(Un)common Properties of Discrete Choice Models with Multi-purchasing*

Ce Matthew Shi†
Department of Economics
Chinese University of Hong Kong

Xinyi Xu‡
Department of Economics
Sun Yat-sen University

December, 2019
(*Prelimineary and Incomplete*)

Abstract While traditional discrete choice models assume that economic agents choose only one option among the alternatives, more recently, discrete choice models that relax the single-unit assumption have been developed and used widely to analyze buyers' multi-purchasing behavior, especially in empirical contexts. We present and analyze two modeling frameworks of multi-purchasing discrete choice that contain as special cases the traditional binary and multinomial choice models and important empirical models, such as Gentzkow (2007) and Fan (2013). We analyze the the models with unspecified, general distribution of randomness and focus on a set of fundamental theoretical properties, including monotonicity, cross product effects and their relationship to traditional single-choice models. We discuss in detail the implications for supply-side modeling and identification and estimation in empirical contexts.

*We thank Simon Anderson, Arthur Lewbel, Patrick Rey, Matt Shum, Alex White, and Jidong Zhou for helpful comments and discussions as well as the participants of 2019 CUHK IO Workshop and International Conference on Industrial Economics (Hangzhou). The draft contains preliminary analysis and is incomplete. Please do not cite without permission.

† *Corresponding author.* Address: 910 Esther Lee Building, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong SAR. E-mail: *shice@cuhk.edu.hk*.

‡ Address: J. T. Wu Hall (Lingnan Administration Centre), 135 West Xingang Road, Guangzhou, China. E-mail: *xuxinyi3@mail.sysu.edu.cn*.

1 Introduction

Traditional discrete choice models assume that economic agents choose only one option among the alternatives. More recently, models of discrete choice that relax the single-unit-purchase assumption have been developed and used widely to analyze consumer multiple-purchasing behavior and other important decisions of this kind, especially in empirical contents. In this paper, we present and analyze two general frameworks of random-utility, multi-purchasing discrete choice that contain as special cases the traditional binary and multinomial discrete choice models and prominent empirical models, such as Gentzkow (2007) and Fan (2013). We study the models for unspecified, general distribution of randomness and focus on a set of theoretical properties, such as monotonicity and negative cross effects for gross substitutes, that have important economic implications, are known to hold in simpler models but often only assumed in more complex applications. In particular, we show some of the properties to hold in the general cases considered while some foundational properties, such as gross substitutes, may not hold even under common distributional assumptions, such as Normal or Type-I Extreme Value (Gumbel) distribution. We discuss in detail how those properties – and their failure – matter for supply-side modeling (e.g., existence of pricing equilibrium and product bundling) and identification and estimation of structural models in empirical contexts.

In empirical applications, owning to increasing availability of consumer-level micro data, it is often observed in data that buyers make multi-unit and/or multi-brand purchases, especially in marketing contexts (e.g. Dube, 2004); thus, the single-purchase assumption becomes inadequate as it is directly rejected by patterns in data. This is in contrast with the classic demand-estimation approach á la Berry *et al.* (1995) that mainly uses aggregate product-level data, such as market shares. Structural empirical models that allow decision makers to choose multiple options at once have been widely used in a variety of applications, including the studies of technology adoption (Augereau *et al.*, 2006), household choice of elderly care (Hiedemann *et al.*, 2013), illegal drugs (Deza, 2015) and in agriculture (Perry *et al.*, 2016) – to name a few. Another important area is the study of platform markets, such as advertising and media. In media markets, assuming that consumers only adhere to a single platform, or single-home, implies that the platform has monopoly power over the consumer attention in the advertising market (e.g. Ambrus *et al.*, 2017; Anderson *et al.*, 2017); that is, it takes away any interesting platform competition *for* advertisers in structural modeling. Recent empirical works, such as Fan (2013), Gentzkow *et al.* (2014), and Shi (2017), all attempt to resolve this inconsistency. Nonetheless, properties of many important empirical frameworks are not very well understood by researchers. We in this paper examine three foundational properties: monotonicity, negative cross effects for substitutes and links to traditional binary and multinomial models. These properties are critical for both theoretical and empirical reasons. In theory, monotonicity and products being gross substitutes are crucial for existence of pricing equilibrium in Bertrand-Nash games. In empirics, they are often sufficient for demand invertibility and needed for the contraction mapping routine in estimation (e.g. Berry *et al.*, 1995, 2013).

