iPhone Trade-in Project

Mingao Xie

March 2023

I build a random coefficient discrete choice model which allows for consumers to trade in their old iPhone model for a new iPhone model. I want to estimate the demand for different iPhone models and estimate the optimal product update cycle on the firm's side.

1 Model

The industry starts at time t = 0. The consumer has an infinite horizon and discounts the future at rate β. The consumer can benefit from at most one iPhone in a period.

Trade-in market assumptions (not necessarily true, but I want to make my model simple):

- Consumer can trade her old iPhone in for credit, and the credit has an upper bound.
- Consumer can only use the trade-in credit to buy iPhone, but not other Apple product.
- No other second-hand market existed.
- If consumer has an iPhone, she can not switch to another brand. Once an iPhone user, always an iPhone user.
- The trade-in value of iPhone depends only on two things: the model of the iPhone, and the market price for that model.

At each time period t, there is a set of models j = i, ..., J_t. Each model has a net flow utility f_{jt} and a disutility from price P_{jt}. The consumer chooses one of the available models or chooses to continue using her current iPhone. If she trades in her current iPhone m and buys model j at time t, then at time t she receives utility of:

$$u_{jt} = f_{jt} - P_{jt} + P_{0.4mt} + \epsilon_{jt} \quad \text{for} \quad j = 1, ..., J_t \tag{1}$$

where ϵ_{jt} is an idiosyncratic type I extreme value term, i.i.d across models and time periods. Let $P_{jt} = \alpha^p \ln(p_{jt})$, where p_{jt} is price of model j in time t. p_{mt}

is the market price of her current iPhone model m at time t if it were new, but since it is a trade in, she can only trade in for 40% the market price, so the utility of trading her old iPhone model m at time t is $P_{0.4m_t}$.

A consumer who does not purchase and trade receives utility of:

$$u_{mt} = f_{mt} + \epsilon_{mt} \tag{2}$$

where flow utility comes from the iPhone m she currently owns plus another idiosyncratic type I extreme value term ϵ_{mt}.

Consider now the consumer dynamic optimization decision. At time t, the consumer is faced with $J_t + 1$ choices and chooses the option that maximizes the sum of the expected discounted values of future utilities conditional on her information at time t. Furthermore, the set of models and prices vary across time, and the market price of existed model decreases in time. Further assume that consumers know that firm will introduce a new model each k period.

Same notation as Gowrisankaran and Rysman (2012), define the state variables and Bellman equation. Let $\vec{\varepsilon_t} \equiv (\varepsilon_{0t}, \ldots, \varepsilon_{J_t t})$ and let $g_{\vec{\varepsilon_t}}$ denote the joint density of $\vec{\varepsilon_t}$. Then, the purchase decision for the consumer depends on $\vec{\varepsilon_t}$, endowment f_{mt}, attributes of currently available models, and expectations about future model attributes. Ω_t denotes the current industry state and includes the number of models J_t, price disutility and mean flow utility for each model. Assume Ω_t evolves according to Markov process that accounts for firm optimizing behavior. The state vector at time t is $(\vec{\varepsilon_t}, f_{mt}, \Omega_t)$.

Bellman equation:

$$\begin{aligned} V\left(f_m, \Omega\right) = \int \max \Big\{ & f_m + \beta E\left[V\left(f_m, \Omega'\right) \mid \Omega\right] + \varepsilon_m, \\ & \max_{j=1,\ldots,J} \left\{ f_j - P_j + P_{0.4m} + \beta E\left[V\left(f_j, \Omega'\right) \mid \Omega\right] + \varepsilon_j \right\} \Big\} g_{\vec{\varepsilon}}(\vec{\varepsilon}) d\vec{\varepsilon}, \end{aligned} \tag{3}$$

where the first term is the value of keeping the current iPhone and the second term is the value of trading in for the optimal replacement today. Note that here $J = J_t$ and $J_t = 1 + \left\lfloor \frac{t}{k} \right\rfloor$, so new model pop up every k periods and consumer knows it.

