Capturing Substitutions using Flexible Utility Function for Outside Good

Chul Kim
Baruch College, City University of New York

Adam Smith
University College London

Jaehwan Kim
Korea University Business School

Greg Allenby
Ohio State University

2019 Data Science Symposium, Tohoku University, December 7, 2019

- **Introduction**
- **Model**
- **Empirical analysis**
- **Concluding remarks**

Introduction

- ✓ **Pricing decision**
- ✓ **Substitution**
- ✓ **Multiple discreteness**

$$\left(x_1, x_2, \dots, x_J\right)$$
$$\left(p_1, p_2, \dots, p_J\right)$$

- ✓ **Outside good**

$$\left(x_1, x_2, \dots, x_J, {\color{red}z}\right)$$

- ✓ Price effect
- ✓ Elasticity

Introduction

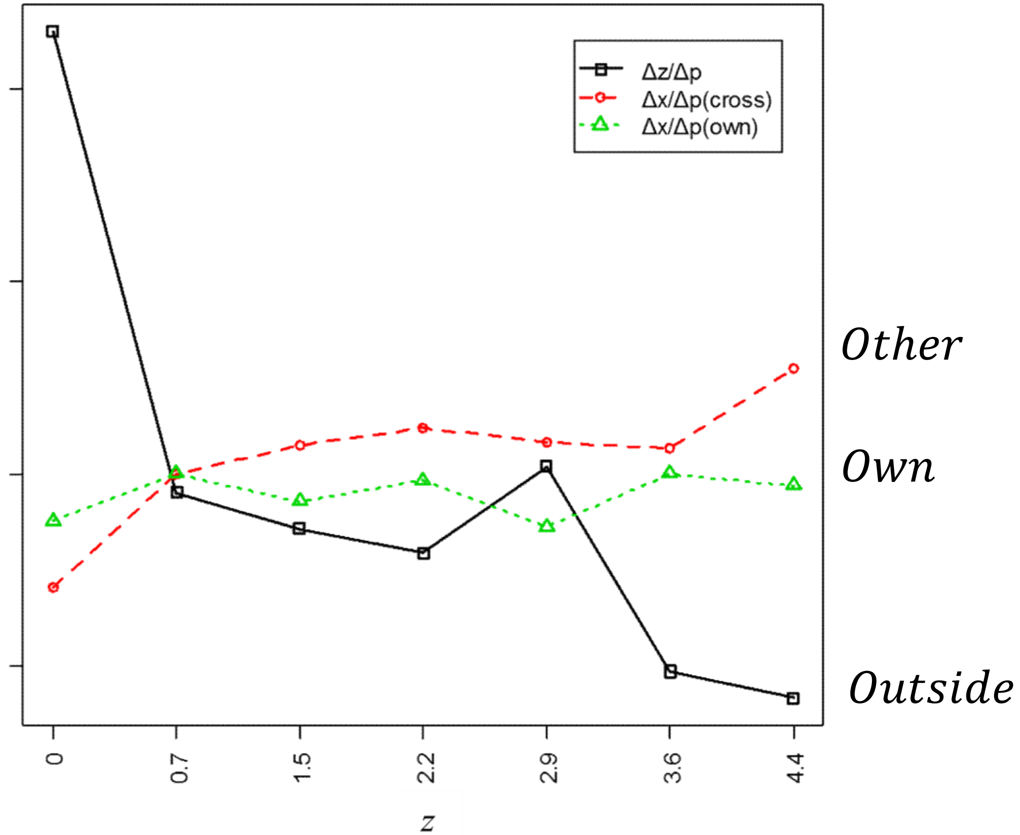
- **Substitution** among product offerings occurring through *price responses*

- In direct utility model, price change by one good affects demand on
  - other goods
  - outside good

- Often, the price effects are **dependent** on the level of **outside good**

Evidence from Data (Potato chip)

IRI panel dataset (Bronnenberg et al. 2008)



The demand changes in response to a price increase of a good k (p_k)

Introduction (cont'd)

