Fire Fox project Theory

Summary = short

Quick math   organized

① Modeling Consumer learning from online product reviews (Zi Zhao Georgia State, Sha Yang USC, ...)

ⓐ Structured model (experiential products) - limited repeat

ⓑ Significant product category experiment (e.g. same genre) learn from experience, credibility modeling (Panel dataset)

ⓒ result: ① learn more from others than their experience in the same genre

ⓓ ② Profit impact of product reviews (and how varies with # reviews): diminishing return: # reviews

ⓔ additional reviews may lower profit (Counter Factual)

ⓕ estimate optimum # of reviews for representative product

ⓖ review credibility: degree of reflect consumer own experience (eval.) ≡ percision (credible info) ~ larger effect on choice probability

ⓗ Bayesian learning { ① product quality ② review credibility

ⓘ fake reviews increase consumer uncertainty (lower effect of negative reviews)

model

$$\underbrace{A_{ij}}_{\text{Category quality}} | \underbrace{I_{i,t-2}}_{\text{info set}} \sim N(E_{i,t-2}(A_{ij}), \sigma^2_{v_{ij},t-2})$$

→ Eval. after consumption; Perception variance; → uncertainty

$$\underbrace{A_{Eikj,t-1}}_{\text{evaluation after consumption}} \sim N(A_{ij}, \sigma^2)$$

→ Consumer belief of category e.g. Horror books

Consumer updating Expected Qual of Categ

$$\begin{cases} A_{Eikj,t-1} \in (-\infty, 1.5] & \text{if } R_{Eikj,t-1} = 1 \\ A_{Eikj,t-1} \in (1.5, 2.5] & \text{if } R_{Eikj,t-1} = 2 \\ A_{Eikj,t-1} \in (2.5, 3.5] & \text{if } R_{Eikj,t-1} = 3 \\ A_{Eikj,t-1} \in (3.5, 4.5] & \text{if } R_{Eikj,t-1} = 4 \\ A_{Eikj,t-1} \in (4.5, +\infty) & \text{if } R_{Eikj,t-1} = 5 \end{cases}$$

$R_{Eikj,t-1}$ → observed evaluation/rating of the purchased bk

k_j = product item

use k_j item @ t-1: $A_{ij} | I_{i,t-1} \sim N(E_{i,t-1}(A_{ij}), \sigma^2_{vij(t-1)})$ ← posterior belief

$$E_{i,t-1}(A_{ij}) = E_{i,t-2}(A_{ij}) + P_{ikj,t-1} B_{ij,t-1}(A_{Eikj,t-1} - E_{i,t-2}(A_{ij}))$$

Dummy of consumption ← ; → relating weight of Experience signal (greater for greater ...)

$$\frac{1}{\sigma^2_{vij,t-1}} = \frac{P_{ikj,t-1}}{\sigma^2} + \frac{1}{\sigma^2_{vij,t-1}} \qquad B_{ij,t-1} = \frac{\sigma^2_{vij,t-2}}{\sigma^2_{vij,t-2} + \sigma^2}$$

$A_{ij} | I_{i1} \sim N(E_{i0}(A_{ij}), \sigma^2_{vij0})$: prior

Rational Expectation: $E_{i0}(A_{ij}) = A_j$ $\sigma^2_{vij0} = \sigma_A^2 \Rightarrow A_{ij} \sim N(A_j, \sigma_A^2)$ ← true quality

$MR_{kj} - A_{Eikj,t-1} \sim N(0, \sigma_R^2)$ For credibility of info

→ population mean of evaluation (item k_j): non symmetric deviation

⊛ Credibility is unobserved ≡ higher percision

$R_{ikj,t-1} - MR_{kj} \sim N(0, \tilde\sigma_R^2)$ ⊛ each review unbiased signal of pop. dist. of evals

→ eval of specific review   → variance of deviation (representativeness of mean)

- since indiv. both cons. & rev: $\tilde\sigma_R^2 = \sigma_R$

$$\underbrace{\bar R_{kj,t-1}}_{\text{AVG eval of all reviews}} - \underbrace{A_{Eikj,t-1}}_{\text{Eval Cons i}} \sim N\left(0, \left(1 + \frac{1}{n_{kj,t-1}}\right)\sigma_R^2\right), \quad \frac{1}{\sigma_R^2} \text{ Percision}$$

$n_{kj,t-1}$: # reviews product k_j at or before t-1

- more distance in evaluation ≡ less credible

evolution of credibility of reviews for consumer i over time:

$\frac{1}{\sigma_R^2} | I_{i,t-2} \sim \Gamma(\alpha_{i,t-2}, \beta_{i,t-2})$ prior belief σ_R

$\frac{1}{\sigma_R^2} | I_{i,t-1} \sim \Gamma(\alpha_{i,t-1}, \beta_{i,t-1})$ posterior (after reviewe $n_{kj,t-1}$ eval, mean eval signal $\bar R_{kj}$)

$$\alpha_{i,t-1} = \alpha_{i,t-2} + \frac{P_{ikj,t-1}}{2}$$

$$\beta_{i,t-1} = \beta_{i,t-2} + P_{ikj,t-1}\frac{(\bar R_{kj,t-1} - A_{Eikj,t-1})^2}{2\left(1 + \frac{1}{n_{kj,t-1}}\right)}$$

Integration of Consumer's own Exper & Reviews

2 src of info { own ① Prod rev. ②   Purchase decision on k_j^* $N(E_{i,t-1}(A_{ij}), \sigma^2 ...)$

① evaluation of cons. i: $A_{Eikj^*t} \sim N(A_{ij}, \sigma^2)$ Prior cat. belief: $A_{ij} | I_{i,t-1} \sim N$

$$\Rightarrow A_{Eikj^*t} \sim N(E_{i,t-1}(A_{ij}), \sigma^2 + \sigma^2_{vij,t-1}) \quad \text{(I)}$$

② src: eval by revs (others) → form Expect. of Cred of revs

$$\frac{1}{\sigma_R} \sim \Gamma(\alpha_{i,t-1}, \beta_{i,t-1}) \Rightarrow E_{i,t-1}(\sigma_R^2) = \frac{\beta_{i,t-1}}{\alpha_{i,t-1} - 1} \quad ⊛$$

mean eval k_j^* by revs ⟵ $\Rightarrow \bar R_{kj,t} \sim N\left(A_{Eikj^*t}, \left(1 + \frac{1}{n_{kj^*,t-1}}\right)\frac{\beta_{i,t-1}}{\alpha_{i,t-1} - 1}\right)$ (II)

⇒ Expected qual item k_j^* cons i time t Dist Norm:

$$E_{it}(A_{Eikj^*t}) = W_{ikj^*,t-1}E_{i,t-1}(A_{ij}) + (1 - W_{ikj^*,t-1})\bar R_{kj^*t}$$

$$Var_{it}(A_{Eikj^*t}) = \frac{1}{\left(\frac{1}{\sigma^2 + \sigma^2_{vij,t-1}} + \frac{\alpha_{i,t-1} - 1}{\beta_{i,t-1}}\right) / \left(1 + \frac{1}{n_{k^*j,t-1}}\right)}$$

$$W_{ikj^*,t-1} = \frac{(\sigma^2 + \sigma^2_{vij,t-1})^{-1}}{\frac{1}{\sigma^2 + \sigma^2_{vij,t-1}} + \frac{\alpha_{i,t-1} - 1}{\beta_{i,t-1}}\left(1 + \frac{1}{1 + n_{kj^*,t-1}}\right)^{-1}}$$

⊛ ⇒ $\beta_{i0} = (\alpha_{i0} - 1)E_{i0}(\sigma_R^2)$ → magnitude of richness of Cons. i's initial expr. (inexp ⇒ ↓α_{i0})

initial uncertainty of revs

$E\alpha_{i0}, E_{i0}(\sigma_R^2) > 0$, estim ⇒ $\log(\alpha_{i0}) \sim N(M_{\alpha 0}, V^2_{\alpha 0})$
$\log(E_{i0}(\sigma_R^2)) \sim N(M_{R0}, N^2_{R0})$

hetrogen. in initial cred. rev.

modeling consumer decisions

Own knowledge → is about product cat.

utility: $V_{ikj^*,t} = \exp(-\hat r_i(\omega_i A_{Eikj^*t} + X_{ikj^*t}\gamma_i + \varepsilon_{ikj^*t}))$

Exterem val Error; $\hat r$ risk aversion; ω quality weight; X Covariate vector e.g. price; γ coeff

not know product quality prior purchase ⇒ decide based on $E_{it}(V_{ikj^*t}) = -E_{it}[e^{-\hat r_i \omega_i A_{Eikj^*t}} e^{\hat r_i (X_{ikj^*t}\gamma_i + \varepsilon_{ikj^*t})}]$

MGF:

$$E_{it}(V_{ikj^*t}) = -e^{-\hat r_i\{[\omega_i E_{it}(A_{Eikj^*t})] - \frac{1}{2}\hat r_i(\omega_i)^2[Var_{it}(A_{Eikj^*t})] + X_{ikj^*t}\gamma_i + \varepsilon_{ikj^*t}\}}$$