We need to solve the Bellman but have to face curse of dimensionality. Proceed by using aggregation properties of extreme value distribution to express (3) in a simple form. Rewrite (3) as:

$$V\left(f_m, \Omega\right) = \ln\left[\exp\left(f_m + \beta E\left[V\left(f_m, \Omega'\right) \mid \Omega\right]\right) + \exp(\delta(\Omega))\right] \tag{4}$$

where $\delta(\Omega)$ is logit inclusive value and is defined as:

$$\delta(\Omega) = \ln\left[\sum_{j=1,\ldots,J} \exp\left(f_j - P_j + P_{0.4m} + \beta E\left[V\left(f_j, \Omega'\right) \mid \Omega\right]\right)\right] \tag{5}$$

Define the net augmented utility as ϕ (see next chapter for specific expression). Similarly to Gowrisankaran and Rysman (2012), δ and ϕ together forms consumer's belief and consumer makes decision based on them instead of Ω. Adopting the Inclusive value sufficiency [IVS] in their paper, rewrite (4) as

$$V\left(f_m, \delta, \phi\right) = \ln\left[\exp\left(f_m + \beta E\left[V\left(f_m, \delta', \phi'\right) \mid \delta, \phi\right]\right) + \exp(\delta)\right] \tag{6}$$

and (5) as

$$\delta = \ln\left[\sum_{j=1,\ldots,J} \exp\left(f_j - P_j + P_{0.4m} + \beta E\left[V\left(f_j, \delta'\right) \mid \delta\right]\right)\right] \tag{7}$$

Then, assume δ_t follows a linear autoregressive specification:

$$\delta_{t+1} = \gamma_1 + \gamma_2 \delta_t + v_{t+1} \tag{8}$$

where v_{t+1} is normally distributed with mean zero and unobserved at time t.

2 Aggregation and Demand

Now let us consider the aggregation across consumers and market demand. I will make a very restricted assumption: at the beginning of time t, every consumer in this market has an iPhone, but you do not know which model he holds. Assume there is a fixed amount of infinitely lived consumers who can hold at most one iPhone each period. Assume the population size is M. Consumers differ in mean flow utility f_{ijt}, disutility from price P_{ijt}, idiosyncratic shocks ϵ and inclusive logit value $\delta(\Omega)$.

Define flow utility as a random coefficient model:

$$f_{ijt} = \alpha_i^x x_{jt} + \xi_{jt} \tag{9}$$

where x_{jt} are the observed iPhone characteristics while ξ_{jt} is the common utility from model j that everyone agrees on but not observable. For a given model (say iPhone 7), the observed characteristics are fixed over time and the unobserved characteristic is not fixed and is generally decreasing with time. Since it is random coefficient, $P_{ijt} = \alpha_i^p \ln\left(p_{jt}\right)$, further assume α_i^p and α_i^x are time invariant.

For consumer i holding model m in period t, the conditional probability of trading in and purchasing model j is d_{itj}^m:

$$\frac{\exp\left(\delta_{it}\right)}{\exp\left(\mathcal{V}_i\left(f_{imt}, \delta_{it}\right)\right)} \times \frac{\exp\left(f_{ijt} - P_{ijt} + P_{0.4imt} + \beta E\left[\mathcal{V}_i\left(f_{ijt}, \delta_{i,t+1}\right) \mid f_{ijt}, \delta_{it}\right]\right)}{\exp\left(\delta_{it}\right)} \tag{10}$$

where the first term is the probability of choosing to trade in and the second term is the probability of trading in for model j conditional on trading in.

Define the net augmented utility flow of buying model j as

$$\phi_{ijt} = \alpha_i^x x_{jt} + \xi_{jt} - \left(P_{ijt} - \beta E[P_{ijt+1}]\right) \tag{11}$$

where $P_{ijt} - \beta E[P_{0.4ijt+1}]$ is like the rental price of model j in period t. The net augmented utility flow captures the per-period quality adjusted by the price that consumers take into account when making choices. For example, even though market price of iPhone 7 is low, consumer may still buy iPhone 12 which has a much higher price with the faith that he can trade in his iPhone 12 for more credit in the future and use it to buy iPhone 13, but if he buys iPhone7, the trade-in credit will be pretty low because he expects the market price of iPhone 7 to decrease dramatically in the future. Define the mean net augmented utility flow as:

$$\hat{\phi}_{jt} = x_{jt}\alpha^x + \xi_{jt} \tag{12}$$

Assume ϕ_{ijt} follows a linear autoregressive specification:

$$\phi_{ijt+1} = \kappa_{1i} + \kappa_{2i}\phi_{ijt} + \mu_{t+1} \tag{13}$$

where μ_{t+1} is also normally distributed with mean zero and unobserved at time t. So combine this equation and equation (8), consumer is generally correct about the evolution of quality and the price of products.