- Demand (x, z) from constrained maximization:

$$\max. U(x,z) = \sum_{j=1}^{J} u_x(x_j) + u_z(z)$$

$$s.t. \sum_{j=1}^{J} p_j x_j + z \leq M$$

$$\frac{\partial z}{\partial p_k}, \quad \frac{\partial x_k}{\partial p_k}, \quad \frac{\partial x_j}{\partial p_k}$$

→ Substitution to be inferred accounting for the **interaction** with the level of **outside good**(z)

Introduction (cont'd)

- **How?** $u_z(z)$
  - Satiation ! (a.1)
  - Flexible enough (a.2)

- Literature in direct utility model is silent about this issue
  - $u_z(z)$ is mostly standard log, $ln(z)$
  - while (a.1) is satisfied, (a.2) is not.

  $\rightarrow$ Price effect$\left(\frac{\partial x_j}{\partial p_k}, \frac{\partial x_k}{\partial p_k}, \frac{\partial z}{\partial p_k}\right)$ not dependent on z

  $\rightarrow$ Elasticity estimates in question

Goal

- To propose a direct utility model with proper specification for $u_z(z)$ that allows for the dependence of price effect on the outside good
  - Flexible substitution
  - Proper elasticity-based pricing

- Introduction
- Model
- Empirical Analysis
- Conclusion

Literature on DUM: *structure*

- Direct utility function

$$U(\boldsymbol{x}_t, z_t) = \sum_{j=1}^{J} u_x(x_{jt}) + u_z(z_t)$$

$$s.t. \sum_{j=1}^{J} p_{jt}x_{jt} + z_t \leq M_t$$

- K-T condition

$$\varepsilon_{jt} = g_{jt} \text{ if } x_{jt} > 0$$

$$\varepsilon_{jt} < g_{jt} \text{ if } x_{jt} = 0$$

$$g_{jt} = -\ln\left(\frac{\partial u_x(x_{jt})}{\partial x_{jt}}\right) + \ln\left(p_{jt} \cdot \frac{\partial u_z(z_t)}{\partial z_t}\right)$$

Literature on DUM: *inference*

- Likelihood

$$L(\boldsymbol{x}_t, z_t) = \prod_{\{x_{jt}>0\}} f_\varepsilon(g_{jt}) \prod_{\{x_{jt}=0\}} F_\varepsilon(g_{jt}) \cdot |J_t|$$

where

$$J_{t(i,j)} = \frac{\partial g_{it}}{\partial x_{jt}}$$

→ Note : $|J_t| = \left| J_t\left(\frac{u''(z_t)}{u'(z_t)}\right) \right|$

Substitution pattern

$$U(\boldsymbol{x}, z) = \sum_j \frac{\psi_j}{\gamma_j} \ln\left(\gamma_j x_j + 1\right) + u(z)$$

- $m_{zk}(z) \equiv \frac{\partial z}{\partial p_k} = \frac{1}{\gamma_k} \cdot \frac{1}{1+h(z)\sum_{l\in R} \psi_l/\gamma_l}$
- $m_{jk}(z) \equiv \frac{\partial x_j}{\partial p_k} = \frac{1}{p_j \gamma_k} \cdot \frac{h(z)\psi_l/\gamma_l}{1+h(z)\sum_{l\in R} \psi_l/\gamma_l}$
- $m_{kk}(z) \equiv \frac{\partial x_k}{\partial p_k} = -\frac{1}{p_k \gamma_k} \cdot \frac{1+h(z)\sum_{l\in R, l\neq k} \psi_l/\gamma_l}{1+h(z)\sum_{l\in R} \psi_l/\gamma_l} - \frac{x_k}{p_k}$

where $h(z) = \frac{-u''(z)}{u'(z)^2}$

Condition required $h'(z) > 0$

$$\left.\begin{array}{l} \dfrac{\partial m_{zk}(z)}{\partial z} < 0 \\[2em] \dfrac{\partial m_{jk}(z)}{\partial z} > 0 \\[2em] \dfrac{\partial m_{kk}(z)}{\partial z} > 0 \end{array}\right\} \quad iff \quad h'(z) > 0$$

