Bayesian Estimation of a Dynamic Model of Two-Sided Markets: Application to the U.S. Video Games

Online Appendix

Yiyi Zhou

A Substitution Structure of the Affiliated Product Market

I specify three different regression models to test the substitution between sports games, and I display the results in Table 1.

In the first regression, the dependent variable is a game title's price and the independent variables include: (i) the substitution within a submarket which is measured by the number of existing games in the same market; (ii) the substitution from other submarkets which is measured by the number of existing games in all other submarkets; (iii) observed characteristics including the online rating score and the game age measured by the months after release; (iv) the market size measured by the log of the console owners; and (v) monthly dummies. The first important result is that the price of an existing game is lower by \$0.148 if additional game is released in the same submarket and this impact is statistically significant. It implies that the substitution within a submarket is strong. The second important result is that the substitution from other submarkets is not statistically significant, which implies that submarkets are separate from each other. Besides, the game price is increasing in its online rating score, declining in its age and increasing in the number of console owners.

In the second regression, the dependent variable is the log of a game title's unit sales (measured in thousands). The independent variables are the same as in the first model except that I include an extra independent variable, the current price. To address the potential endogeneity of the price, I use the lagged prices as instruments for the current price. The second column in table 1 lists the estimation results which are consistent with the assumption of strong competition within a market and weak competition across markets.

The last regression mimics Nair (2007).[1] I still use the log of a game title's unit sales as the dependent variable. However, I use the log of the total unit sales of all existing games in the same submarket to measure the substitution within a submarket, and the log of the total unit sales of all existing games in other submarkets to measure the substitution from other submarkets. To address the potential endogeneity problem, I use the lagged prices as instruments for the current price, the number of existing games within the a submarket as an instrument for the within-submarket sales,

---

[1] In Nair (2007), the dependent variable is $ln(s_{jt}/s_{0t})$, where s_{jt} is the market share of game j and s_{0t} is the share of the outside good. He uses $ln(s_{jt|g}/s_{0t})$ to measure the effect within a market, where $s_{jt|g}$ is the share of units sales of the game within its genre, g.

and the number of existing games in other submarkets as an instrument for the outside-submarket sales. The results also show that the substitution effect within a submarket is significant while the substitution from games sold in other markets is insignificant.

Table 1: Empirical Results of Testing Games' Substitution Structure

	Model I Price ($)	Model II $ln(q_{jt})$	Model III $ln(q_{jt})$
Game Price		-.0126***	-.009***
		(.0019)	(.002)
Substitution within a submarket	-.148**	-.0359***	-.219***
	(.060)	(.0045)	(.019)
Substitution from other submarkets	-.011	.0002	-.012
	(0.066)	(.0004)	(.024)
Online rating score	1.417***	.3419***	.261***
	(.047)	(.0099)	(.010)
Product age (months)	-1.141***	-.1988***	-.199***
	(.015)	(.0099)	(.004)
Age square	.013***	.0015***	.002***
	(.002)	(.0000)	(.000)
Market size (million)	3.353***	.2438***	.803***
	(.088)	(.0466)	(.034)
R-square	0.68	0.63	0.53
# Observations	13779	13024	12794

Notes: * indicates significance at 10 percent level; ** indicates significance at 5 percent level; and *** indicates significance at 1 percent level.

B Jacobian Matrix

$J_{(\mathfrak{q}_{\mathfrak{mt}}, p_{mt}) \to (\xi_{mt}, \varsigma_{mt})}$ is the Jacobian matrix corresponding to the transformation-of-variables from $(\mathfrak{q}_{\mathfrak{mt}}, p_{mt})$ to $(\xi_{mt}, \varsigma_{mt})$. Here, the Jacobian

$$J_{(\mathfrak{q},p)\to(\xi,\varsigma)} = \left\| \begin{array}{cc} \nabla_{\xi}\mathfrak{q} & \nabla_{\varsigma}\mathfrak{q} \\ \nabla_{\xi}\mathbf{p} & \nabla_{\varsigma}\mathbf{p} \end{array} \right\| = \left\| \begin{array}{cc} \nabla_{\xi}\mathfrak{q} & \nabla_{\varsigma}\mathfrak{q} \\ \mathbf{0} & \mathbf{I} \end{array} \right\| = \|\nabla_{\xi}\mathfrak{q}\| = J_{\mathfrak{q}\to\xi},$$

where the matrix elements are

$$\frac{\partial \mathfrak{q}_{jt}}{\partial \xi_{kt}} = \begin{cases} \sum_{i=1}^{I} n_{mit}\mathfrak{b}_{ijt}(1-\mathfrak{b}_{ijt}), & if\, k = j, \\ -\sum_{i=1}^{I} n_{mit}\mathfrak{b}_{ijt}\mathfrak{b}_{ikt}, & if\, k \neq j. \end{cases}$$

