Demand Estimation with Complementarity and Variety Effect

Heng Ju, Yifu Shao

College of Business, Shanghai University of Finance and Economics

November 16, 2024, at ***

I. Motivation

Complex cross-market strategy

- Merger/acquisition: GE and Honeywell, KE Holdings and iKongjian, Qualcomm and Intel
- Business practice: bundling (e.g., 88VIP, Apple One, App Store + iPhone)
- Platform: cross-market externalities (e.g., Armstrong and Wright, 2007)

Concerns arising from demand correlations

- (Countering) public policy (e.g. antitrust scrutiny)
- Business strategy
- Industrial policy
  - Digital/platform economy
  - Broader policies, e.g., FDA "Healthy" label redefinition

I. Motivation

Suppose a new product enters the choice set, $\{\mathrm{A}\} \Rightarrow \{\mathrm{A}, \mathrm{B}\}$

- Share of A
  - Decrease due to substitution
- A higher share?
  - Complement consumption (A + B); "quality" increase (A-alone)
- The **demand correlations** of A and B are fundamental to address many antitrust inquiries
  - E.g., bundling, platform pricing, cross-market mergers, spillover in digital platform, ...

Demand correlations

- Complementarity: "interaction" from consuming B on top of A
- Variety effect: introduction of a B on A-alone
- Structural error: one likes A may like/dislike B
- **Under-studied in antitrust research but important in practices**

I. Motivation

Empirical challenges

- Proper modeling for both concepts
  - Model suitability; dimensionality
- Identification
  - Lack of moment; endogeneity; heterogeneity
- Data requirement
  - Aggregated data (market share) fails. Consumer-level data is scarce.
- Computationally feasible tool for merger prediction
  - Internalizing externality; bundling incentive

I. Motivation

Empirical challenges

- Proper modeling for both concepts
  - Model suitability; dimensionality
- Identification
  - Lack of moment; endogeneity; heterogeneity
- Data requirement
  - Aggregated data (market share) fails. Consumer-level data is scarce.
- Computationally feasible tool for merger prediction
  - Internalizing externality; bundling incentive

Roadmap

- Empirical framework to model a hypothetical demand system
- Misidentification when using order-level data
- Monte-Carlo experiment
- Bundle incentives in a stylized merger

II. Related Literature

Complementarity:

- Superadditivity framework (Gentzkow, 2007; Berry et al., 2017)
- Recent works using market share data (Iaria and Wang, 2019; Wang, 2021)
  - Supply-side assumptions
  - Demand inverse techniques

Variety effect:

- Taste for diversity or diversity in taste (Lancaster, 1990)
- System competition (Katz and Shapiro, 1994)
- Uncertainty, or preference for flexibility (Kreps, 1979)
- Brand perception (Berger et al., 2007)

III. Empirical Framework: Model Setup

- $t = 1$: $j \in \{A_1, A_2\}$
- $t = 2$: introduce B, $j \in \{A_1, A_2, B\}$
- Quasilinear utility for single product j: $U_{ij} = \delta_j + \alpha(y_i - p_j) + v_{ij}$
  - Outside option: $U_{i0} = \alpha y_i$
- Normalization: $u_{ij} = \delta_j - \alpha p_j + v_{ij}$
- Free covariate matrix in unobervable utility $\mathbf{v}_i \sim N(0, \Sigma)$

$$\begin{bmatrix} v_{iA_1} \\ v_{iA_2} \\ v_{iB} \end{bmatrix} \sim N\left(0, \begin{bmatrix} 1 & \sigma_{A_1 A_2} & \sigma_{A_1 B} \\ \sigma_{A_1 A_2} & 1 & \sigma_{A_2 B} \\ \sigma_{A_1 B} & \sigma_{A_2 B} & 1 \end{bmatrix}\right)$$

III. Empirical Framework: Model Setup

Consumer make bundle-level choice r

- At $t = 1$, $r \in \{0, A_1, A_2\}$
- At $t = 2$, $r \in \{0, A_1, A_2, B, A_1B, A_2B\}$

VE as demand shifter:

- $\Phi_{A_j} := u_{A_j}^{t=2} - u_{A_j}^{t=1}$
- Utility difference of A_j resulting from introducing B, ceteris paribus.