Summary : $U_z(z)$ in literature

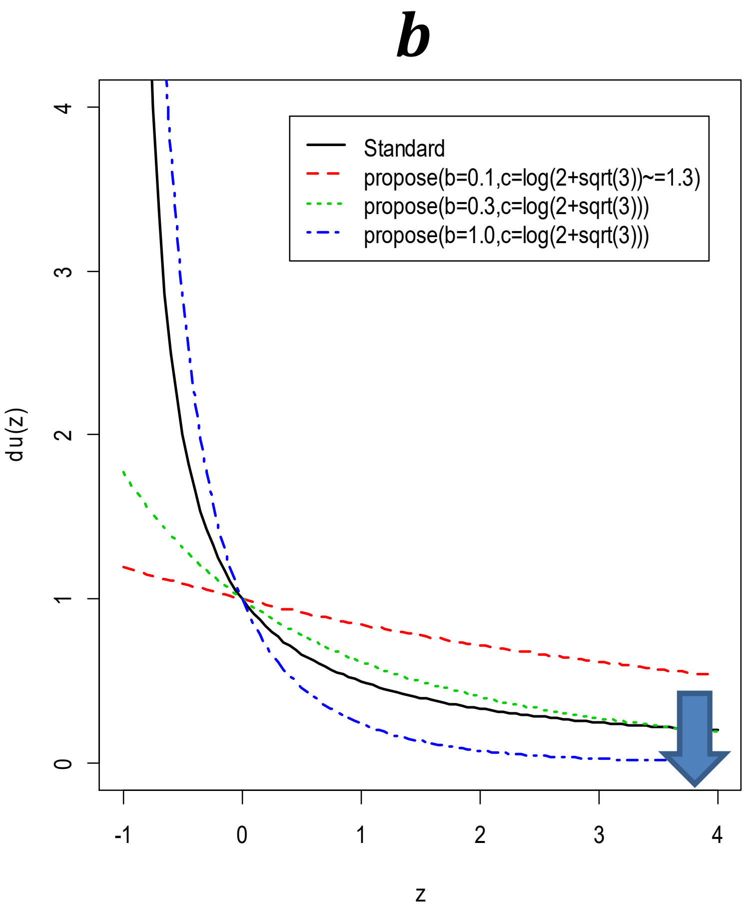
U(z)	Satiation	$h(z)$	
Log function[1]	*yes*	**Constant**	$h'(z) = 1$
This study	*yes*	Flexible	$h'(z) > 0$ guaranteed

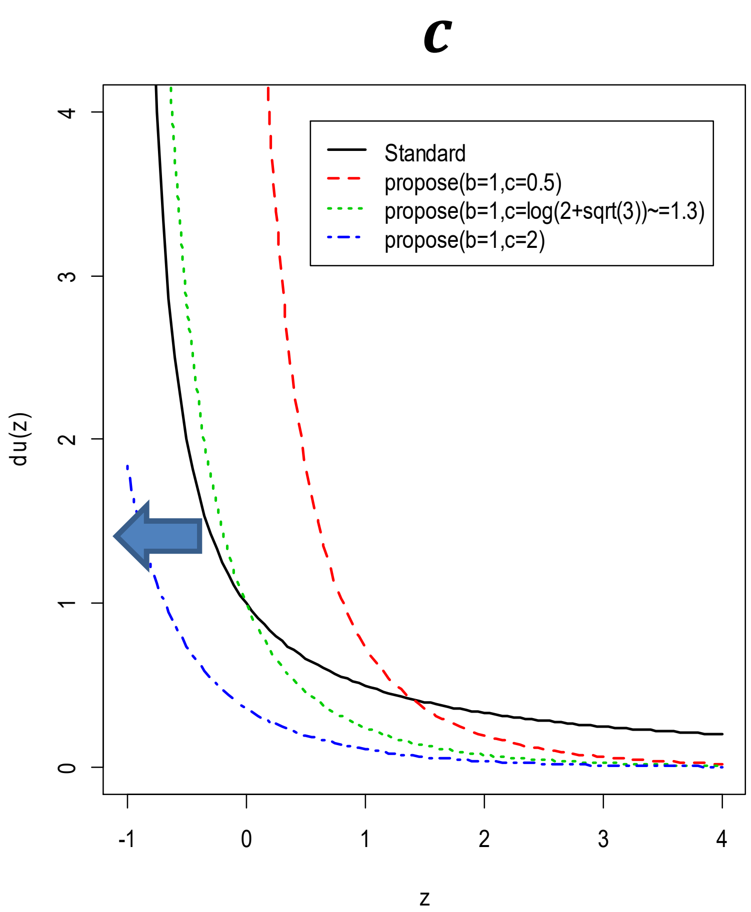
[1]Satomura et al(2011), Hasegawa et al(2012), Lee et al(2013), Howell et al(2016), Kang et al(2016), Kim et al(2017)