And the probability of not trading in but holding the current iPhone model m is $\tilde{d}_{itm}$:

$$1 - \frac{\exp(\delta_{it})}{\exp(\mathcal{V}_i(f_{imt}, \delta_{it}))} \tag{14}$$

Equation (9) can be rewritten as

$$f_{ijt} = \hat{\phi}_{jt} + (\alpha_i^x - \alpha^x)x_{jt} \tag{15}$$

where $(\alpha_i^x - \alpha^x) \sim N(0, \Sigma)$.

Let $\alpha_i = (\alpha_i^x, \alpha_i^p)$ and has mean $\alpha \equiv (\alpha^x, \alpha^p)$ with variance matrix Σ which is a diagonal matrix. For the consumer's discount rate, instead of estimate it, set it to $\beta = 0.99$.

3 Supply

Firm is playing a game. At time 0, firm has to decide the frequency of introducing new iPhone model and once determined, has to make an announcement so all consumers will know and firm has to keep his commitment. At the beginning of each period, firm also has to set the prices for all the iPhone models on the menu. Assume firm has a constant marginal cost $c \geq 0$ which is time invariant. Introducing a new model will incur a fixed cost $\phi > 0$, which is also time invariant. So at time t, the number of models available for consumers is $J_t = 1 + \lfloor \frac{t}{k} \rfloor$.

Let s_{jt}^n denote the market share of consumers that purchase j in period t and s_{jt}^o be the proportion of consumers holding model j which they bought sometime ago in period t. So the total proportion of consumers having model j at the end of period t is

$$s_{jt} = s_{jt}^n + s_{jt}^o \tag{16}$$

where

$$s_{jt}^{n} = \int_{vi} \sum_{m \neq j} d_{itj}^{m} s_{m,t-1} dP_v \tag{17}$$

and

$$s_{jt}^{o} = \int_{vi} \tilde{d}_{itj} s_{j,t-1} dP_v \tag{18}$$

The state space of the firm in period t is the menu of current available model $\vec{j}_t \equiv (j_0, \ldots, j_{J_t})$and the market share of each model $\vec{s}_t \equiv (s_0, \ldots, s_{J_t})$. And the market demand for model j is Ms_{jt}.

Let $\beta = 0.99$ be the discount rate for the firm and firm's profit-maximizing problem is:

$$V(\vec{j}_t, \vec{s}_t) = \max_{\{p_{jt}\}} \sum_{j=1}^{j=J_t} [(p_{jt} - c)Q_{jt} - 0.4 \int_v \sum_{m \neq j} d_{itj}^{m} s_{m,t-1} p_{itm} dv] - d_t \phi + \beta V(\vec{j}_{t+1}, \vec{s}_{t+1}) \tag{19}$$

where $d_t = 1$ if t is divisible by k, meaning it is time to make the commitment and introduce a new product, and $d_t = 0$ otherwise.

On the RHS of equation (16), the first term is the money firm gets for selling model j; because people trade in for iPhone, the second term is the money lost from trading-in program; the third term is the fixed cost from introducing a new model if it is time to do so by commitment; finally the last term is the discounted value from future.

But then the model is too complicated, for now let us assume firm can not set prices, instead, price is set by a raccoon and in each period, this raccoon will adjust the prices of all iPhone models and generally, he will lower the price of old iPhone models. Hence, the only decision firm has to make is the product update cycle k which should be announced at the beginning of period 0. So the profit-maximizing problem becomes:

$$V_t(k) = \max_{k} \sum_{j=1}^{j=J_t} [(p_{jt} - c)Q_{jt} - 0.4 \int_v \sum_{m \neq j} d_{itj}^{m} s_{m,t-1} p_{itm} M dv] - d_t \phi + \beta V_{t+1}(k) \tag{20}$$

4 Solution Strategy

4.1 Fake/ideal data

Suppose in each period, I can observe those things:

- population M, which is time fixed.
- each model's characteristics (weight, size, battery life, camera resolution)
- market price of all the existing iPhone models (exogenous)

- update circle chosen by Apple: k
- market share of people holding each model
- each individual's purchasing history (this is a strong requirement for data)

4.2 Identification

4.2.1 Intuition

In fact, firm is increasing competition by introducing new model and consumer is making substitution across time periods. Consumer heterogeneity Σ is captured by consumer substitute between similar products (iPhone model with similar characteristics). Change in market share of model associated with a change in model characteristics captures the distribution of α.

4.2.2 Demand side

Further assume the unobserved common utility for model j ξ_{jt} evolves according to AR(1):

$$\xi_{jt+1} = -\lambda\xi_{jt} + \zeta_{jt} \tag{21}$$

There is a negative sign in front of λ because the idea is consumer's unobserved preference towards model j in decreasing in time, which is intuitive because people tend to lose interest in obsolete product. λ is also a parameter to be estimated. ζ_{jt} is the error term.