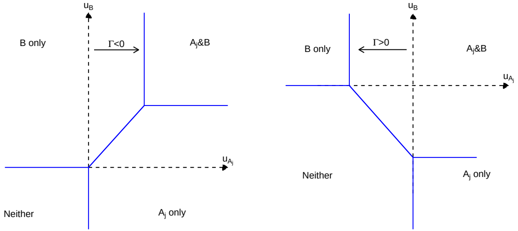
CMP as superadditivity:

- $\Gamma_{A_jB} := (u_{A_jB} - u_B) - (u_{A_j} - u_0)$
- "incremental utility from consuming A_j on top of B" minus "utility from consuming A_j alone when B is available"
- From Gentzkow (2007), A_j and B are complements iff $\Gamma > 0$

Bundle-level utility: $u_{i,A_jB} = u_{i,A_j} + u_{i,B} + \Phi_{A_j} + \Gamma_{A_jB}$

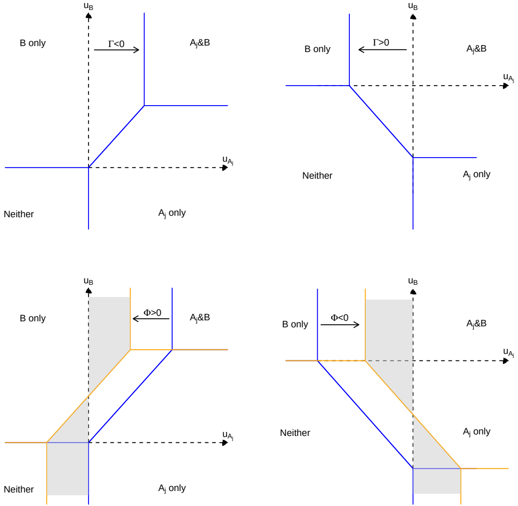
III. Empirical Framework

Aggregated share fails in identification



III. Empirical Framework

Aggregated share fails in identification



III. Empirical Framework: Identification

When $t = 2$:

- 5 order-level demands: $P^2_{A_1}$, $P^2_{A_2}$, P^2_B, $P^2_{A_1B}$, $P^2_{A_2B}$
- 8 parameters: $\underline{\delta_{A_1} + \Phi}$, $\underline{\delta_{A_2} + \Phi}$, δ_B, α, Γ, $\sigma_{A_1A_2}$, σ_{A_1B}, σ_{A_2B}
- $Q^2_j = \sum_{j \in r} P^2_r$ identifies the mean utility
- P_{A_jB} only identifies a mixture $f(\Gamma, \sigma)$
  - "Exclusive variable": $\frac{\partial Q_j}{\partial x_j} \cdot \frac{\partial Q_k}{\partial x_j} > 0$
  - Consumer panel: $y^1_i \perp y^2_i$
- The "nest" share $(Q^2_{A_1} + Q^2_{A_2})/Q^2_B$ identifies $\sigma_{A_1A_2}$
- (**Exogenous**) price changes identifies α

III. Empirical Framework: Identification

When $t = 2$:

- 5 order-level demands: $P^2_{A_1}$, $P^2_{A_2}$, P^2_B, $P^2_{A_1B}$, $P^2_{A_2B}$
- 8 parameters: $\underline{\delta_{A_1} + \Phi}$, $\underline{\delta_{A_2} + \Phi}$, δ_B, α, Γ, $\sigma_{A_1A_2}$, σ_{A_1B}, σ_{A_2B}
- $Q^2_j = \sum_{j \in r} P^2_r$ identifies the mean utility
- P_{A_jB} only identifies a mixture $f(\Gamma, \sigma)$
  - "Exclusive variable": $\frac{\partial Q_j}{\partial x_j} \cdot \frac{\partial Q_k}{\partial x_j} > 0$
  - Consumer panel: $y^1_i \perp y^2_i$
- The "nest" share $(Q^2_{A_1} + Q^2_{A_2})/Q^2_B$ identifies $\sigma_{A_1A_2}$
- (**Exogenous**) price changes identifies α

When $t = 1$:

- 2 product-level demands: $P^1_{A_1}$, $P^1_{A_2}$
- 4 parameters: δ_{A_1}, δ_{A_2}, α, $\sigma_{A_1A_2}$
- $P^1_{A_j}$ identifies δ_{A_j}; mean utility difference identifies Φ