cf. Kim et al(2002, 2007), Bhat(2008), Luo et al(2012),

$U_z(z)$

$$u(z) = \frac{1}{b} \cdot \frac{1 + exp(bz + c)}{1 - exp(bz + c)}$$





$U_z(z)$

$$u(z) = \frac{1}{b} \cdot \frac{1 + exp(bz + c)}{1 - exp(bz + c)}$$

$$\rightarrow \quad u'(z) = \frac{2 \cdot exp(bz + c)}{\left(1 - exp(bz + c)\right)^2}$$

$$\rightarrow \quad h(z) = b \cdot sinh(bz + c)$$

$$h'(z) = b^2 \cdot cosh(bz + c) \; > 0$$

- Introduction
- Model
- **Empirical Analysis**
- Concluding remarks

Data

	LAYS	WAVY LAYS	RUFFLES	OLD DUTCH
Average Price ($)	2.21	2.38	2.67	0.87
Total Quantity	4,598	2,447	1,072	755
Total Incidence	3,246	1,869	839	591
1 unit	2,147	1,397	626	446
2 units	899	390	196	131
more than 2 units	200	82	17	14

including UPCs with m/s > 1% and size being medium (12 – 16 oz).

- 476 households with at least 3 purchase occasions
- Total 5,941 purchase occasions,
- On average, 1.49 units purchased per trip.

Estimation

$$U(\mathbf{x}, z) = \sum_j \frac{\psi_{h,j}}{\gamma_{h,j}} \ln\left(\gamma_{h,j} x_j + 1\right) + \frac{1}{b_h} \cdot \frac{1 + exp(b_h z + c_h)}{1 - exp(b_h z + c_h)}$$

- Heterogeneity

$$\boldsymbol{\theta}_h = \left(\tilde{\psi}_{1,h}, \dots, \tilde{\psi}_{4,h}, \tilde{\gamma}_{1,h}, \dots, \tilde{\gamma}_{4,h}\right)' \sim N(\bar{\boldsymbol{\theta}}, \mathbf{V}_\theta)$$

$$\boldsymbol{\alpha}_h = \left(\tilde{b}_h, \tilde{c}_h\right)' \sim N(\bar{\boldsymbol{\alpha}}, \mathbf{V}_\alpha)$$

where

$$\tilde{\psi}_h = ln(\psi_h),\ \tilde{\gamma}_h = ln(\gamma_h),\ \tilde{b}_h = ln(b_h),\ \tilde{c}_h = ln(c_h)$$

Estimation

$$U(\mathbf{x}, z) = \sum_j \frac{\psi_{h,j}}{\gamma_{h,j}} \ln\left(\gamma_{h,j} x_j + 1\right) + \frac{1}{b_h} \cdot \frac{1 + exp(b_h z + c_h)}{1 - exp(b_h z + c_h)}$$

- MCMC

$$\{\boldsymbol{\theta}_h, h = 1, .., H\} \middle| \ \bar{\boldsymbol{\theta}}, \mathbf{V}_\theta, \{\mathbf{x}_h, \mathbf{p}_h\}$$

$$\bar{\boldsymbol{\theta}} \ \middle| \{\boldsymbol{\theta}_h, h = 1, .., H\}, \mathbf{V}_\theta, \{\mathbf{x}_h, \mathbf{p}_h\}$$

$$\mathbf{V}_\theta \ \middle| \{\boldsymbol{\theta}_h, h = 1, .., H\}, \bar{\boldsymbol{\theta}}, \{\mathbf{x}_h, \mathbf{p}_h\}$$