→ depend on 2 src ←

maximize expected utility: $U_{ikj^*t} = \hat U_{kj^*t} + \varepsilon_{kj^*t} = r_i\omega_i[E_{it}(A_{Eikj^*t})] - r_i[Var_{it}(A_{Eikj^*t})] + X_{ikj^*t}\gamma_i + \varepsilon_{ikj^*t}$, $r_i = \frac{1}{2}\hat r_i(\omega_i)^2$

unobserved hetrogeneity:
$$\begin{cases} \log(\omega_i) \sim N(\bar\omega, \sigma_\omega^2) \\ \log(r_i) \sim N(\bar r, \sigma_r^2) \\ A_{ij} \sim N(A_j, \sigma_A^2) \rightarrow \text{diagonal} \\ \gamma_i \sim MVN(\bar\gamma, \Sigma) \end{cases}$$

Firefox project theory

Seq Srch

① $U_{it} = \alpha_j + \beta_i S_{jt}$

specific need — stars — risk aversion

$E\{U_{iK_j^* t}\} = - \hat{r}_i \{[w_i E_{it}(A_{EiK_j^* t})]$

$- \frac{1}{2}\hat{r}_i (w_i)^2 [Var_{it}(A_{EiK_j^* t})] - r_i [Var_{it}(A_{EiK_j^* t})]$

$+ X_{iK_j^* t} \gamma_i + \varepsilon_{iK_j^* t}$

$$w_{iK_j^*, t-1} = \frac{(\sigma^2 + \sigma^2_{vij, t-1})^{-1}}{\frac{1}{\sigma^2 + \sigma^2_{vij,t-1}} + \frac{a_{i,t-1}^{-1}}{\beta_{i,t-1}} \left(1 + \frac{1}{1 + n_{K_j^*, t-1}}\right)^{-1}}$$

$E_{it}(A_{EiK_j^* t}) = w_{iK_j, t-1} E_{i,t-1}(A_{ij}) + (1 - w_{iK_j^*, t-1}) \bar{R}_{K_j^* t}$

AVG eval of all reviews

$\bar{R}_{K_j, t-1} - A_{EiK_j, t-1} \sim N(0, (1 + \frac{1}{n_{K_j, t-1}}) \sigma_R^2)$

$\Rightarrow \frac{1}{\sigma_R^2} | I_{i,t-1} \sim \Gamma(a_{i,t-1}, \beta_{i,t-1})$

$\underbrace{A_{ij}}_{\text{Cat quality}} | \underbrace{I_{i,t-1}}_{\text{info set}} \sim N(E_{i,t-2}(A_{ij}), \sigma^2_{vij,t-2})$

$\underbrace{A_{EiK_j, t-1}}_{\text{eval of the consump}} \sim N(A_{ij}, \sigma^2)$

↳ Cons. belief of Cates

Relative weight Expr signal

$E_{i,t-1}(A_{ij}) = E_{i,t-2}(A_{ij}) + D_{iK_j, t-1} B_{ij,t-1} (A_{EiK_j, t-1} - E_{i,t-2}(A_{ij}))$

$$B_{ij,t-1} = \frac{\sigma^2_{vij,t-2}}{\sigma^2_{vij,t-2} + \sigma^2}$$

$$\frac{1}{\sigma^2_{vij}} = \frac{D_{iK_j,t-1}}{\sigma^2} + \frac{1}{\sigma^2_{vi,j,t-1}}$$

$\underbrace{MR_{K_j}}_{\text{pop mean of eval item } K_j} - A_{EiK_j, t-1} \sim N(0, \sigma_R^2)$

$\underbrace{R_{iK_j, t-1}}_{\text{eval spefic Rev}} - MR_{K_j} \sim N(0, \tilde{\sigma}_R^2)$

{① prod qual ② review cred

Correl util item within cat: ρ

$$\Rightarrow P(D_{iK_j^*} = 1 \mid D_{ijt} = 1) = \frac{\exp(\hat{u}_{iK_j t})}{\sum_{K'} \exp(\hat{u}_{iK'_j t}/1-\rho)}$$

$$P(D_{ijt} = 1) = \frac{[\sum_{K'} \exp(\frac{u_{iK'_j t}}{1-\rho})]}{\sum_{j'} [\sum_{K'} \exp(\hat{u}_{iK'_j t}/1-\rho)]^{1-\rho}}$$

Covars {① price ② time elapse since public.

Results: {
① profit impact of prod Rev. : diminishing return
② learn more from others than own Exp (cmps)
③ estimate opt # review
④ Fake rev increase cons. uncert.

$\boxed{\text{Firefox project thery}}$ $\boxed{\text{Summary = Short}}$ $\boxed{\text{Quick math}}$ $\boxed{\text{organized}}$

- utility of other good cat j: (O_j) : $U_{io_jt} = \underbrace{O_j}_{\text{Cat. spec. intercept othr gds}} + \varepsilon_{io_jt}$

- Correl. util. item's within categ (nested logit): ρ

$$P(D_{ik_j^*t} = 1 \mid D_{ijt} = 1) = \frac{\exp\left(\frac{\hat{U}_{ik_jt}}{1-\rho}\right)}{\Sigma_{k'} \exp\left(\frac{\hat{U}_{ik_jt}}{1-\rho}\right)}$$

unconditional prob. purchase Cat. j:

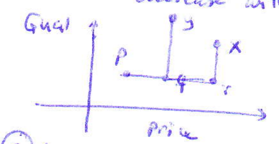
$$P(D_{ijt} = 1) = \frac{\left[\Sigma_{k'} \exp\left(\frac{\hat{U}_{ik_jt}}{1-\rho}\right)\right]^{1-\rho}}{\Sigma_{j'} \left[\Sigma_{k'} \exp(U_{ik'_jt}/1-\rho)\right]^{1-\rho}}$$

Covars { (1) price
(2) time elapse since publication

FFP Prospect theory, Loss Aversion & Ref Dep effect Brand choice

② Bruce Hardie, Eric Johnson, Fader 1993

① multi attrib. gen. of prospect theory value func
② position of brand relative to multiattrib. ref. point
③ Cons. weight loss from ref point more than equiv sized gain (loss aversion)
④ multinom. logit of ref dependent choice model, Calib scanner data
⑤ Signific. loss aversion in coeff (assymetric responses to changes in prod. choice)
⑥ Trav. Kahneman theor. framework multi attrib basis: (Price & Qual...)
(1) each choice altrntv decompose into set of valuesom attrib.
(2) each attrib. can be desc. by its own val. func. with own value on attrib.
(3) altrntv eval. relativ to ref points
⑦ three charc. of val func:
(1) Ref depend: carrier of an attrib val. not on absolute but rather deviation from ref. level ⇒ (gain or loss)
(2) Loss aversion: value func. steeper for loss than for gain; loss decr. value more than an equiv gain will incr val
(3) Diminishing sensitivity: marg. val. of both gains & losses decrease with size (first dollar hurts the most)



$x \succ_r y$ $x \approx_q y$ Diminish sens: $y \succ_p x$

ref func assoc with two attrib
⑧ decomposability: e.g. additive $U_r(x_1, x_2) = R_1(x_1) + R_2(x_2)$
⑨ Constant loss Aversion: $R_i(x_i) = \begin{cases} u_i(x_i) - u_i(r_i) & \text{if } x_i \ge r_i \\ \lambda_i[u_i(x_i) - u_i(r_i)] & \text{if } x_i < r_i \end{cases}$
Coeff of loss Aversion
⑩ reference brand: most recent brand purchased by each household (avail. in memory, ② status quo)
⑪ loss aversion:
$v_{hjt} = \beta_1[Qualgain_{hjrt} + \lambda_q Qualloss_{hjrt}] + \cancel{\beta_2[Pricegain_{hjrt}}$
$+ \beta_2[Pricegain_{hjrt} + \lambda_p Priceloss_{hjrt}] + \beta_3 Feature_{hjt}$
$+ \beta_4 loyalty_{hjt}$
household; ret

$Qualgain_{hjrt} = Q_j - Q_r$ $Qualloss_{hjrt} = 0$ $Q_j \ge Q_r$
$Qualgain_{hjrt} = 0$ $Qualloss_{hjrt} = Q_j - Q_r$ $Q_j < Q_r$
$Pricegain_{hjrt} = P_{hrt} - P_{hjt}$ $Priceloss_{hjrt} = 0$ $P_{hjt} \le P_{hrt}$
$Pricegain_{hjrt} = 0$ $Priceloss_{hjrt} = P_{hrt} - P_{hjt}$ $P_{hjt} > P_{hrt}$

⑫ tempered ref price model (smooth of past observed)
$Price_{hj(t+1)} = \gamma\, price_{hjt} + (1-\gamma) P_{hjt}$

When is versioning optimal for info goods | Bhargava, Choudhary 2008

① when versioning optimal strategy for info goods
② info good: var. costs invar. with qual.
③ monopoly firm exist prod in mkt, opport. to seg mkt by introd. additional lower qual vers. (single prod. max prof.)
④ optimal when mkt share of lower-qual vers. (offered alone) > mkt share of high qual vers. alone
⑤ Incentive Compatibility
⑥ var cost ↗ ⇒ Consider only offering one version
⑦ var cost ↘ ⇒ Explore adding lower qual. version