Either consumer- or order-level data identifies the model

III. Empirical Framework: Identification

Misidentification occurs when firm endogenously set price

- At $t = 2$, firms know (Φ, Γ) and set price $p_j^2(\Phi, \Gamma \mid \cdot)$
- Consumer-level data still identifies the model (Villas-Boas and Winer, 1999)
- Order-level data fails: $Q^2_{A_j} - Q^1_{A_j} \propto \Phi + \alpha(p^1_{A_j} - p^2_{A_j})$

III. Empirical Framework: Identification

Misidentification occurs when firm endogenously set price

- At $t = 2$, firms know (Φ, Γ) and set price $p_j^2(\Phi, \Gamma \mid \cdot)$
- Consumer-level data still identifies the model (Villas-Boas and Winer, 1999)
- Order-level data fails: $Q^2_{A_j} - Q^1_{A_j} \propto \Phi + \alpha(p^1_{A_j} - p^2_{A_j})$

2-step estimation: time-invariant Θ_1

- When $t = 1$, get estimated $\hat{\Theta}_1 = (\hat{\delta}_{A_1}, \hat{\delta}_{A_2}, \hat{\alpha}, \hat{\sigma}_{A_1 A_2})$
- When $t = 2$, given $\hat{\Theta}_1$, estimate $\Theta_2 = (\delta_B, \Phi, \Gamma, \sigma_{A_1 B}, \sigma_{A_2 B})$

Any other potential endogeneity caused by $\mathbf{v}$

- Mean utility with instruments (Berry et al., 1995; Goolsbee and Petrin, 2004)

III. Empirical Framework: Identification

Or, add supply side moments

- Prices contain information about the demand
- From FOC of j, $Q_{jt} + \frac{\partial Q_{jt}}{\partial p_{jt}} (p_{jt} - c_{jt}) = 0$:

$$\mathbb{E}\left[p_{A_j,t} - c_{A_j} + Q_{A_j,t}\left(\frac{\partial Q_{A_j,t}}{\partial p_{A_j,t}}\right)^{-1}\right] = \mathbb{E}\left[\epsilon_{A_j,t}\right] = 0$$

  where $\epsilon_{A_j,t}$ is a iid cost shock
- Pro: one-step estimation by assuming FOC and *constant* c_{A_j}
- Con: wrong assumptions; optimal instruments and weight in GMM
- Our 2-step process can be estimated by ML, and avoid supply assumptions.

III. Empirical Framework: Specification

- Bundle-level utility

$$u_{irt} = \sum_{j \in r} (\delta_j + \Phi_{jt} - \alpha p_{jt} + v_{ij}) + \Gamma_r + \varepsilon_{irt}$$

- B is introduced at T
- $r \in \{0, A_1, A_2\}$ when $t < T$ and $r \in \{0, A_1, A_2, B, A_1B, A_2B\}$ when $t \geq T$
- $\varepsilon_{irt} \sim \text{Gumbel}(0, 1)$. $u_{i0t} = \varepsilon_{i0t}$.
- Assume $\Phi_{A_j, t \geq T} = \Phi$ and $\Phi_{jt} = 0$ otherwise
- Assume $\Gamma_{A_1B} = \Gamma_{A_2B}$. $\Gamma_r = 0$ for all single good

III. Empirical Framework: Specification

Conditional choice probability:

$$P_{irt}(\mathbf{v}_i;\Theta) = \frac{\exp\left\{\sum_{j\in r}\left(\delta_j + \Phi_{jt} - \alpha p_{jt} + v_{ij}\right) + \Gamma_r\right\}}{1 + \sum_{k\in R_t/0}\exp\left\{\sum_{j\in k}\left(\delta_j + \Phi_{jt} - \alpha p_{jt} + v_{ij}\right) + \Gamma_k\right\}}$$

Estimation strategy: Simulated Maximum Likelihood

- Order-level market share $S_{rt}(\Theta) = \int P_{irt}\left(\mathbf{v}_i;\Theta\right) dF\left(\mathbf{v}_i;\Theta\right)$

$$\ln L(\mathbf{q},\Theta) = \sum_t \sum_r q_{rt} \ln S_{rt}(\Theta)$$

- Consumer-level choice sequence $y_{it} \in \mathbf{y}_i$ with unconditional probability $\tilde{P}(\mathbf{y}_i, \mathbf{v}_i;\Theta) = \prod_t P_{it,r=y_{it}}(\mathbf{v}_i;\Theta)$

$$\ln L(\mathbf{y},\Theta) = \sum_i \ln \int \tilde{P}(\mathbf{y}_i, \mathbf{v}_i;\Theta) dF\left(\mathbf{v}_i;\Theta\right)$$