$$\{\boldsymbol{\alpha}_h, h = 1, .., H\} \middle| \ \bar{\boldsymbol{\alpha}}, \mathbf{V}_\alpha, \{\mathbf{x}_h, \mathbf{p}_h\}$$

$$\bar{\boldsymbol{\alpha}} \middle| \{\boldsymbol{\alpha}_h, h = 1, .., H\}, \mathbf{V}_\alpha, \{\mathbf{x}_h, \mathbf{p}_h\}$$

$$\mathbf{V}_\alpha \middle| \{\boldsymbol{\alpha}_h, h = 1, .., H\}, \bar{\boldsymbol{\alpha}}, \{\mathbf{x}_h, \mathbf{p}_h\}$$

Model Fit

Model Fit		Benchmark Model	Proposed Model
In-sample	LMD	-11,581	-11,015
	SSE	15,645	7,237
Predictive	SSE	17,020	15,062

$$U(\mathbf{x}, z) = \sum_j \frac{\psi_{h,j}}{\gamma_{h,j}} \ln\left(\gamma_{h,j} x_j + 1\right) + \psi_z \cdot log(z)$$

$$U(\mathbf{x}, z) = \sum_j \frac{\psi_{h,j}}{\gamma_{h,j}} \ln\left(\gamma_{h,j} x_j + 1\right) + \frac{1}{b_h} \cdot \frac{1 + exp(b_h z + c_h)}{1 - exp(b_h z + c_h)}$$

Posterior distribution

Parameter	Benchmark Model			Proposed Model		
	Mean	*Std.*	Hetero	Mean	*Std.*	Hetero
$\bar{\tilde{\psi}}_1$	set to 0	-		set to 0	-	
$\bar{\tilde{\psi}}_2$	-0.65	*0.09*	2.45	-0.56	*0.07*	1.54
$\bar{\tilde{\psi}}_3$	-1.47	*0.09*	2.24	-1.36	*0.08*	1.28
$\bar{\tilde{\psi}}_4$	-3.10	*0.12*	4.00	-3.00	*0.10*	2.64
$\bar{\tilde{\gamma}}_1$	-0.71	*0.07*	1.04	-1.10	*0.06*	0.46
$\bar{\tilde{\gamma}}_2$	-1.48	*0.11*	1.11	-1.71	*0.14*	1.26
$\bar{\tilde{\gamma}}_3$	-4.05	*0.28*	2.58	-4.46	*0.46*	2.20
$\bar{\tilde{\gamma}}_4$	-1.50	*0.15*	1.11	-1.66	*0.13*	1.18
$\bar{\tilde{\psi}}_z$	0.53	*0.05*	0.90			
$\bar{\tilde{b}}$				-1.21	*0.04*	0.50
$\bar{\tilde{c}}$				-0.12	*0.02*	0.11

Price Elasticity $(\eta_{j \leftarrow k})$

Benchmark model

$$u(z) = \psi_z \cdot log(z)$$

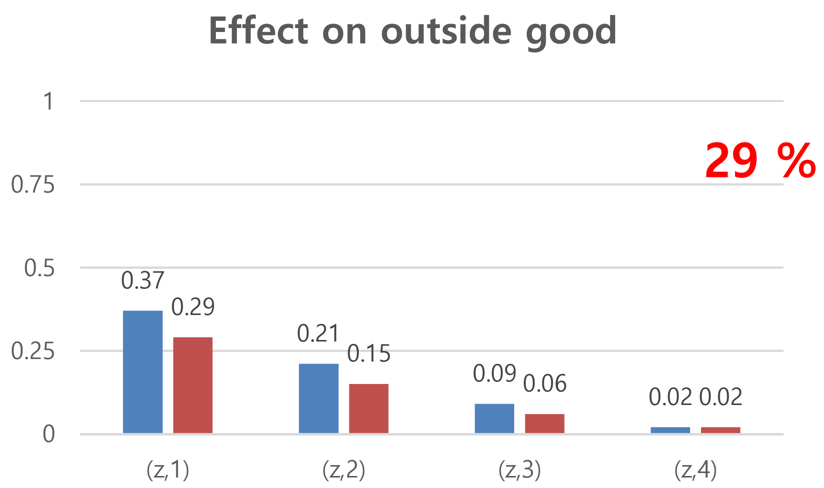
	1	2	3	4	z
1	-1.20	.30	.10	.02	.37
2	.61	-1.63	.14	.02	.21
3	.48	.41	-2.01	.00	.09
4	.26	.18	.06	-1.53	.02