(1) Feasible quality: $\max_{P_1,\dots,P_H} \pi = (P_H - c)(1 - F(v_H)) + \sum_{j=1}^{H-1} (P_j - c)(F(v_{j+1}) - F(v_j))$ ②
→ Cumul. dist. of types
s.t. $U(v_j, q_j) - P_j = U(v_j, q_{j-1}) - P_{j-1}$ ($\forall j = 1, \dots, H$) → qual level, price
$v_j \ge v_{j-1} \ge 0$ ↳ valuation of type v_j
v: index cons. types (arranged increasing order)
$U_v(v,q) > 0$
$c(q) = c \ \forall q$ q_j low qual ⇒ q_1 $q_0 = 0$ $U(v, q_0) = P_0 = 0$
own demand elast:
$\varepsilon(v, q_j) = \frac{\%\Delta q_j}{\%\Delta P_j} = \frac{U(v,q_j)}{U_v(v,q_j)} \left(\frac{F'(v)}{1 - F(v)}\right)$
~~Cross~~ elast. (elast q_j, presence of lower qual q_{j-1})
$\eta(v; q_{j-1}, q_j) = \frac{\%\Delta q_j}{\%\Delta(P_j P_{j-1})} = \frac{U(v,q_j) - U(v,q_{j-1})}{U_v(v,q_j) - U_v(v,q_{j-1})}$
versioning optimal when $\exists q_i: (1 - F(\hat{v}_i)) > (1 - F(\hat{v}_H))$
≠ versioning not optimal linear valuation $U(v,q) = v \cdot q$
⑧ Heterogeneity makes versioning optimal (in marginal value)
revenue loss by Cannibalization & info rent
$c=0$, versioning optimal when $\exists q_j \in (0, q_H)$ $U(v,q_j)/U(v,q_H) \searrow v$
$c>1$, optimal when: $\exists q_j \in (0, q_H)$ rel. surplus: $\frac{U(v,q_j) - c}{U(v,q_H) - c} \searrow v$
How? disable features

How Does The Variance of Prod Rating matter | Monic Sun 2012

① info role of prod rating. How rat. help cons. fig. out how much enjoy
② high Avg rating ≡ high prod. qual.
③ high var rating ≡ niche prod ⇒ higher subsey demand ⇔ AVG rating low
④ monopoly, risk neutral cons. hetrogen. tastes; seller's prod. two attrib: (1) qual. (2) mismatch cost (Category how niche, small group)
⑤ taste space: line of size 2, product mid point
cons. taste dist. x from prod. buys at price P
prod quality: $v - t \cdot x - P$ = utility, high t product: niche product
mismatch cost; low t prod: mainstream, all cons same utility
⑥ beginning of game neither seller nor consumer know realization of v and t (exogen. given ≡ 95% fail)
$f(v,t)$ common knowledge
⑦ joint prob dist. (no info on qual or mismatch cost)
⑧ period 1: unit mass of early cons. $x \sim U(0,1)$; seller chooses price: P_1
Cons. decides to buy & quickly consum & publish rating $s(x) = v - tx$
Period 2: unit mass of late cons. $x \sim U(0,1)$; observe 1st period demand, AVG rating, Var ratings. Seller chooses P_2, Cons. decides whether to buy the product
⑨ Subgame equilibrium: first period no info ⇒ use Expectation $E(v), E(t)$, (the $f(v,t)$) common knowledge
indiff cons.: $E(v) - E(t) \cdot D_1 - P_1 = 0$ ⇒ $x \in [0, D_1]$ purchase, first period demand D_1
distance D_1
⇒ rating uniformly dist.: $[v - t \cdot D_1, v]$
AVG rating: $M = v - \frac{t \cdot D_1}{2}$ Var rating: $V = \frac{1}{12}(t \cdot D_1)^2$
V, M not indep. ≡ high AVG rating ⇒ $V = 0$ ⇒ $v = M + \sqrt{3V}$ $t = \frac{2\sqrt{3V}}{D_1}$
low AVG ⇒ { ① low quality or ② high mismatch cost

MKT sci | UGC

Users generated Content { ① product review ② Description of product usage ③ home made advertising blogs

Q { ① How? ② why? ③ methods? ④ impact?

Challanges { ① processing ② qualitative ③ life time of UGC

New performance metric: (Just means toward an end)

{ ① likes ② followers ③ views

① individual's decision to Contribute a product rating
① whether to Cont.?
② what to Cont.?

② wikipedia
③ online sponsored UGC platform
④ effort to increase UGC
⑤ UGC network effect (impact)
⑥ Financial implication of investments in the retention and acquisition of Content Contributers and Consumers (AR Problem) - Network dynamic
⑦ Sequential and temporal dynamics of online opinion Amazon.com
why rating get worse over time?
⑧ UGC Contribution in NPD of open source ① Social Capital
② time to product release) B2C ≡ B2B
⑨ Designing ranking systems for Hotel and travel search engines
mining Cont of UGC or Crowd src
⑩ Text mining ⇒ Competitive market structure (develop map)

① online product opinions: Incidence, Evaluation, and Evolution
Wendy Moe, David Schweidel 2012

- individual decision to Contribute (provide product rating)
- Previously posted behavior affect < whether: incidence (Selection effects) / what: evaluation (adjustment effects)
- negative environment ⇒ discourages posting (posit → Encorgs)
{ - less frequent posters → Band wagon behavior
- active Customers → differentiation behavior

- how evolution shapes? will determine
① V_{ij}: indiv. Postpurchase eval. for product j (outcome of two decisions) (latent constructs)
② condition on time
③ latent measure of Experience ($experience_j$): product specific

Evaluation model — post product rating: Ordered Probit

Postpurchase eval $V_{ij} = \gamma_{i0} + k_j \sim N(0, \sigma_k^2)$ → Variation across products (prod qual)
↓ different level of baseline positivity
→ Cov desc. rating env

Respondent adjustment → $V^*_{ijk} = V_{ij} + \gamma_{i,1:N} X_{j,k(i)}$ [post eval diff from postp eval]
→ rating session

$Pr(y_{ijk} = r \mid Z_{ijk} = 1) = Pr(\mu_{i,r-1} < V^*_{ijk} + \epsilon_{ijk} < \mu_{i,r})$ $r = 1, 2, \ldots, S$
→ rating is Contributed → idiosyncratic

Incidence model — conditioning:
① experience the product ③ curr stat. ratings env.
② vary in baseline tendency ④ post purchase eval.

$Pr(Z_{ijk}=1) = pr(experience_j = 1)\, Pr(Z_{ijk}=1 \mid experience_j = 1)$ (Contribute)

① Selection effect of rating env. ② not having Experienced the product like market penetration

- Conditioned on having Experience prod j, indiv i submits rating if

$\beta_{i0} + \beta_{i,1:N} X_{j,k(i)} + \beta_{i,N+1} GT_{j,k(i)} + \delta_1 V_{ij} + \delta_2 V_{ij}^2 + \omega_{ijk} > 0$ ← idiosync

$\Rightarrow pr(Z_{ijk}=1 \mid exp_j = 1) = \Phi(\beta_{i0} + \beta_{i,1:N} X_{j,k(i)} + \beta_{i,N+1} GT_{j,k(i)} + \delta_1 V_{ij} + \delta V_{ij}^2)$

baseline Prop. to submit prod rating; Selection effect env.; Google trend Proxy of Category level interest; Extreme opinion more likely to express opn

$\Rightarrow pr(Z_{ijk}=1) = \Phi(v_j)\, \Phi(\beta_{i0} + \beta_{i,1:N} X_{j,k(i)} + \beta_{i,N+1} GT_{j,k(i)} + \delta_1 V_{ij} + \delta_2 V_{ij}^2)$

→ Prob. avg indiv has Exp. with product j

$v_j = \theta + \tau_i \sim N(0, \sigma_\tau^2)$

② Network effects val. Collab. UGC Ransbotham, Kane 2012

① Value of UGC func { ① direct effect of Cont. ② embeddedness in Contrib/Cont. ntwrk
(inverted U)
② wikipedia: # Contrib ↗ viewership { locally central / Globally Central: shortest path
③ two mod social ntw and , hetrogenity
④ network effect stronger for new UGC
⑤ Closeness Centrality $C_c(v_a) = \frac{(|V_a| - 1)}{\sum_{v_b \in V_a \setminus v_a} d(v_a, v_b)}$
⑥ General latent mixed models:
⑦ $\ln(v_{i,t}) = (\beta_0 + \zeta_{0,i}) + \beta_c C_{i,t} + \beta_n N_{i,t-1} + \varepsilon_{i,t}$
article view; monthly Control; monthly total network cov
⑦ hetrogenity: $\ln(v_{i,t}) = (\beta_0 + \zeta_{0,i}) + \beta_c C_{i,t} + (\beta_n + \zeta_{n,i}) N_{i,t} + \varepsilon_{i,t}$