We first consider a "bundle" model, which is a generalization of the empirical model in Gentzkow (2007) that studies a three-good case and assumes Type-I extreme value distribution. In this model, consumers' multi-purchasing choices are treated as choices over bundles of product. The value of a bundle is the sum of the stand-alone value of each product and a *bundle-specific* taste shock minus some utility loss due to product substitution in joint consumption. Subsequent empirical applications include – but not limited to – Gentzkow *et al.* (2014); Song and Chintagunta (2006) and Liu *et al.* (2010) in marketing; Pereira *et al.* (2013); Deza (2015) that incorporates consumer inertia; and Perry *et al.* (2016) that assumes normal distribution. We show that while this framework has a negative utility term to capture consumption substitution, products (instead of bundles) may not be gross substitutes. We present counter examples of Normal distribution with two goods and Type-I extreme value with three goods. We discuss in great detail how the properties can hold in more restricted cases of Type-I extreme value distribution with specific assumptions on parameter values. We then consider a "Hendel"-type model, which we name after Fan (2013)'s application of Hendel (1999)'s pioneering framework. In the framework of Hendel (1999), a product is being considered to meet the consumption need on M distinct and independent occasions. On each consumption occasion, only one option is chosen over the alternatives.

But at the aggregate level, multiple purchases can be made. In Fan (2013)'s innovative application, each occasion is the event that a product takes a particular position in a decision-maker's ranking of n products. Shi (2017) also uses this model. We show that despite of the apparent technical difficulties, all the properties hold for general distribution of randomness. We aim to provide detailed and useful guidance for empirical applications of the two important frameworks.

The paper is organized as follows. In the next section, we describe two models of multi-purchasing discrete choice with generally distributed randomness. In section 3, we closely examine the main theoretical properties for each framework. Section 4 focuses on the specific case of Type-I extreme value distributions. In Section 5, we discuss in detail the implications for structural modeling and empirical implementation. Section 6 concludes.

2 Two Models of Multi-Purchasing Discrete Choice

In this section, we describe two modeling frameworks of multi-purchasing discrete choice (mDCM) with random taste shocks from some unspecified, general distribution. In both frameworks, we consider n underlying products, indexed by $i \in \mathscr{I}$, and a outside option called 0. Each product has some non-random, stand-alone consumption value, δ_i, common to all consumers. In IO and marketing applications, δ_i represents the overall perceived product quality and may contain product price. In each model, there is a distinct way to incorporate an idiosyncratic, consumer-product match value and consumers' decision rule, which leads to some expression for the probability that a product i is chosen. We refer to this probability, $\mathbb{P}(i)$, colloquially as the *demand* for i. We are interested in analyzing the theoretical properties of the demand functions that arise from the following two modeling frameworks.

Bundle model

We first consider a "bundle" model, which is a generalization of the empirical model in Gentzkow (2007).[1] In this model, consumers' multi-purchasing choices are treated as choices over bundles of product. For n underlying products, consumer face 2^n bundle choices including the outside option, which cannot be bundled together with other options. Each bundle choice is indexed by $b \in \mathscr{B}$, where $\mathscr{B}$ is the power set of $\mathscr{I}$. An element of b is denoted $b^{(r)}$. For example, when $\mathscr{I} = \{A, B\}$, $n = 2$, and $\mathscr{B} = \{0, \{A\}, \{B\}, \{A, B\}\}$.

A consumer c's indirect utility from choosing bundle b, u_{cb}, consists of bundle b's quality, v_b, and a bundle-specific idiosyncratic term, ϵ_{cb}. In particular, suppressing the subscript c in what follows, we write

$$u_b = \underbrace{\sum_r \delta_{b^{(r)}} - \Gamma_b}_{v_b} + \epsilon_b, \tag{1}$$

where the parameter Γ_b captures any difference between the value of a bundle and the sum of the stand-alone value of its elements with $\Gamma_b = 0$ if $|b| \leq 1$, where $|b|$ is the cardinality of bundle b (i.e., b is a singleton bundle or the outside option). It is further assumed that $\Gamma_{b'} \geq \Gamma_b \geq 0$ for any two distinct bundles b and b' with $|b'| > |b| > 1$. In the literature, Γ_b is used to parsimoniously capture product *consumption substitutibility*. While v_b is common to all consumers, consumers' preference heterogeneity is captured by ϵ_{cb}, which is bundle-specific and independent and identically distributed (i.i.d.) across bundles with some unspecified, general distribution function F. For the outside option, $u_0 = \delta_0 + \epsilon_0$, where δ_0 is conventionally normalized to zero without loss of generality. In many empirical applications, ϵ_b is assumed to follow the Type I extreme value distribution, which case we specifically study in Section 4.