C A Monte Carlo Exercise: Model and Estimation Algorithm

I adopt the discrete choice dynamic model of Dube, Fox and Su (2012). There is a mass N_0 potential consumers at date $t = 1$. Consumers drop out of the market once they make a purchase. Abstracting from the supply side specifics, prices evolves over time as a function of the lagged prices of both firms:

$$p_{jt} = \eta_0 + \eta_1 p_{jt-1} + \eta_2 p_{-j,t-1} + \omega_{jt} = \mathbf{p}_{t-1}\eta_j + \varsigma_{jt},\ j = 1, ..., J \tag{1}$$

where ς_{jt} is a random supply shock.

A consumer with tastes indexed by $i = 1, ..., I$ derives the following flow utility from adopting product $j = 1, ..., J$ at time t:

$$u_{ijt}(p_t; \alpha_i) = \alpha_{ij}^x + \alpha_i^p p_{jt} + \xi_{jt} + \nu_{ijt},$$

where $\alpha_i = (\alpha_{ij}^x, \alpha_i^p)$ are taste parameter, ξ_{jt} is a time-varying product characteristics, and ν_{ijt} is an independent and identically distributed Type I extreme value utility shock. The expected flow utility from no-purchase are normalized to be zero. Assume that the demand shocks ξ_{jt} is jointly distributed with the supply shock ς_{jt}. Denote the distribution as $F_{\xi,\varsigma}(\xi, \varsigma)$. Furthermore, assume that shocks are normally distributed independent across time periods and markets, that is, $(\xi_t, \varsigma_t)' \sim N(\mathbf{0}, \Omega)$.

A discrete distribution with I mass points is used to characterize the consumer population's tastes at data $t = 1$. Let w_i denote the initial share of type-i consumers, with $\sum_{i=1}^{I} w_i = 1$. This heterogeneity allows for certain types of consumers to systematically purchase earlier than others. Let θ summarizes all parameters that we need to estimate, $\theta = (\alpha, \eta, w, \Omega)$.

Consumers are forward looking and have rational expectations in the sense that their beliefs about future prices coincide with (1). A consumer i's expected value of the outside option at a given state vector $\mathbf{p}$ can be defined by the following Bellman equation:

$$\begin{aligned} EV_i(\mathbf{p}; \theta) \quad &= \quad \beta \int log\Big\{ \sum_j exp\left[\alpha_{ij}^x + \alpha_i^p(\mathbf{p}\eta_j + \varsigma_j) + \xi_j\right] \\ &\qquad + exp\left[EV_i(\mathbf{p}\eta + \varsigma; \theta)\right] \Big\} dF_{\xi,\omega}(\xi, \varsigma). \end{aligned} \tag{2}$$

In a given period t, the market share of product j is

$$\mathfrak{s}_j(\mathbf{p}_t; \theta) = \sum_{i=1}^{I} w_{it} \frac{exp\left[\alpha_{ij}^x + \alpha_i^p p_{jt} + \xi_{jt}\right]}{exp\left[EV_i(\mathbf{p}_t; \theta)\right] + \sum_{j'=1}^{J} exp(\alpha_{j'}^x + \alpha_i^p p_{j't} + \xi_{j't})}, \tag{3}$$

for $j = 1, ..., J$. Here w_{it} is the share of remaining type i consumers who have not yet adopted at the beginning of period t.

The total remaining mass of consumers of type i who have not yet adopted at the beginning of period t, N_{it}, is

$$N_{it} = \begin{cases} N_0 w_i, & t = 1, \\ N_{it-1}\mathfrak{s}_{0i}(p_t;\theta), & t > 1, \end{cases}$$

where $\mathfrak{s}_{0i}(\mathbf{p}_t;\theta)$ is the fraction of consumers of type i who purchase the outside good. Hence, the shares of each consumer type are

$$w_{it} = \begin{cases} w_i, & t = 1, \\ \frac{N_{it}}{\sum_i N_{it}} & t > 1. \end{cases}$$

The assumption of finite types eases dynamic programming because there is only one unknown value of waiting function for each type.