③ Albuquerque, pavlidis 2012
Eval prom. activ. in online two sided MKT UGC

① indiv. choice visiting the platform, creating, purchasing Content = f(indiv char, MKT activities)
② allow interdep. dec within & across users, { Agg: Brand qual., H.puc mag. (indiv)
③ Compare referral & Firm based actions: Price, Prom, Public Rel
↳ affect all cons. dec platform ↳ purchase dec impact
④ ⇒ motivate free prom activity by content creator (invest)
⑤ Cons decision stages: ① whether visit the platform ② produce Content? purchase content? (3 decisions) ⇒ Backward induction
⑥ utility by: ① personal pref ② expectation about actions' utility
⑦ Cons. choices: ① Browsing (no purchase/creation) ② purchase Content ③ creating Content ④ both creat. & purchas.
⑧ percieved quality under the impression of firm's signaling
⑨ used AVG Price, reason of creation { ① mony ② gift Economy
⑩ assump. user's correct expect. of outcome of creation
⑪ not intertemporal dec: ① Content perish. ② low monetary benefit (diminish postpone)

⑫ Cons. dec purchase $d_{1it} = 1$ if $u_{1it} > 0$ $d_{1it} = 0$ otherwise
Cons. dec creation $d_{2it} = 1$ if $u_{2it} > 0$ $d_{2it} = 0$ otherwise

⑬ util of purchasing: $u_{1it} = \varsigma_{1i} Z_{it} + \alpha_{1i} X_{1t} - \beta_{1i} P_t + \delta_i d_{2it} + \lambda D_{1t} + \varepsilon_{1it}$ (Price)

util of creation: $u_{2it} = \varsigma_{2i} Z_{2it} + \alpha_{2i} X_{2t} + \lambda_{2i} E(D_{1t+\tau}) + \varepsilon_{2it}$

indiv specific or Cons see specific charac.

appear var. platform e.g. prom. activ. of owner & Content. creator (daily issues, sales of offline

Kim 2010

①

Online Demand under Limited Consumer Search

Kim 2010 Albuquerque Bronnenberg

① amazon.com, durable

② optimal seq srch proc. (Aggreg. indiv. opt srch seq)

③ dynamic programming + choice model (Weitzman); hetrogen.

④ answers ques (1) mkt struc (2) competition (3) policy maker issue (lower search cost)

⑤ own & cross price elast., actual product search rank predict

⑥ Benefit product recomm. via lower search cost

⑦ $u_{ij} = V_{ij} + e_{ij}$ $\quad V_{ij} = x_j b_i$ $\quad b_i \sim N(b, B)$ $\quad e_{ij} \sim N(0, \sigma_{ij}^2)$

outside good ~~j~~ alternative (not req search: Cons aware) (Zero determin. part) — Cons. knows (general cat. pg, not detail pg) — diagonal — uncertainty — costly resolved; Cons. not know; goal: to resolve by srch diff v.r b/w cons. know

⑧ Search cost lognormal: $c_{ij} \propto \exp(\lambda_j \gamma_i)$, $\gamma_i \sim N(\gamma, \Gamma)$ — diagonal (to have positive)

accessibility of Prod. known to cons (freq in store, # recommended)

⑨ obj func: expected util − total search cost

⑩ seq srch (stp or cont): expected marginal benefit of srch:

$$B_{ij}(u_i^*) = \int_{u_i^*}^{\infty} (u_{ij} - u_i^*) f(u_{ij})\, du_{ij} \rightarrow \text{density of } u_{ij}$$

left tail — highest utility among searched products

⑪ Cons. cont search if at least one j: $c_{ij} < B_{ij}(u_i^*)$

⑫ Search strategy: $S_i \cup \bar{S}_i$ (partition, searched options vs. unsearched)

⑬ all info. about S_i contained in $u_i^* = \max_{j \in S_i} \{u_{ij}, e_{i,J+1}\}$

⑭ state of system $(u_i^*, \bar{S}_i)$

⑮ value func: $W(u_i^*, \bar{S}_i) = \max\big(u_i^*, \max_{j \in \bar{S}_i} \{-c_{ij} + \beta_i [\underbrace{F(u_i^*) \cdot W(u_i^*, \bar{S}_i - \{j\})}_{u_{ij} \le u_i^*}$ (cdf u_{ij})

$+ \underbrace{\int_{u_i^*}^{\infty} W(u_{ij}, \bar{S}_i - \{j\}) f(u_{ij})\, du_{ij}}_{u_{ij} > u_i^*}]\}\big)$ Bellman equation

⑯ assump: ① Cons. has full info prod attrib V_{ij}

② $E(e_{ij} e_{kl}) = 0$ (well defined pref prior search, not learn about pref during search process)

③ no context & ref. effect

⑰ reservation utility (indif b/w search & not search): z_{ij} def:

$$c_{ij} = B_{ij}(z_{ij}) = \int_{z_{ij}}^{\infty} (u_{ij} - z_{ij}) f(u_{ij})\, du_{ij} \qquad z_{ij} = B_{ij}^{-1}(c_{ij}) \text{ monotone}$$

⑱ Component of search:

① selection rule: compute all z_{ij} & sort descend. ordr (search highest)

② stopping rule: stop when $u_i^* > \max_{j \in \bar{S}_i}(z_{ij})$

③ choice rule: once search stopped collect u_i^* by choosing max util altntv in S_i.

⑲ assump: e_{ij} indep V_{ij} (search of one product does not affect search of other)

⑳ not curse of dimensionality, since cons. samples descend. order. seq search

㉑ π_{ij} inclusion j in optimal search set cons. i $\quad \pi_{i,r(1)}$ inclu. highest rank

$$\pi_{i,r(j)} = \Pr\left(\max_{k=1}^{j-1} (V_{i,r(k)} + e_{i,r(k)}) < z_{i,r(j)}\right) = \prod_{k=1}^{j-1} F(z_{i,r(j)} - V_{i,r(k)}) \quad j > 1$$

all $j-1$ draws fell short ⇒ inclusion probability, $\pi_{i,r(1)} = 1$ at least search one

properties (1) $\pi_{i,r(j)} > \pi_{i,r(j+1)}$ (2) $\pi_{i,\{r(j) \text{ and } r(j+k)\}} = \pi_{i,r(j+k)} = \min(\pi_{i,r(j)}, \pi_{i,r(j+k)})$

prob $r(j+k)$ is in the set

(3) $\Pr(S_{i,r(k)}) = \pi_{i,r(k)} - \pi_{i,r(k+1)}$

optimal set of size k for indiv i ← S_{ik}

㉒ Data: view rank data, average sales price, appearance on other prod list

percentile ranking $\hat{r}_{jt} = \dfrac{r_{jt}}{\max_k \{r_{kt}\}}$ highest rank j $\quad \hat{r}_j = \dfrac{\Sigma_t \hat{r}_{jt}}{T}$

㉓ Conclude: # appearances on view rank depend on attrib & price

㉔ # product detail pageview λ_j

㉕ appearance on different colum → affect search cost

㉖ Commonality index: relationship b/w products:

$$CI_{jk} = \frac{n_{jk}}{\sqrt{n_j}\sqrt{n_k}} \qquad (j,k) > (j,l) \longleftrightarrow CI_{jk} > CI_{jl}$$

㉗ $u_{ij} = X_j \beta_i - \alpha_i p_j + e_{ij}$ $\qquad \log(\alpha_i) \sim N(\beta_p, \sigma_p^2)$

$\beta_i \sim N(\beta_0, \Sigma_\beta)$

Search cost: $c_{ij} = \exp(\gamma_{0i} + \gamma_{1i} \lambda_j)$

$\gamma_{0i} \sim N(\gamma_0, \sigma_{\gamma_0}^2)$, $\gamma_{1i} \sim N(\gamma_1, \sigma_{\gamma_1}^2)$

Firefox project theory

① 3 type players { developer, reviewer, user (agent based)

② developer: max $R \times d - q \cdot c$ (rate, download, Cost of versioning)

Given opportunity cost Constraint

③ user: decide to download given type

mixture ← type (signaling) { review searcher ⇒ last reviews weight; general signal searcher: { # Download, # users, # stars

③ Reviewer: find Error → Exponentially distributed time

decide to share or not: Poisson split p

⇒ ① erlang time of 2nd & 3rd....