IV. Monte Carlo Experiments

Data structure example

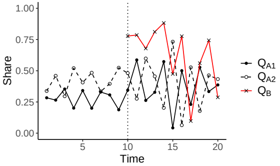
- Consumer-level: tracked sequence of choices (e.g. Nielsen Homescan)
- Order/Transaction-level: anonymous consumer panel (e.g., store scanner, online retail)

cust_id	tran_id	prod_id	tran_t	qty	price	prod_chr_1	...
cust 1	tran 1	prod 1	day 1	1	0.5	...	...
cust 1	tran 1	prod 2	day 1	2	1.5	...	...
cust 2	tran 2	prod 1	day 1	1	0.5	...	...
cust 2	tran 3	prod 2	day 2	3	2	...	...

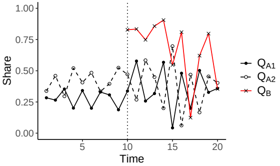
Monte Carlo experiments use *generated* data

IV. Monte Carlo Experiments

Exogenous pricing, 20 periods (introduce B at 10), $N = 2000$, $S = 700$



(a) $\Gamma = 0$ and $\Phi > 0$



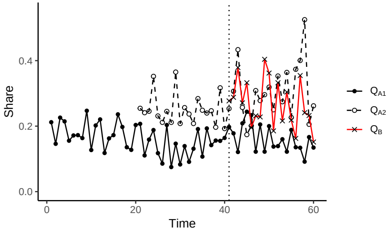
(b) $\Gamma > 0$ and $\Phi = 0$

	Models with Φ & Γ				Models without Γ				Models without Φ			
Parameters	Real	NL	Order	Consumer	Real	NL	Order	Consumer	Real	NL	Order	Consumer
δ_{A_1}	1	0.93^{***}	1.013^{***}	0.991^{***}	1	0.94^{***}	1.019^{***}	0.991^{***}	1	0.932^{***}	1.021^{***}	0.994^{***}
		(0.191)	(0.027)	(0.032)		(0.155)	(0.028)	(0.032)		(0.18)	(0.028)	(0.032)
δ_{A_2}	2	1.701^{***}	2.04^{***}	1.974^{***}	2	1.711^{***}	2.053^{***}	1.975^{***}	2	1.703^{***}	2.053^{***}	1.981^{***}
		(0.195)	(0.028)	(0.035)		(0.16)	(0.026)	(0.035)		(0.184)	(0.031)	(0.035)
δ_B	3	4.089^{***}	2.972^{***}	2.923^{***}	3	3.802^{***}	3.017^{***}	2.93^{***}	3	3.755^{***}	3.008^{***}	2.928^{***}
		(0.268)	(0.061)	(0.049)		(0.219)	(0.052)	(0.046)		(0.253)	(0.051)	(0.046)
α	-1	-0.806^{***}	-1.01^{***}	-0.999^{***}	-1	-0.806^{***}	-1.023^{***}	-0.998^{***}	-1	-0.806^{***}	-1.016^{***}	-1.002^{***}
		(0.095)	(0.022)	(0.011)		(0.078)	(0.022)	(0.011)		(0.09)	(0.028)	(0.011)
$\sigma_{A_1A_2}$	0.5	-	0.507^{***}	0.498^{***}	0.5	-	0.474^{***}	0.507^{***}	0.5	-	0.514^{***}	0.5^{***}
		(-)	(0.089)	(0.022)		(-)	(0.101)	(0.021)		(-)	(0.095)	(0.021)
σ_{A_1B}	0.7	-	0.633^{***}	0.673^{***}	0.7	-	0.609^{***}	0.685^{***}	0.7	-	0.635^{***}	0.682^{***}
		(-)	(0.11)	(0.027)		(-)	(0.076)	(0.024)		(-)	(0.102)	(0.026)
σ_{A_2B}	0.3	-	0.231	0.351^{***}	0.3	-	0.242^{**}	0.344^{***}	0.3	-	0.247	0.352^{***}
		(-)	(0.13)	(0.032)		(-)	(0.089)	(0.031)		(-)	(0.171)	(0.032)
Φ	0.3	1.629^{***}	0.256^{***}	0.31^{***}	0.5	1.446^{***}	0.456^{***}	0.502^{***}	0	1.257^{***}	-0.033	-0.007
		(0.158)	(0.073)	(0.035)		(0.129)	(0.036)	(0.033)		(0.149)	(0.074)	(0.032)
Γ	0.7	-	0.75^{***}	0.746^{***}	0	-	0.039	0.053	0.5	-	0.528^{***}	0.55^{***}
		(-)	(0.096)	(0.039)		(-)	(0.06)	(0.037)		(-)	(0.115)	(0.035)