As is standard in the literature, we assume the observed market shares equal to the predicted market shares, that is,

$$s_{jt} = \mathfrak{s}_j(\mathbf{p}_t, \xi_t; \theta), \tag{4}$$

for all $j = 1, ..., J$ and $t = 1, ...T$. Then, we can solve for the demand shock by inverting the market share functions, that is, $\xi_t(\mathbf{p}_t, \mathbf{s}_t;\theta) = \mathfrak{s}^{-1}(\mathbf{p}_t, \mathbf{s}_t;\theta)$.

The empirical model consist of the system (1) and (3),

$$\epsilon_t \equiv \begin{bmatrix} \xi_t \\ \varsigma_t \end{bmatrix} = \begin{bmatrix} \mathfrak{s}^{-1}(\mathbf{p}_t, \mathbf{s}_t;\theta) \\ p_t - p_{t-1}^{'}\eta \end{bmatrix}.$$

The multivariate normal distribution of ϵ_t induces the following density on the observable outcomes:

$$f(\mathbf{p}_t, \mathbf{s}_t;\theta) = \frac{1}{(2\pi)^J|\Omega|^{1/2}} exp\Big(-\frac{1}{2}\epsilon_t^{'}\Omega^{-1}\epsilon_t\Big)|J^{-1}_{t,(\mathfrak{s},p)\to(\xi,\varsigma)}|,$$

where $J_{t,(\mathfrak{s},\mathbf{p})\to(\xi,\varsigma)}$ is a $(2J \times 2J)$ Jacobian matrix corresponding to the transformation-of-variables from $(\mathbf{s}_t, \mathbf{p})$ to ϵ_t. where the Jacobian

$$J_{(\mathfrak{s},p)\to(\xi,\varsigma)} = \begin{Vmatrix} \nabla_\xi \mathfrak{s} & \nabla_\varsigma \mathfrak{s} \\ \nabla_\xi \mathbf{p} & \nabla_\varsigma \mathbf{p} \end{Vmatrix} = \begin{Vmatrix} \nabla_\xi \mathfrak{s} & \nabla_\varsigma \mathfrak{s} \\ \mathbf{0} & \mathbf{I} \end{Vmatrix} = \|\nabla_\xi \mathfrak{s}\| = J_{\mathfrak{s}\to\xi},$$

where the matrix elements are

$$\frac{\partial \mathfrak{s}_{jt}}{\partial \xi_{kt}} = \begin{cases} \sum_{i=1}^I w_{it}\mathfrak{s}_{ijt}(1-\mathfrak{s}_{ijt}), & if\, k = j, \\ -\sum_{i=1}^I w_{it}\mathfrak{s}_{jt}\mathfrak{s}_{kt}, & if\, k \neq j. \end{cases}$$

Suppose that all the data we have is the market shares and observed product prices over T time periods across M markets, denoted as $data = \{\{\mathbf{p}_{mt}, \mathbf{s}_{mt}\}_{t=1}^T\}_{m=1}^M$. The maximum likelihood estimation is

$$L(data|\theta) = \underset{\{\theta\}}{max} \prod_{m=1}^M \prod_{t=1}^T f(\mathbf{p}_{mt}, \mathbf{s}_{mt};\theta). \tag{5}$$

Let $\pi(\theta)$ be the prior distribution of θ. Then the posterior distribution of θ is

$$P(\theta \mid data) \propto L(data \mid \theta) \times \pi(\theta).$$

The conventional methods typically nest two inner loops. Each stage of the outer-loop evaluation of the likelihood function for a given parameter vector θ in (5) nests a call to compute the fixed point of the Bellman equations (2) so as to obtain the expected value of delaying adoption. There is also a nested call to compute the demand shocks ξ_t as the fixed point of the BLP contraction mapping (4).

The main difference between the modified algorithm and the conventional algorithm is that during each MCMC step, we do not solve for the fixed point of the Bellman operator (2) nor the fixed point of BLP contraction mapping (4). In contrast, we approximate the demand shocks ξ_t by averaging over a subset of past iterations and we also approximate the expected value of delaying adoption by averaging over a subset of past iteration. Therefore, during each MCMC step k, we iterate the contraction mapping and the Bellman equations only once.