② depend on # users ③ selection model

⑤ if review searcher search cost C_s: Sequential

search ⇒ learning

Stopping rule: explanation close to the need (remove uncertainty)

⑥ prospect theory: different effect of positive vs. negative → How model { MKSci, mgmnt Sci

Reference stars as # stars of top 2 cates

⑦ out come of paper: { what signaling, search cost, policy } by each ① party ② platform owner

⑧ Cost of usage of open src product

FFP Optimal Number of versions (JMIS): Hui, Yoo, TAM 2008 ①

①

① pricing ⇒ as many versions as possible

② marginal benefit of an additional version decreases rapidly

③ Cognitive Costs have more profound effects on versioning than menue Cost summer project | what drives decision to download? | Experience Good /review | if in first look found not related discard | [illegible]

① List of Covariates:

① interaction significant b/w reviews & product variety (context popular vs. niche products)

↑ Prod. Var ⇒ ↓ positive neg. vs. rev (mor on popular Prod)

↑ " ⇒ helps niche Prod (get more dwnlds)

(Zhao, Duan 2012 EC research & apps) Quantile Reg

② Micheal trusove, Decker 2010 (Journal of research in marketing) → percieved strength & weakness of product

- aggregate cons. pref (product development, Conjoint)
- preprocessing NLP
- relative role of ① product attribute ② Brand name on overall evaluation (heterogeneity) - mobile phone Mkt
- clarify pros & Cons (remov unrelated, aggregate redundant, implicit → Explicit, eleminate less frequent, binary Coding (Dummy))
- discrete distribution of pref. (latent class)
- 3 model: (1) Homogen (2) Hetregen(disc) (3) Heterog (Contin)
- $\tilde{x}_{km}$ (=1) brand m (of) - x_{kh}: func. attrib get level

$Prob^{PR}(Y=y) = \frac{\lambda^k}{y!} e^{-\lambda}$ (Poisson) $\lambda = e^{\alpha + \sum_{h=1}^{2}(\beta_{h1} x_{h1} + \beta_{22} x_{22}) + \sum_{m=1}^{M} \delta_m \tilde{x}_m}$ ← Given Certain Pro/Con summary

↳ observing evaluation

2. func. (prod. attrib) k: indiv rev h = {1,2} m: # brands

$LL(\Phi^{PR}) = \ln\left(\prod_{k=1}^{K} \frac{\lambda^{y_k}}{y_k!} e^{-\lambda}\right) = \sum_{k=1}^{K} y_k \ln(\lambda) - \lambda - \ln(y_k!)$ (AIC)

hetrogen: ① Neg binom. poisson Process (Conditional mean $\Lambda = \lambda e^{\varepsilon}$)

② latent class poisson: $LL = \sum^{K} \ln\left(\sum^{I} \pi_i \frac{\lambda_i^{y_k}}{y_k!} e^{-\lambda_i}\right)$

AVG sensitivity of product eval to var qual of func attrib: Φ, Φ attrib Brand Contin brand

③ Quantifying social influence in an online cultural market (Pols 1, 2012)

① two step process { ① decision to sample (social influence) ② decision to download

② role of placement in mediating social signals

→ informational rather than normative

④ availability: prob of being sampled: $V_i = \sum_p \frac{I_{i,p}}{\sum_i I_{i,p}}$

appeal function of final listen: $A_i = l_i / \sum_k b_k$

- Conditional prob of download Curr Dwnld avail. score

- prob song i sampled: $\frac{V_{i,t}(D_{i,t} + \alpha A_i)}{\sum_j V_{j,t}(D_{j,t} + \alpha A_j)}$ (Poulya urn)

④ what makes a helpful online review (Mudambi 2010)

percieved helpfulness (① Extremity ② review depth ③ product type)

less helpful in Experience good (Star rating) | positive (more for search goods) (word Count) | (moderator) Search or Experience

③ MKT Sci. UGC

User Generated Open Src Products: Founder's Soc. Capital & Time to Prod. Release: Mallapragada, Grewal, Lilien 2012

① success measure : time to prod release.

② how ① location of Proj founder in soc. netw. ② interplay of developers & end users & ③ Proj prod charac. ⇒ affect time to prod. release ⇒ (hazard model)

Scratch

①
$prob(prod) = \sum p(\text{signal}_{prod} \mid \text{core signal}) \cdot \pi(\text{core signal})$
$prob(\text{core signal}) = \sum prob(\text{core signal} \mid \text{contribute}) \cdot p(\text{contrib})$

② ⓐ Decision to use open src prod.
ⓑ Plot decision making proc & Equations
ⓒ Game theory (competition: $f(x_2)$)

③ Candidate theory: modules available
① prospect theory (show model)
② multi stage decision making
③ social learning
④ Bass model?
⑤ platform decision making (versioning) → ① policy ② signaling
⑥ search theories (fix sample size, sequential)
⑦ advertising, signaling

④ what info set available?

---

① depth of review → information seeker (first paper match)
② credibility of reviewer
③ type of consumers → ① CNF: consult but not follow
② NC: not consult: since have more info to process
③ CF: consult & follow prod recomm
⑥ intensity of prod rev.

Entities: ① Risk aversion
② product category
③ product attribute
④ search cost: navigation
⑤ Business model, firm related

---

two main related for vertical
① sha Yang 2012 ③ Monic 2012 ⑤ Kim 2010
② NYU paper ④ Bhargava, chaudhary (verse)

Math to model your own theory | Theoretical framework: architecture

two sides
① supply: versioning | Demand: product reviews & signaling
downloads

idea → use the literature here to create research proposal for Genca as well

intuition, ⓒ Theory

Corresponding
popularity contest
① prices → stars (as indicator of cost & usage)
② credibility → variance of stars
③ versioning → like markdown shock (Best response)
④ time as variable
⑤ segmentation { ① price ≡ stars ② credibility
⑥ sequential search of products & reviews
① restructure website ← (lower search cost) → implication of structure
② add new browsing for reviews (search)

ⓘ

(*) proposed list of covariates to explain mean:
① often used with (theory?)
② part of collection (theory?)
③ other by this author (theory?) } popularity # users
④ Tags (theory?) → ⑬ from which channel? (Dwnld src)
⑤ Category (theory?)
⑥ Featured (theory?) ≡ in first page of Category
⑦ business model (theory)
⑧ ~~age of the~~ tenure of reviewer (theory?)
⑨ # reviews of reviewer (theory?)
⑩ AVG rate of reviewer ② our estimator vs ① understanding
⑪ is reviewer developer? (theory?) × popularity of developer
Plug in
→ ⑫ platform, language, application, usage, download, status

Proposed category → ① web development
② condition on ① most popular & ② featured
③ top rated

Result – No direct marketing theory except those papers for tag, and often used
⇒ find computational processes that are currently used & include customer perspective for profit maximization best respons

idea – The same way that postpurchase affects reviewing behavior
– reviewing consumption affects, download behavior (Moe, Schweidel 2012) decision to download [like household production]

Data – evolution is available from start but need method to scrap → not difficult (Javascript ≡ all avail on one page

challenge products not independent & may be purchased together

explain
① Download (BH)
② usage (BH)
③ source (BH)
④ application
⑤ language (BH)
⑥ addon status (BH)

two observable shocks:
① own version
② platform version
③ learning & stars

model ① prob of finding a fault (expon)
② hazard of time b/w versions.

Data – two groups of often used with products

How to solve simple: ① transform/translate what you don't know to what you know/have (side by side)
② limit & concrete problem ③ having even wrong goal is more efficient than not having any (e.g. eye to search)

Summer Project idea

Speed is important by conditions

intuition

1. Core sha Yang's paper
2. BLP model
3. prospect theory of different response
4. observable versioning shock as supply side process based on market (like Gonca's paper) ⇒ Data Reuse
5. poisson download process (?) logit
6. mixture model (what signal to consider)

Next Step
1. write utility using these papers [Deliverable]
2. write likelihood using these papers
3. find courses explaining + theory
4. Solve problem like math problem
5. teach what you learn (multiple summary)
6. write fiction novel

Priority
1. Data
2. theory (e.g model of fairness & prospect) (decision to download)
3. Decision's & what will affect them
4. Hypothesis to check & theory Behind
5. integration of Components (Story)
6. simulation of Best response
7. Divide data into two parts & check prediction

① 2 what affect? 3 recall Noris adv paper (process)

Contribution
1. Role of Company is affecting each variable e.g feature products, clarification (variance), ...
2. problem of endogeneity & instrument

Key words
1. marketing, MIS
2. Decision making process (Experience good)
3. product review, social learning

two types of search
1. Horizontal
2. vertical
3. targeted for module

Experience
1. university website is much better to find src's since it checks EBSCO, MBSQ, JMR, MS, ...
2. [illegible]
3. method: ① Quick abstract ② math

Theory
1. is key
2. is about process
3. organizes thought (just like software methodology)
4. simplifies things
5. make things Quick, efficient, time reduction
6. rationalizes tree of solution
7. helps you to not get lost
8. Best response & utility max, profit max, cost min (FOC) rights
9. modular thinking

e.g
1. availability
2. segmentation
3. two utilities
4. mark down time
5. mark down Depth

Decision?

- Same problem in Bip Data

- Price, promotion, Display Feature → Store level & household
- People type? → searcher → Non searcher

unobservable → Theory

- what theory in Experience good?

time series
- Version time
- Downloads of two
- # Stars

What likelihood?

p(type)
- Gamma
- Exponential
- Beta
- → probit, ~~~~ logit

selection

Decision of Developer: when to respond to reviews

problem: Data lack of Demand time series for the rest of products

Decision of Customer

Stars as reputation system

- Threshold that customer changes his mind about product → unobserved

key: Translation

even wrong code is good, since it is complex & he has to debug

signaling theory perfect Bayesian Equilibrium

idea: read R code understand & convert to MATLAB

Run Bugs code on ~~real~~ Simulated data check the Bias & then can real data

Plan A

Good Category → web development → 5 products download available

instruments could be download of other plug ins

→ Simply define reputation elasticity rather than Price elasticity: # stars on # downloads → daily

Robust analysis
1. last page
2. last two page
3. last three page
4. overall

Combine with time series → Independent shock Questionable

Plan B ☆ → or even simply just run time series on download & stars

Steps rather than Price

① Econometrics theory? $y_{it} = y_{it-1} + \mu_{it-1} + u_{it}$?