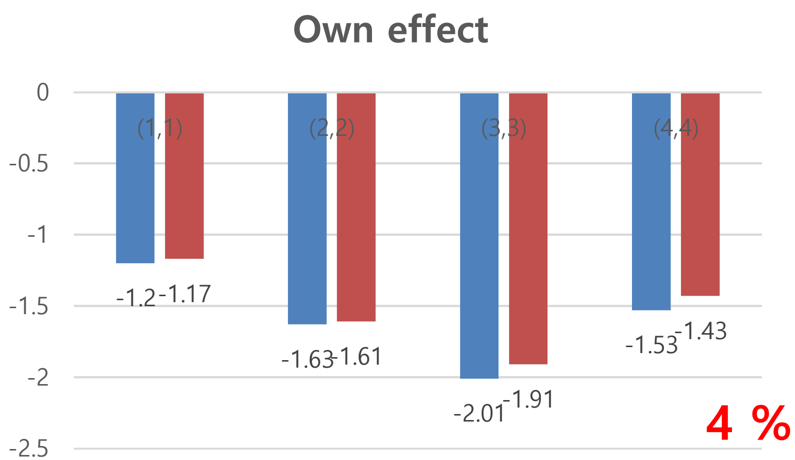
Proposed model

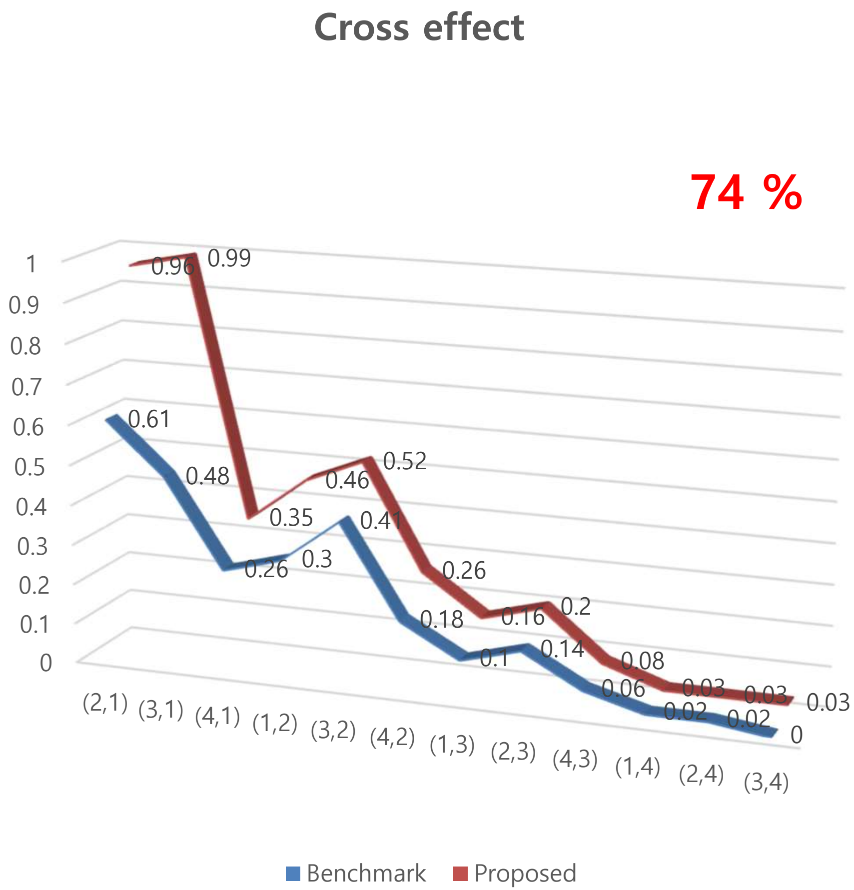
$$u(z) = \frac{1}{b} \cdot \frac{1 + exp(bz + c)}{1 - exp(bz + c)}$$