Consider the example for $i \in \{A, B\}$. Under this formulation, $u_A = \delta_A + \epsilon_A$, $u_B = \delta_B + \epsilon_B$, and $u_{AB} = \delta_A + \delta_B - \Gamma_{AB} + \epsilon_{AB}$. Note that although $v_{AB} < v_A + v_B$, or v_b is submodular, u_b is not necessarily submodular due to the bundle-specific, i.i.d. shock. While the bundle-specific randomness makes it easier for probability expression, as we show in the next section, this unique

[1] In Gentzkow (2007), there is also an illustrative theoretical model, which is subtly different from the empirical model for structural estimation, as we notice and discuss in later sections.

modeling feature nonetheless gives rise to a number of uncommon theoretical properties, which are often not present in the special case of the Type I extreme value distribution.

Observing the conditional utility of each choice, consumers choose a bundle of products that yields the highest utility among the alternatives. Therefore, for any party who does not directly observe the random match value, the probability of *a bundle b being chosen*, $\mathbb{P}(b)$, is given by

$$\mathbb{P}(b) = Pr(u_b \geq u_{b'}, \forall b' \in \mathscr{B}). \tag{2}$$

In many important contexts, such as oligopoly analysis, researchers care more about i's product demand or the probability of a product i – instead of a bundle – being chosen. In this model, we write

$$\mathbb{P}(i) = \sum_{b \in \mathscr{B}} \mathbb{1}(b \ni i) \cdot \mathbb{P}(b), \tag{3}$$

where $\mathbb{1}(b \ni i)$ is an indicator function that equals to 1 if a bundle b contains i and 0 otherwise; that is, the probability of a product being ever purchased is the sum of the probabilities of choosing all bundles that contain that product.

Using the semi-parametric form of the utility function and the i.i.d. shock, we have specifically

$$\begin{aligned}
\mathbb{P}(i) &= \sum_{b \in \mathscr{B}} \mathbb{1}(b \ni i) \cdot Pr(v_b - v_{b'} + \epsilon_b \geq \epsilon_{b'}, \forall b' \in \mathscr{B}) \\
&= \sum_{b \in \mathscr{B}} \mathbb{1}(b \ni i) \cdot \int \Big(\prod_{\substack{b' \in \mathscr{B} \\ b' \neq b}} F(v_b - v_{b'} + \epsilon_b) \Big) dF(\epsilon_b).
\end{aligned} \tag{4}$$

"Hendel" model

Second, we consider a Hendel-Fan model, which we name after Fan (2013)'s application of Hendel (1999)'s pioneering framework. In the framework of Hendel (1999), a product is being considered to meet the consumption need on M distinct and independent occasions. On each consumption occasion, only one option is chosen over the alternatives. But at the aggregate level, multiple purchases can be made. Let $\mathbb{P}^m(i)$ denote the probability that i is chosen on occasion m, and given that the occasions are independent, the probability of i being chosen is

$$\mathbb{P}(i) = \sum_{m=1}^{M} \mathbb{P}^m(i). \tag{5}$$

In Fan (2013)'s innovative application, each occasion is the event that a product takes a particular position in a decision-maker's ranking of n products. A natural way to think about this is that consumers make a series of sequential purchase decisions. For example, if a newspaper reader can purchase up to $M = 2$ newspapers, occasion $m = 1$ refers to the event that the reader picks her *first best* choice for consumption among the products, and occasion $m = 2$ refers to when the reader picks her *second best* choice for consumption. In this case, the stand-alone value of a product is (with the subscript c suppressed thereafter)

$$u_i = \delta_i + \varepsilon_i, \tag{6}$$

where ε_i is product-specific. The utility of the outside option is $u_0 = \delta_0 + \epsilon_0$, where δ_0 is normalized to zero.

To capture diminishing utilities from multiple purchases, the stand-alone utility of any product i will decrease by Γ_m with $\Gamma_m = 0$ for $m \leq 1$, and $m' > m > 1, \Gamma_{m'} \geq \Gamma_m \geq 0$.

Therefore, the probability of i being chosen as the m^{th} best choice is

$$\mathbb{P}^m(i) = Pr\Big(\min_{j \in \{j^{(1)}, \cdots, j^{(m-1)}\} \subset \mathscr{I}} \{u_j\} \geq u_i \geq \max_{\substack{k \neq i,j \\ k \in \mathscr{I}}} \{u_k\}; u_i - \Gamma(m) \geq u_0 \Big). \tag{7}$$

For instance, when $n = 1$, it follows automatically that $m = 1$, so

$$\mathbb{P}(i) = \mathbb{P}^1(i) = Pr\left(u_i \geq \max_{\substack{k \neq i \\ k \in \mathscr{I}}} \{u_k, u_0\} \right),$$

which is the choice probability expression in a traditional multinomial discrete choice model.