IV. Monte Carlo Experiments

Endogenous pricing, 60 periods (introduce A_1 at 21, B at 41)

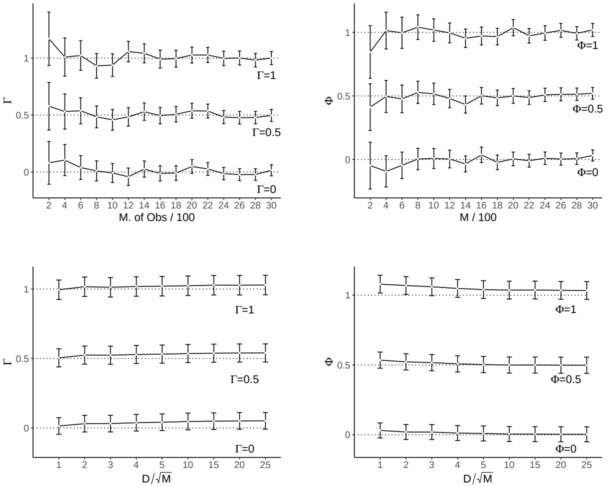
- $c_{A_1} = c_{A_2} = 1.5$, $c_B = 3$
- Firm knows the dist. of $\mathbf{v}$
- $\epsilon_{jt} \sim N(0, 0.3)$
- $\hat{c}_{A_1} = 1.487$, $\hat{c}_{A_2} = 1.493$



Parameters	Real	NL	Order	Consumer	2-Step
δ_{A_1}	1	0.514^{***}	0.979^{**}	0.881^{***}	1.023^{***}
		(0.118)	(0.409)	(0.114)	(0.057)
δ_{A_2}	2	1.065^{***}	1.872^{***}	1.846^{***}	1.963^{***}
		(0.127)	(0.459)	(0.129)	(0.063)
δ_B	3	2.174^{***}	2.611^{***}	2.778^{***}	3.034^{***}
		(0.222)	(0.911)	(0.203)	(0.086)
α	-1	-0.584^{***}	-0.948^{***}	-0.954^{***}	-1.003^{***}
		(0.047)	(0.168)	(0.043)	(0.021)
$\sigma_{A_1A_2}$	0.5	-	0.99	0.501^{***}	0.445^{***}
		(-)	(1.609)	(0.025)	(0.065)
σ_{A_1B}	0.7	-	0.323	0.67^{***}	0.684^{***}
		(-)	(1.136)	(0.027)	(0.209)
σ_{A_2B}	0.3	-	-0.014	0.33^{***}	0.427^{**}
		(-)	(1.109)	(0.03)	(0.216)
Φ	0.3	0.535^{***}	0.233	0.289^{***}	0.293^{***}
		(0.022)	(0.232)	(0.026)	(0.047)
Γ	0.7	-	1.024	0.755^{***}	0.677^{***}
		(-)	(0.829)	(0.038)	(0.164)

IV. Monte Carlo Experiments

SML is asymptotically equivalent to ML only if $S/\sqrt{N} \to \infty$ as N rising (Gourieroux and Monfort, 1993)



V. Merger Simulation

Suppose A_1 and B propose to merge

- independent pricing: p_{A_1}, p_B, p_{A_2}
- pure bundling: p_{A_1B}, p_{A_2}
- mixed bundling: $p_{A_1}, p_B, p_{A_1B}, p_{A_2}$
- mixed bundling w/o B: $p_{A_1}, p_{A_1B}, p_{A_2}$
- mixed bundling w/o A_1: p_B, p_{A_1B}, p_{A_2}

Two-stage game (Whinston, 1990)