Inner Loop

For a given candidate parameter $\tilde{\theta}^{(k)}$ at MCMC iteration k, the approximation of expected value function is defined as

$$\hat{E}^{(k)}_{\epsilon'}\left[V_i(p', \epsilon'; \tilde{\theta}^{(k)})\right] = \sum_{l=1}^{B(k)} V_i^{(k-l)}(p', \epsilon^{(k-l)}, \tilde{\theta}^{(k-l)}) \frac{K_h(\tilde{\theta}^{(k)} - \tilde{\theta}^{(k-l)})}{\sum_{n=1}^{B(k)} K_h(\theta^{*(k)} - \theta^{*(n-l)})}, \tag{6}$$

where $K_h(\cdot)$ is a multivariate kernel with bandwidth $h > 0$.

Given this approximated expected value function, we update the value functions as

$$\begin{aligned} V_i^{(k)}(p, \epsilon^{(k)}, \tilde{\theta}^{(k)}) &= log\Big\{ \sum_j exp(\tilde{\alpha}_{ij}^{x,(k)} + \tilde{\alpha}_i^{p,(k)} p_j + \xi_j^{(k)}) \\ &+ exp\Big(\beta \sum_{p'} \hat{E}^{(k)}_{\epsilon'}\left[V_i(p', \epsilon'; \tilde{\theta}^{(k)})\right] f(p' \mid p, \tilde{\theta}^{(k)})\Big)\Big\}, \end{aligned} \tag{7}$$

and we store them for future MCMC iterations

For a given candidate parameter $\tilde{\theta}^{(k)}$ at MCMC iteration k, the approximation of ξ is

$$\hat{\xi}^{(k)}(\mathbf{s}, \mathbf{p}, \tilde{\theta}^{(k)}) = \sum_{l=1}^{B(k)} \xi^{(k-l)}(\mathbf{s}, \mathbf{p}, \tilde{\theta}^{(k-l)}) \frac{K_h(\theta^{*(k)} - \theta^{*(k-l)})}{\sum_{n=1}^{B(k)} K_h(\theta - \theta^{*(n-l)})}. \tag{8}$$

Then, we update the ξ and store them for future MCMC iterations:

$$\xi^{(k)}(\mathbf{s}, \mathbf{p}, \tilde{\theta}^{(k)}) = \hat{\xi}^{(k)}(\mathbf{s}, \mathbf{p}, \tilde{\theta}^{(k)}) + log(\mathbf{s}) - log[\mathfrak{s}(\hat{\xi}^{(k)}, V_i^{(k)}, \tilde{\theta}^{(k)}))], \tag{9}$$

where the predicted market share $\mathfrak{s}(\hat{\xi}^{(k)}, V_i^{(k)}, \tilde{\theta}^{(k)})$ is computed by plugging the approximated $\hat{\xi}^{(k)}$

and updated value functions $V_i^{(k)}$ into the equation (3) .

Priors

The priors of the coefficients in the utility equation and the pricing equation are specified as multivariate normals:

$$\begin{aligned} \alpha &\sim MVN(\bar{\alpha}, \Lambda_{\bar{\alpha}}), \\ \eta &\sim MVN(\bar{\eta}, \Lambda_{\bar{\eta}}). \end{aligned}$$

To enforce positive-definiteness, we parameterize the variance-covariance matrix of $\epsilon_t = [\xi_{1,t}, \xi_{2,t}, \varsigma_{1,t}, \varsigma_{2,t}]$ in terms of the log of the diagonal elements of the roots, that is, $\Omega = U'U$ where

$$U = \begin{bmatrix} \sigma_\xi & 0 & \sigma_{\xi\varsigma} & 0 \\ 0 & \sigma_\xi & 0 & 0 \\ \sigma_{\xi\varsigma} & 0 & \sigma_\varsigma & 0 \\ 0 & 0 & 0 & \sigma_\varsigma \end{bmatrix}.$$

The prior distributions of $ln(\sigma_\xi)$, $ln(\sigma_\varsigma)$,and $\sigma_{\xi\varsigma}$ are specified as normal distributions with mean zero and diffuse variance. To guarantee the initial share of consumer types, w_i for $i = 1, ..., I - 1$, is located in the interval $[0, 1]$, we reparameterize $w_i = e^{\varpi_i}/(1 + e^{\varpi_i})$ where the prior of ϖ_i is a normal distribution with mean zero and diffuse variance.