② Game theory? what decision?

③ CGL model? mixture, random ~~coefficient~~

④ Data analysis? Fixed effect of product?

Shocks { ① Versioning
② Product review

idea ① clustering → topic

② anthology

Question ① what is main goal to be curious abot?
② what is decision?
③ how use dates? → Solution Convert # days

sample answer ① how people search product reviews?

Sample Research Question ① how developer respond to reviews and how this define success

② which strategy is better? lower version - higher version
more frequent, less frequent

Key ① Compare your data with scanner data ①
other data
② Think about other's models also helps you to adjust your versioning

→ ③ how versioning by Competitors affects

② Convert to days: $\underline{t}$ and look as dummy?

idea → hetrogeneity in process of usage is a key?

Next step: ① write mixture model in MATLAB
② write BLP (random coefficient) in MATLAB
③ write selection model in Bayesian

Lesson learned ⓪ key is translation → to understand

① Lock to subject - screw to understand
Core

② go forward & Backward b/w model, reality intuition, but more weight on model & intuition

③ when people are intelligent = they have practiced more & ruled out problems & know them (e.g Structural modeling)

modeling: $u_t = \rho u_{t-1} + e_t$

→ mixture: # obs per household is not important
= even one could help

Task list

✓ ① Rewrite normal BLP in MATLAB & Get same results

② write Bayesian in MATLAB & Get same results

③ Run Code & Download data

④ write time series Baysian in MATLAB

⑤ run time series on download data

⑥ write mixture model in MATLAB

⑦ run mixture on signaling step of MATLAB

main trick: write modular and if not work Just Submit modules & result of normal BLP

$\boxed{\text{BuG identified}}$ ↓ solve

- effect of release on product succ[e]ss

$\boxed{\text{type}}$ $\dfrac{e^{v_1}}{1+e^{v_1}}$ $v_1 \longrightarrow$ # reviews

activeness in reviewing others product

$\boxed{\text{Assump}}$ today is equilibrium

everybody has made Best choice

$\boxed{\text{Q}}$ - how model best response? to BuG?

- how identify BuG is found?

$\boxed{\text{Short answer}}$ check how Stars evolve? $\boxed{\text{Curves}}$

$\boxed{\text{Task}}$ → Date Refinement → first get dates

→ 2nd put if condition in excels

& check how evolves

problem → how calculate for

→ cumulative visible window

↓

$\boxed{\text{higher Exposure}}$

$\boxed{\text{idea}}$ $\boxed{\text{Dummy of versioning}}$

Quick look → first mover advantage

Date extraction "on_" " ." ✓

Activity list

① website process & data audit: 1 day

② Gonca paper and model: 0.5 day

③ Hortascu paper: 0.5 day

④ BLP audit: 0.5 day

⑤ text mining { ling pipe / Word net 1 day

⑥ modeling, significance, unobservable 1 day

Next step – Download and extract

① Versioning info (Date, version, topic)
② Review (Date, version, Content, Person)
③ all person's review for reviewers
④ Download data for one plugin

first priority
- how time to solve bug will affect marketing success
measure: # downloads

lesson learned - degree of freedom of system is decision making controls

Person choice ⟷ Product chosen
① Consistent → pref
② mixture → product selection

method / strategy

① Framework fit

② interdisciplinary nice

Problems

① Good data (one layer unobservable) → (sol) aggregate demand
② Cool model (Combination, interdisciplinary)
(sol) aggregate demand:
③ significance & implication for decision making
(sol) ① Bias? ② Good question
④ Good Game theory behind
(sol) Best response? (I)

idea

① what problem to solve first as developer, prioritize?
② which product review is more valuable to consider? ⟷ Best response
2.1 How timely (hazard)?
⇒ Content of review is more important than # of stars
→ text mining

Logit problem → ① Consumer choices? (II) ⟷ (III) unobservab

③ whether as developer I have to be active reviewer?

types ① those only stars to date
② those who read content

Complementary tools (why?) → intuition

① wordnet → java interface

② Yago

idea: email them and ask for stats

→ help them decide better

→ my contribution to open source community

2 main Panel Data

① Reviewing → person # of reviews

② versioning, review time

search

slice of Data with lots of info

① developer
② collection

type

① developer
② user

problem:

Download

① not exist

② aggregate demand

Sol 1 product category selection (narrow down & simplify)

Sol 2 Granular to product

Sol 3 → one product → effect of its ① versioning ② review

& competitor's en usage ① entrance & versioning ② reviews

mixture ① panel ② multiple observation per item

① Collection Dummy

② used with Dummy

main question now ① what is likelihood and model (hint you can use instruments)

Signals of Quality

① Download: # users

② star distribution

③ #reviews

④ # versions

objective: explain a behavior and predict it

$\boxed{\text{Statistics}}$ $\boxed{\text{May 1st}}$ (5)

per $\underbrace{\text{Household/individual}}$ (over T)

$$\prod_{i=1}^{N} \left[q \times L_i(S) + (1-q) L_i(NS) \right]$$

over # indivs

$$\frac{e^{V_i}}{\sum e^{V_i}} \qquad \frac{e^{U_i}}{\sum e^{U_i}}$$

$$q = \frac{e^{Z_i}}{e^{Z_i}+1}$$

Likelihood = mixture

$$\prod_{t=1}^{T} \frac{\pi_1}{\sqrt{2\pi}\sigma} e^{\frac{-(y_t - X_t\beta_1)^2}{2\sigma^2}} I[s_t=0] + \frac{\pi_2}{\sqrt{2\pi}\sigma} e^{\frac{-(y_t - X_t\beta_2)}{2\sigma^2}} I[s_t=1]$$

→ what weighted average?

(6)

review	users	
3,811	6,811,652	video download helper
1,243	2,213,551	Easy Youtube video
1,285	1,826,848	Download them all
1,207	1,889,185	Download Statusbar
380	1,721,742	FlashGot
331	1,218,912	Flash video downloader
116	1,213,672	Microsoft .NET
775	1,006,915	FireFTP
334	590,552	ant video downloader
X 161	540,096	Download flash & video

Q: opinion leaders affect usage?

type of Comments

$\boxed{\text{2SLS}}$ endogeneity

Good product → more version → more opinion leader

Dummy Developer

whether → $\boxed{\text{activity}}$

↓ unobservable

$\boxed{\text{Structure = theory → Common knowledge}}$

review → version → review → success

روش جمع تحقیق ← استفاده در حوزه های زبانشناسی و کامپیوتر ها

IO و تجزیه بازارها و اقتصادی مدلهای ترجمه و تبدیل ←

mixture استفاده از ← برای تبیین و توضیح پدیده

ترکیب با علم دیگر niche

→ $\boxed{\text{Niche of text mining}}$

— location instrument

(1) what decisions? strategic interactions?

(2) what types unobservable

(3) what bias? or significance?

(1-1) what competing forces & equilibrium?

Questions:

1. why download variation in different days AR(P)?

2. effect of number of version on downloads and reviews → { indogeneity? Game theory behind }

3. why some products often used with others?

4. how revealing stat about product affects number of users?

5. why popular addons are used with non popular?

6. Decision: Product category

when to release → 1. Version history

Category decision ← 2. Category: same - not same : Explain

work to collaborate ← 3. Common reviewer → opinion leader

4. ~~base centrality~~ # stars

Remove bugs ← 5. positive versus negative reviews

6. # reviews

packaging decision ← 7. Same collections?

↑ implication

need crawler

7. Question: Cool part of model? → key: unobservable theory

draw bs nature unobservable

type? → Extensive searcher 3 pages → non searcher

segments: Novice, professional → normal searcher → { 1. searcher/non searcher 2. elastic & not price elastic 3. satisfaction & loyalty }

Product → Good → Bad, inferior → # stars → Behavioral type

Product attribute: Category? rating? Versions?

8. type signals → read review (unobservable) → passed with Product list

9. unobservables { 1. elasticity (high vs. low) 2. searcher non searcher 3. utility - preference } → type

10. learning - Equilibrium - time - thresholds

11. relevant: 1. Distance (cost) naving 2. time cost

{ 1. opinion leaders 2. intrinsic preference 3. product category only candidate → (# downloads)? 4. timing effect (versioning) 5. packaging effect }

Q. why less popular products are used with high popular? → Based on usage

Q. why some products have small number of reviews, but high number of usage? / category

explain:

1. Role of products in product category
2. number of stars 1. main 2. distribution → negative idea ← loss aversion
3. number of product reviews
4. versioning history (time) { 1. number 2. time }
5. part of Collections (3 collections)

Dependent:

1. often used with Collection: 6 products
2. # users b

unobservable | theory

- Technical vs. social → 1. Contributor: (Product review Contribution) # review lifetime → 2. free rider

-

why interesting:

- what decisions? Controls

{ 1. packaging? 2. engage customer review? 3. which product category to compete in? 4. release management aggregate or frequently? }

1. Bayesian mixed logit - Probit

unobservable behavior type → Behavior 1 → Consider data item → Behavior 2 → or not → attention to different signals

↓ segments

2. Data: aggregate Demand ↓ usage

Plan B — tentative

BLP in Bayesian? → pradeep code of Gauss run & then convert Bayesian & run on this data

- significance: if not taken into account ⇒ biased

Reference

interesting Question
interesting model — mixture hidden markov
— time series
— Neural network
— multinomial probit

intuition

① How download is going on

② is there a trend or correlation

③ what to Explain?

