mathrm{E}_s^{(k-1)}, \Sigma_s^{(k-1)}, \cdots) \quad \text{for } s = 1, \cdots, S$$

$$\mathrm{E}_s^{(k)} \veebar p(\mathrm{E}_s \mid \beta_s^{(k)}, \Sigma_s^{(k-1)}, \phi_s^{(k-1)}, \cdots) \quad \text{for } s = 1, \cdots, S$$

$$\phi_s^{(k)} \veebar p(\phi_s \mid \beta_s^{(k)}, \mathrm{E}_s^{(k)}, \cdots) \quad \text{for } s = 1, \cdots, S$$

$$\Sigma^{(k)} \veebar p(\Sigma_s \mid \beta_s^{(k)}, \mathrm{E}_s^{(k)}, \cdots)$$

$$\bar{\beta}^{(k)} \veebar p(\bar{\beta} \mid \beta_1^{(k)}, \cdots, \beta_S^{(k)}, \Lambda^{(k-1)}, \cdots)$$

$$\bar{\phi}^{(k)} \veebar p(\bar{\phi} \mid \phi_1^{(k)}, \cdots, \phi_S^{(k)}, \lambda_\phi^{(k-1)}, \cdots)$$

$$\Delta^{(k)} \veebar p(\Delta \mid \mathrm{E}_1^{(k)}, \cdots, \mathrm{E}_S^{(k)}, \beta_1^{(k)}, \cdots, \beta_S^{(k)}, \phi_1^{(k)}, \cdots, \phi_S^{(k)}, \cdots)$$

$$\Lambda^{(k)} \veebar p(\Lambda \mid \beta_1^{(k)}, \cdots, \beta_S^{(k)}, \bar{\beta}^{(k)}, \cdots)$$

$$\lambda_\phi^{(k)} \veebar p(\lambda_\phi \mid \phi_1^{(k)}, \cdots, \phi_S^{(k)}, \bar{\phi}^{(k)}, \cdots)$$

Where the symbol $x \veebar p(x)$ means that the x is a simulated realization or draw from the density $p(x)$ and k denotes the iteration number. The above conditional distributions are understood to also depend upon the prior parameters and the data.

3. Use the last M_2 sets of draws to estimate the posterior marginal distributions.

This means that the problem reduces to solving the conditional distributions of each of the parameters in the posterior distribution. These solutions are readily available due to the hierarchical structure of our model and the affine nature of the normal and Wishart distributions. The solution of the conditional densities are:

1. Draw the parameter vector in the first-stage in two parts to avoid the nonlinearity induced by the additive separable prior:

    (a) Since we know the price elasticities, we can rewrite the model as below:

$$\left[\ln(q_{its}) - \sum_j \epsilon_{ijs} \ln(p_{jts})\right] = \alpha_{is} + \mu_{is} \ln(x/P_{ts}) + \theta_{is} f_{its} + \phi_{is} d_{its} + e_{its}$$

    The $\boldsymbol{\beta}_s$ vector can be drawn using the usual SUR result.

    (b) Since we know the $\boldsymbol{\beta}_s$ vector we can rewrite the model as below:

$$\left[\ln(q_{its}) - \left\{\alpha_{is} + \mu_{is} \ln(x/P_{ts}) + \theta_{is} f_{its} + \psi_{is} d_{its}\right\}\right] = \sum_j \epsilon_{ijs} \ln(p_{jts}) + e_{its}$$

    The E_s matrix can be drawn using the usual multivariate regression result.

2. Draw the ϕ parameter. Notice that conditional upon $\mathbf{E}_s$ and $\boldsymbol{\mu}_s$ we have the following univariate regression:

$$\epsilon_{ijs} = (\delta_{ij} - \mu_{js} w_{js}) \mu_{is} \phi_s + u_{ijs} \;,\; \boldsymbol{u}_s \sim N(\mathbf{0}, \Delta)$$

    Hence, ϕ_s can be drawn using the usual univariate regression result.

3. Σ_s is drawn from an inverted wishart distribution

$$\Sigma_s^{-1} \sim W(\nu_\Sigma + T_s \;,\; (V_\Sigma + \hat{E}_s^{\prime} \hat{E}_s)^{-1}) \quad , \quad \hat{E}_s[,i] = \boldsymbol{y}_{is} - X_{is}^{\prime} \boldsymbol{\beta}_{is}$$

4. $\bar{\beta}$ is a multivariate regression

$$\bar{\beta} \sim N(\; H(\textstyle\sum_s \Lambda^{-1} \beta_s + V_\theta^{-1} \theta) \;,\; H \;) \quad , \quad H = (S\Lambda^{-1} + V_\theta^{-1})^{-1}$$

$\bar{\phi}$ is a univariate regression

$$\bar{\phi} \sim N\left(H\left[\frac{\sum_s \phi_s}{\lambda_\phi} + \frac{\bar{\bar{\phi}}}{V_{\bar{\phi}}} \right], H \right) \quad , \quad H = \left[\frac{S}{\lambda_\phi} + \frac{1}{V_{\bar{\phi}}} \right]^{-1}$$

5. Since Δ and Λ are independent they can be drawn separately from inverted Wishart distributions:

$$\Delta^{-1} \sim W(\nu_\Delta + S \,,\, V_\Delta + \sum (\boldsymbol{vec}(\mathrm{E})_s - vec(\bar{\mathrm{E}}_s))(\boldsymbol{vec}(\mathrm{E})_s - vec(\bar{\mathrm{E}}_s))')$$

$$\Lambda^{-1} \sim W(\nu_\Lambda + S \,,\, V_\Lambda + \sum (\beta_s - \bar{\beta}_s))(\beta_s - \bar{\beta}_s))')$$

$$\lambda_\phi^{-1} \sim W(\nu_\phi + S \,,\, V_\phi + \sum (\phi_s - \bar{\phi})(\phi_s - \bar{\phi})')$$