More assumptions! Consumers have different price sensitivity and assume $\alpha_i^p = \frac{\alpha^p}{y_i}$, where y_i is individual's income which is lognormal distributed and can be estimated through maybe census data. But there is selection bias in using census data because iPhone users may be generally richer than other mobile phone users, but since in my model Apple is the monopolist let us ignore this issue for now. For consumer's heterogeneity preferences in other iPhone characteristics, I assume α_i^x is independently normally distributed with mean α^x and standard deviation σ_{α^x}.

So on the consumer side, the parameters I have to recover are $\theta = \{\alpha^p, \alpha^x, \sigma_{\alpha^x}, \lambda\}$. GMM criterion function:

$$G(\theta) = z'\overrightarrow{\xi}(\theta) \tag{22}$$

Let nonlinear parameters be $\theta_1 = \{\alpha^p, \lambda, \sigma_{\alpha^x}\}$ and linear parameters be $\theta_2 = \{\alpha^x\}$. GMM estimator:

$$\hat{\theta} = arg \min_{\theta} G(\theta)'W^{-1}G(\theta) \tag{23}$$

where $G(\theta)$ is vector of stacked moments and W is weighting matrix which depends on consistent estimates of unknowns.

Computation

1. Guess θ_1

2. Get random draws of consumers and hold it fixed during estimation
3. Given θ_1 and random draws, via contraction mapping, find the mean net augmented utility $\hat{\phi}_{jt}$
4. Recover ξ_{jt} from the final value of $\hat{\phi}_{jt}$ via linear regression by equation (12)
5. Compute objective function (23)
6. Check convergence of GMM, return to step 1
7. θ_2 can be expressed by θ_1, so once pin down θ_1, θ_2 is solved.

How to proceed step 3 though?

1. In each iteration, for each value of $\hat{\phi}_{jt}$ and with the given θ_1, calculate the logit inclusive value δ_{it} using equation (7)[1]; calculate the net augmented utility ϕ_{ijt} using equation (11)[2].
2. Use the δ_{it} and ϕ_{ijt} in the previous step to estimate the coefficients in the Markov process in equation (8) and equation (13).
3. Use the estimated coefficients got in step 2 to construct transition matrix.
4. Use the transition matrix to calculate Bellman in equation (6)
5. Compute individual choice probability d^m_{itj} and $\tilde{d}_{itm}$ by equation (10) and (14)
6. Aggregate individual using equations (16), (17) and (18) to get the predicted market share.
7. Use the predicted market shares to update $\hat{\phi}_{jt}$.

   I follow Schiraldi's (2011) approach, inverting the market share to recover ϕ_{jt}:
$$\hat{\phi_{jt}}' = \hat{\phi}_{jt} + \psi(ln(\check{s}^o_{jt}) - ln(\tilde{s}^o_{jt}(\hat{\phi}_{jt}, \theta))) \quad (24)$$
   (a) Note that $ln(\check{s}^o_{jt})$ is the observed market share and is from data.
   (b) $\tilde{s}^o_{jt}(\hat{\phi}_{jt}, \theta)$ is the predicted market share from step 6
   (c) ψ is a tuning parameter and Schiraldi sets it to $1 - \beta$
8. Keep iteration until $\hat{\phi}_{jt}$ converges.

[1] It should add subscript "it" because (7) is at the individual time period now. Also in this step we need to use equation (15)

[2] Remember $\phi ijt = \hat{\phi}_{jt} - (P_{ijt} - \beta E[P_{ijt+1}])$

4.2.3 IV

ξ_{jt} could be correlated with p_{jt} and cause endogeneity.

4.2.4 Supply side

On the firm side, the only decision is the update circle k. Suppose firm knows the consumer's optimization strategy and everything when setting k, then optimal k^* solves:

$$k^* = arg \max_k V_0(k) \tag{25}$$

where $V_0(k)$ is defined in equation (20).

References

Gowrisankaran, Gautam, and Marc Rysman. "Dynamics of consumer demand for new durable goods." Journal of political Economy 120, no. 6 (2012): 1173-1219.

Copeland, Adam, and Adam Hale Shapiro. "Price setting and rapid technology adoption: The case of the PC industry." Review of Economics and Statistics 98, no. 3 (2016): 601-616.

Schiraldi, Pasquale. "Automobile replacement: a dynamic structural approach." The RAND journal of economics 42, no. 2 (2011): 266-291.