④ what are Competing mechanisms (I)
- ① – learning
- ② – signaling prospect, assymmetry
- ③ – Behavioral (trust, fairness / regret)
- ④ – Stock piling
- ⑤ – Breaks (DID) → Releases
- ⑥ – Forward looking myopic
- ⑦ variety seeking

⑤ what Consumer decisions? (i) Stakeholders?

⑥ state of people and how move — searcher / non searcher

⑦ Dummy variable, fixed effect — price elastic / Non price elastic

⑧ Theory underpinning

⑨ panel data (II)

⑩ why interesting? ⟨1⟩ Curiosity → ⟨2⟩ policy implication

⑪ what observab? / what non observable? (III)

⑫ multi data source

⑬ what is hypothesis (IV)

⑭ list of stakeholders: ① Customer ② reviewer ③ platform owner ✓ ④ product owner

⑮ what decisions? ✓ / what implication for policy? ✓ → app store / Competing forces? ✓

⑯ panel data? → for change state mixture

⑰ Granger Causality

⑱ how crawler?

⑲ Likelihood function?

⑳ what is game? strategic interaction?

㉑ interactions?

㉒ objective function of platform owner?

---

① Releases → Dummy

② Nickel bias of time series

---

(3) # stages of the game decision parameters

⑧ matching – Deviation – indifference
⑨ patience time, discount, netrops
⑩ Product diff. → Vertical Cartel
⑪ Coalition – Cartel → single entity
⑫ Search Cost → Core
⑬ non Cooperative vs Cooperative
⑭ limited Resource → Bass model Hazard
⑮ 2SLS – endogeneity / SUR / Best Performance
⑯ multi stage entry

㉙ Distribution cost
⑰ Incentive Compatibility
⑱ moral hazard
⑲ Experience goods
⑳ trust, Punishment
㉑ Contractual Precedent
㉒ optimal policy (recursive)
㉓ aggregate demand
㉔ Cherry Picking
㉕ Basket level
㉖ Household Production
㉗ dispersion
㉘ unobservable type

---

① add on status = c
② plat form down
③ downloads & usage → down on weekends
④ languages the same except drop
→ ⑤ users: downward trend
⑥ Tags – Product Categories
⑦ often used with Collections
⑧ stats
- ① Download trends
- ② opendownload
- ③ download manager trends

Q. Data

→ ① multi product
② Commenter name → Demographic ethnicity → ella site
③ application life time (Bass model)
④ add on versions life time
→ ⑤ panel data?
⑥ share of channels of download { detail page / add on manager
⑦ Version history

user
- expert (in computers)
- novice

Firefox

Platform

Competition → switch to chrome

developer

→ Time series
↳ selection

↳ control
for dev characteristics

1. Need it!
2. Suggested to me
↳ I've seen it sw
~~Stated~~ refer

Addon

1. download without Reading Review
↳ why? → bc ~~there is a~~ the cost
is trivial

novice

2. I may look
at star reviews

Cost of search
price → free

low
risk

expert →

$\boxed{}$ prior knowledge { 1 / 0 → review become more important

Search { narrow / wide

1. Is firefox the primary browser?

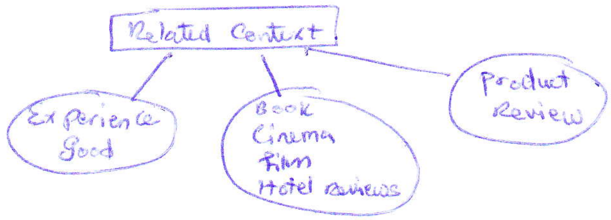
user { تازه‌کار → ~~review~~ review bad خیلی کم تأثیر داره
review خوان → بیشتر تأثیر داره

Add on
تازه‌کار ۲۰ / میزان استفاده بالا → Review بیشتر تأثیر است

۲۰ سن ... ریویوخوان حرفه‌ای → ... Review تأثیر

Developer → تصمیم Public
... Data DL

product Review litratur



1. Compare what?
2. Control everything help to infor Counterintuitive
3. different with other contexts.

Product review literature

①

① Tried and tested: The impact of online hotel reviews on Consumer Consideration (2009, Tourism management, Vermeulen, Seegers)

ⓐ Consideration set theory (impact of online reviews on consumer choice)

includes
① review valance (positive vs. negative rews)
② hotel familiarity (well known vs. lesser known)
③ reviewer expertise (Experts vs. non-expert rews)

H: exposure to online reviews enhances hotel Consideration in Consumers.

positive : improves attitude

positive & Neg : increase awareness

②

$A = [a_{ij}]$ → # sit

$x = (x^1, \ldots, x^m)$ leg Capac. → how many sits

$R_t = (R_t^1, \ldots, R_t^n)$

multiple

④ how many taken out → 1 2 3 4

diff. Resource

1 2 3 4 5

remove index

Simplify

X over backs

Cont. j out

$F_t(r) \to$ dist $(R_t^j(r))$ ~~finite~~ mean

② api → E → src_add on mgr

BW → search

~~AG~~

BR – collection

mst pop: C + Y + AM ~~+AK~~

~~search~~ ~~AJ + AE +~~ C + W + AK + I + J + Q + AA + AL + BD + BH + BE + BZ

282

add on dtl: AG BF

primary detail

① search most popular

C + I + L + Q + W + AA +

AE + AK + BD + BH + BZ

universe likes speed

Don't care result just ~~take~~ adventure

choice of not energetically do is always failure

more than ~~100~~ 1/total # versions

3 days 2013 { 5/25, 5/20, last day

open mind

Solution pro

hash

persistence

greed

give Heart to work

0,1 $B6

hashtable

creativity

Translation

Data is Gold

Guts

Quick

end 6/8/2011

total user disabled

user disabled: AF

User enabled: Z

Status = disable / enabled

sum (linux, mac, wind)

Generalize = more data

G

→ Right – left – enter

→ Control range

for downloads

⇓

spec date first

---

reviews

① path

② # pages

MK sure no blank

Date Extractor → perl

Observable stars

obsStars → mean

① mean ✓

② variance ✓

③ # reviews ✓

3/22/2007

problem of missing dates & old dates

$V \; (E(x^2) - E^2(x)) \times \frac{n}{n-1}$ → ③ match Dates

Versions (later)

multi coll { linux share vs mac share, wind share; AVG-str, std-str; Src Downloads

eye for plugin

only significant:

* Download server down today

last year is fine

Complete data finish.p

* version → output

* spec date

save xls

sum (C6:) shift ← Enter

$B6 × 0,1

dis/enab

last year ⇒ –2 days

look for dates $\frac{D6}{B6}, \frac{F6}{B6}, \frac{E6}{B6}$

win lin osx

0,9 last year

2013 Eng R6 + AK6 + BG6 + BH6 + BY6 + CD6 (eng ttl) / B6

(same sheet)

2012 R6 + AL6 + BF6 + BG6 + BX6 + CC6

2011 N6 + AE6 + AY6 + AZ6 + BG6 + BV6

↳ one leg

(2013) AN6 + BK6 + BL6 + CE6 + CJ6 + T6

Gek URL

Q6 + AJ6 + BD6 + BE6 + BF6 + CA6

Date Extractor → DateExtractor

Stars → Stars output

Orig Date → Date output

→ CompletedataStars → obsStars

obsAllStar

first row empty

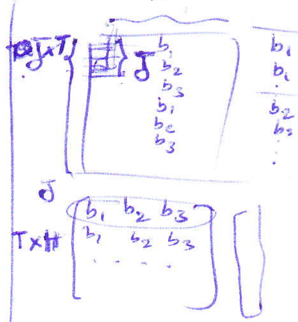
$(\alpha_i, \beta_i) \sim (0, \Sigma)$

$(\alpha, \beta) \begin{bmatrix} 1 \\ X \end{bmatrix}$ $\quad \underline{v_i}$

δ mean

$\Gamma = \text{Cholesc}(\Sigma)$

$\delta \Rightarrow 2SLS \Rightarrow \xi_{jt} \Rightarrow \xi Z'(Z'Z)^{-1}Z'\xi$ unons