	1	2	3	4	z
1	-1.17	.46	.16	.03	.29
2	.96	-1.61	.20	.03	.15
3	.99	.52	-1.91	.03	.06
4	.35	.26	.08	-1.43	.02

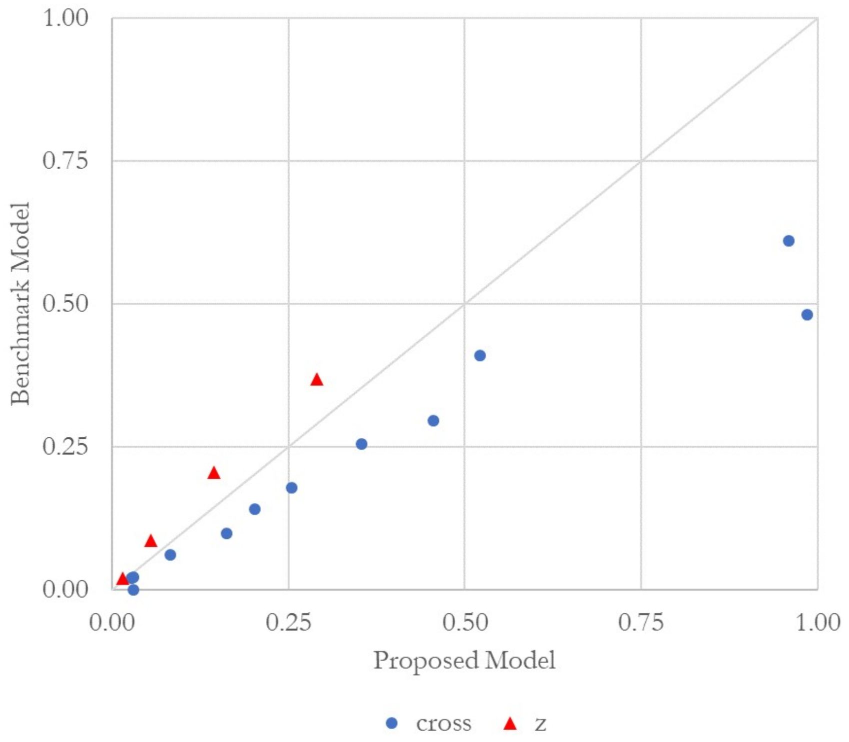
Price Elasticity







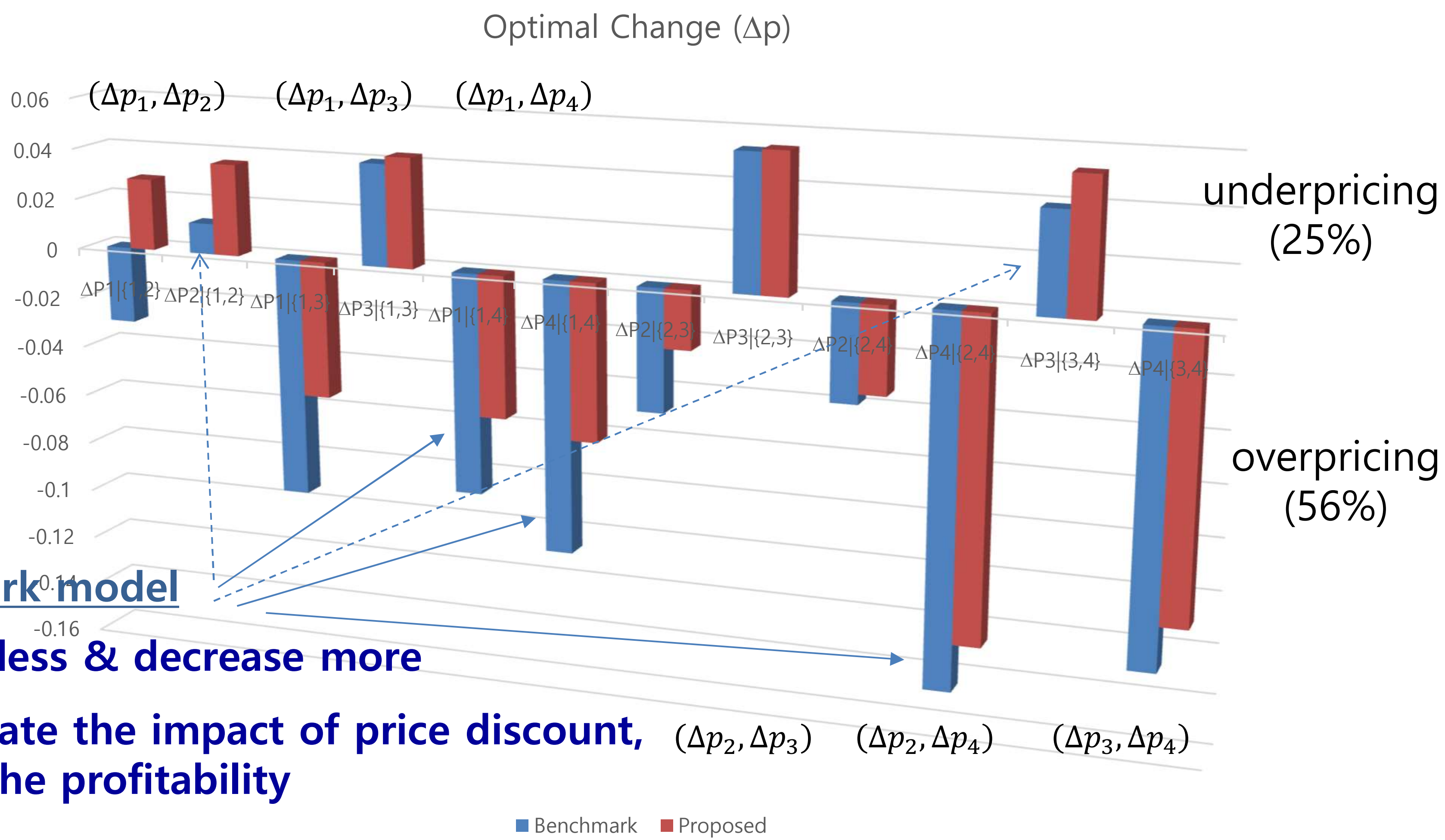
Price Elasticity



Counterfactual : Line pricing

- Changes in price**s** for **multiple** products : manufacturers or retailers to price for a targeted 'group' of offerings (revenue maximization)
  - 4 product offerings :$(x_1, x_2, x_3, x_4), (p_1, p_2, p_3, p_4)$
  - Changing prices for **two** of them simultaneously
  - Optimal change: $\left(\Delta p_i, \Delta p_j\right) \middle| \left(p_i, p_j, \check{\mathbf{p}}_{k \neq i,j}\right)$

Optimal price change



Benchmark model

Increase less & decrease more

Overly state the impact of price discount, hurting the profitability

Counterfactual : Line pricing

- Whether/How much to change (Δp)
  - increase p : own effect (-) vs cross effect (+ vs -)
  - decrease p : own effect (+) vs cross effect (- vs +)

Proposed < Benchmark

Proposed > Benchmark

→ dependent on ***relative*** difference in cross effect between goods 'in' and 'outside' the product line

- **Introduction**
- **Model**
- **Empirical Analysis**
- **Concluding remarks**

Concluding remarks

- **Demand response to price change being dependent on the level of outside good**
- **An approach to capture the flexible substitution via new utility function of outside good**
  - more than satiation (flexibility)
- **Multi-product pricing**
  - Proper cross-elasticity estimates

Concluding remarks

- **More on consumer's constrained u.max decision**
  - Utility structure : preference, satiation
  - Constraint structure: single to multi-, linear to nonlinear
  - Incorporating 'partition': consumer preferences within each group can be described independently of consumption levels in other groups (for larger assortment)

감사합니다

Thank you

ありがとうございました