- Pricing commitment
- Nash-Bertrand equilibrium

$$p_{rt} = S_{rt}(\mathbf{p}_t)\left[-\frac{\partial S_{rt}}{\partial p_{rt}}(\mathbf{p}_t)\right]^{-1} + c_r + \sum_{k \in \mathbb{R}_f / r} (p_{kt} - c_k) \frac{\partial S_{kt}}{\partial p_{kt}}(\mathbf{p}_t)\left[-\frac{\partial S_{rt}}{\partial p_{rt}}(\mathbf{p}_t)\right]^{-1}$$

- **Bundling harms competitors while benefiting the merged firm?**

V. Merger Simulation

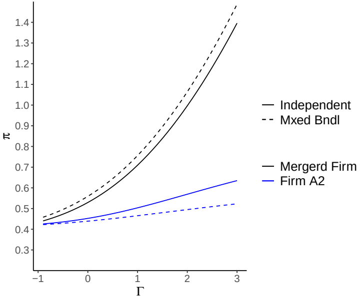
Demand parameters estimated by endogenous DGP

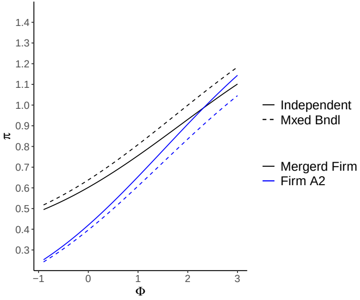
Variable	Pre-Merger	Independent	Pure Bndl	Mxed Bndl	Mxed Bndl w.o B	Mxed Bndl w.o A_1
p_{A_1}	2.835	2.75	-	3.282	3.112	-
p_{A_2}	3.148	3.13	3.024	3.088	2.995	3.117
p_B	4.621	4.573	-	4.885	-	4.822
$p_{A_1 B}$	-	-	6.302	6.61	6.393	6.536
π_{A_1} or π_m	0.221	0.652	0.416	0.694	0.503	0.626
π_{A_2}	0.492	0.472	0.383	0.441	0.361	0.467
π_B	0.423	-	-	-	-	-

- Independent: $0.221 + 0.423 < 0.652$ & all prices drop
- Profit incentive: MB $>$ Independent $>$ MB w/o B
- Foreclosure: MB w/o B $>$ PB $>$ MB $>$ Independent

V. Merger Simulation

Profit and foreclosure in MB with varying Γ and Φ





- (Positive) Γ and Φ: profit incentive $(+)$, foreclosure $(+)$
  - MB: $\Phi \uparrow \Rightarrow \; p_B \uparrow$, **profit from** A_2; Ind: $\Phi \uparrow \Rightarrow \; p_B$ unchanged
- Limited foreclose effect

V. Merger Simulation

Asymmetric variety effect caused by merger

- Complementary consumptions in the future (Kreps, 1979)

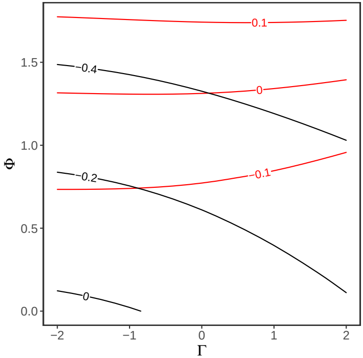
Assumptions for $\Phi_{A_2} = 0$

- **Incompatible** B (Katz and Shapiro, 1994)
- **Fully monopolized** B (Gentzkow, 2007)

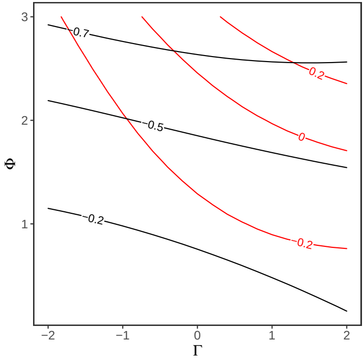
Profit incentive for MB w/o B or PB?

V. Merger Simulation

Asymmetric variety effect



(a) $\pi^{mbwob} - \pi^{mb}$



(b) $\pi^{pb} - \pi^{mb}$

- Upper-right region of the Γ, Φ space
- MB w/o $B >$ PB

VI. Concluding Remarks

- A flexible demand model to estimate complementarity and variety effect, using order-level data without supply assumption.
- Monte Carlo results show good accuracy and robustness.
- Our model can be a valuable tool for antitrust authorities handling cross-market strategies, eg., E-comm platform, ecosystem, product line,. . .
- In cases with tying or complement merger, the motivation and effect of these strategies are important empirical questions (Choi, 2008; Chen and Riordan, 2013; etc.).