① Nevis paper
② ys Andrea

$b_1\ b_1\ b_1\ b_1 \quad 1 \quad b_2 \ldots \quad] \ b_3 \ldots$

$15000 = 300 \times 50 = T \times H$

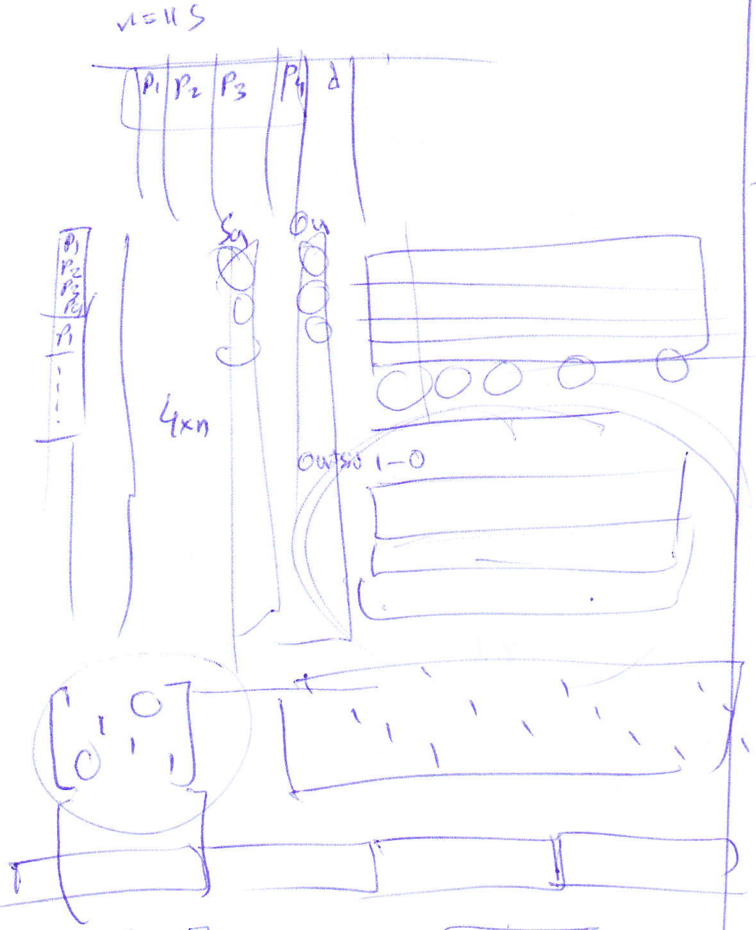
write comments MHz

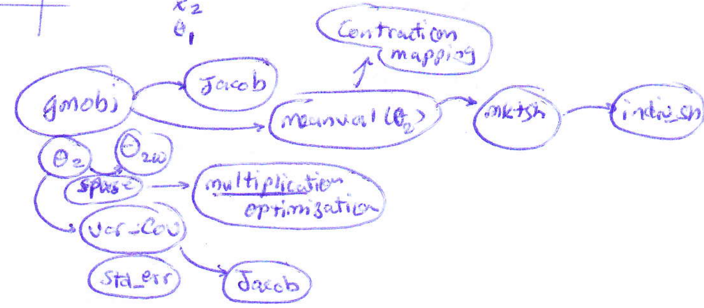
$y_{md} = \theta_1$ [2SJ]

$R^2 = 1 - \frac{MSE}{MST}$ $(res - \mu_{res})^2$ $= (y_{md} - \mu_{y_{md}})^2$

Global variables (Blue)
θ_i
θ_j
x_2
θ_1

$v = 115$





ind_sh: indiv prob. of choosing each Brand
mkt_sh: mkt shr for each prod.
meanval: mean utility level

List of hypothesis's

1. Consumer's variance of Star rating will reduce download
2. Learning ~~increase~~ rate will decrease during time
3. At the beggining of life cycle of product learning is more effective than the later phases
4. Quick versioning will increase the product download
5. Time between versioning decreases as number of ~~products~~ negative product reviews increases

X ⑥ whether consumers learn Bayesian from product reviews of experience goods
⑦ whether quick versioning is a good idea
⑧ whether # tags affect consumer's decision by reducing search cost
⑨ negative product reviews will harm usage & downloads more at the beginning of Prod. life cycle than the end.

⑩ Consumers learn more about negative product reviews than positive product reviews.

$\frac{\text{Share}}{1-P}$ $\quad y = \ln\left(\frac{P}{1-P}\right)$

Logist $= \quad P = \frac{e^{X'\beta}}{1+e^{X'\beta}}$

$\Rightarrow P + Pe^{X'\beta} = e^{X'\beta}$

$\boxed{e^{X'\beta}(P-1) = -P}$

$Err = y - X\beta$

Variance $= (X'X)^{-1} X' e e' X (X'X)^{-1}$

$\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 & z_2 & z_3 & z_4 \end{bmatrix}$

① 3 type player — ① developer — Versioning, Profit max
② review
③ user

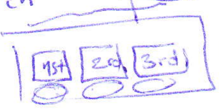
② Review $R \cdot d - qc$ — user Errors Expon

③ user — ① ~~last re~~ general searchers ⇒ # down
② specific → reviews

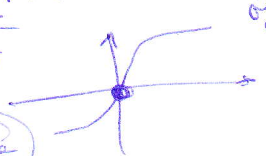
④ review ⇒ sequential

$c = MR_{search}$ $\quad c = \int_{z^*_t}^{t} (z^* - z(t))\,dt$

⑤ Learning user Review ⇒ $\frac{1}{\sigma}$ ↗ match # AVG



⑥ prospect



reference ↳ estimate general # Stars Category

⑦ equalibrium

⑧ Version

① sequential decision Review → user → Reviewer ↳

② extra Cost of feedback from somewhere else → but Cost

③ the 3 last prod review impact

Econom — Freqent, Bayesian ⟩ fit dat

Structural modeling (Chicago) — assumption static; Theoric assumption; Counter factual — Simulation — MCMC, Dynamic Prog

Theory — Game theor, probabilit

BLP

① developer selects — time of version (b/w); whether to version Dummy

① explain # ~~reviewer~~ Developers ≡ Cost of versioning
opportunity Cost ≡ ② # ~~plat~~ plugin developed

③ # total downloads

④ # stars

⑥ Binary? hazard?

Developer Question: what other information to take into account?

⑤ like queuing system of Operating System?
⑥ or like economic theory?

⇒ ① # developers d_j
② total downloads nd_{jt}
~~③ # negative ...~~ ⑤ total users nu_{jt}
③ total time from ~~error~~ low rate ts_{jt}
④ # negative stars from 5 : ns_{jt}

daily: specific version

$$p(v=1 \mid vr=c) = \frac{\exp(\alpha_i + \beta_i d_j + \gamma_i nd_j + \theta_i ts_j + \cdots)}{1+\exp(\cdots)}$$

⇒ Dynamic programming? what is probablistic?
what is Expected rather than real Value?
probablity of the same error or another neg Product review in the next period ((1) how mainstream or niche the error is, (2) how ~~efficient~~ effective the responding of Cons. to neg prod rev.)?)
how estimate this prob? (1) # users (2) # stars
(3) # positive reviews on current version
≡ assume Correct expectation

$$\Rightarrow \begin{cases} u_{jt1} = \alpha_i + \beta_i d_j + \gamma_i nd_{jt} + \theta_i ts_j + \lambda_i ns_j + \eta_i nu_{jt} + \varepsilon_{jt} \\ u_{jt0} = \gamma E[\mu_{i,t+1} \max\{u_{jt1}(s_{t+1}), u_{j0t+1}(s_{t+1})\} \mid s_t] \end{cases}$$

$\mu_{j,t+1} = \varphi(u_{jt}, s_{jt})$

② $u_{ijt} = \alpha_j + \beta_i a_{jt} + \gamma_i s_{jt} + \theta_i c_j + \delta_i t_j + \varepsilon_{ijt}$

(not versioned) # Cat — (search cost) # tags (usability)

(non searcher)
u_{ijt}
Sha Yang paper is indiv level
⇒ I want to do it in aggregate form
how?
how high level estimate of # downloads?

Q6+ R6 +AE6+ AO6 +AU6+ BM6
V6+ W6 +AI6+ AS6+ A26+ BG6
V6+ W6+ AI6+ AS6+ AY6+ BO6
R6+ S6+ AF6+ AQ6+ AV6+ BN6

V6+ W6 AK6+ AT6+ BABS6
S6+ T6 + AF6+ AO6+ AU6+ BM6

L6+ AB6+ AS6+ AT6+ BK6 + BP6
Q6+ R6 + AF6+ AO6+ AT6+ BJ6

Start simple

---

1. Comment should be opened
2. wait for loading of twitter then save
3. sequential file name

---

1. first pilot for one
2. spec date problem ⇒ full date (version)
3. weekly aggregation (stats) ⟺ match sale trend
4. cross sectional search of these people (race, gender, ...)

50 people × >100 observations

*

5. Dummy of whether post Exists & AVG based on that weekly & dummy of week creation
6. # posts per week
7. proposed: UGC →
   1. # tag
   2. ~~title~~ # words
   3. time of the day classified
   4. day of the year
   5. # pages

Respons
1. facebook
2. twitter
3. g+
4. shares
5. Comments

? Control topic
2. life cycle
3. Good will creation during time