In particular, combining (5) and (7) and using the i.i.d. shock, we have

$$
\begin{aligned}
\mathbb{P}(i) = & \int \Big\{ F(\delta_i + \epsilon_i) \cdot \prod_{\substack{j \in \mathscr{I} \\ j \neq i}} F(\delta_i + \epsilon_i - \delta_j) \Big\} dF(\epsilon_i) \\
& + \sum_{\substack{j \in \{j^{(1)}, \cdots, j^{(m-1)}\} \subset \mathscr{I} \\ j \neq i}} \int \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot \prod_{k \neq j,i} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma_2) \Big\} dF(\epsilon_i) \\
& + \sum_{j,k \neq i} \int \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot [1 - F(\delta_i + \epsilon_i - \delta_k)] \cdot \prod_{l \neq i,j,k} F(\delta_i + \epsilon_i - \delta_l) \cdot F(\delta_i + \epsilon_i - \Gamma_3) \Big\} dF(\epsilon) \\
& \vdots \\
& + \int \Big\{ \prod_{j \neq i} [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot F(\delta_i + \epsilon_i - \Gamma_n) \Big\} dF(\epsilon_i). \qquad (8)
\end{aligned}
$$

3 The Common and Uncommon Properties

For the demand system that arises under each of the two modeling frameworks described above, we examine whether the following common properties hold. A detailed discussion of their implications is in Section?

P1 (***Monotonicity***) $\dfrac{\partial \mathbb{P}(i)}{\partial \delta_i} > 0$.

P2 (***Gross substitutes***) When $\Gamma_N > 0$ for $N > 1$, $\dfrac{\partial \mathbb{P}(i)}{\partial \delta_j} < 0$, for any $j \neq i$ and $i, j \in \mathscr{I}$.

P3 (***Limiting cases***) **(a)** When $\Gamma_N \to \infty$ for all $N > 1$, mDCM reduces to DCM;
(b) When $\Gamma_N = 0$ for all N, mDCM reduces to a binary choice model with $\dfrac{\partial \mathbb{P}(i)}{\partial \delta_j} = 0$ for any $j \neq i$ and $i, j \in \mathscr{I}$.

Proposition 1 *In a bundle model, P1 and P3a are always satisfied while P2 and P3b are not.*

Lemma 1 *For a bundle model with* $n = 2$ *and* $\Gamma(N) = 0$ *for all* N*, (P3b) holds only if*

$$\int_{-\infty}^{\infty} 2F^2(\delta_i + x) \cdot F(x) dF(x) = \int_{-\infty}^{\infty} F(\delta_i + x) dF(x). \qquad (9)$$

That property holds with Gumbel. Does it hold for normal? Below there is also a counter-example with exponential.

For monotonicity, first notice

$$\begin{aligned}\mathbb{P}(i) &= \sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\cdot\int\Big(\prod_{b'\in\mathscr{B}} F(v_b - v_{b'} + \epsilon_b)\Big)dF(\epsilon_b)\\ &= \sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\cdot\int\Big(\prod_{b'\ni i} F(v_b - v_{b'} + \epsilon_b)\cdot\prod_{b'\not\ni i} F(v_b - v_{b'} + \epsilon_b)\Big)dF(\epsilon_b); \qquad (10)\end{aligned}$$

In particular, given $b \ni i$, the term $\prod_{b'\ni i} F(v_b - v_{b'} + \epsilon_b)$ does not depend on δ_i; however, for each $b' \not\ni i$, $\frac{\partial v_{b'}}{\partial \delta_i} = 0$, so $\frac{\partial F(v_b - v_{b'} + \epsilon_b)}{\partial \delta_i} = f(v_b - v_{b'} + \epsilon_b) > 0$. Therefore, it can be shown that $\frac{\partial \mathbb{P}(i)}{\partial \delta_i} > 0$.

Intuitively, for any bundle b that contains i, its choice probability increases as i's value, δ_i, increases. Based on the above formulation, this is because when δ_i increases, the probability that b is chosen over some other bundle that also contains i remains the same while the probability it is chosen over some bundle that does not contain i becomes larger.

For gross substitutes, we write alternatively that for some $j \neq i$,

$$\begin{aligned}\mathbb{P}(i) &= \sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\big[\mathbb{1}(b\ni j)+\mathbb{1}(b\not\ni j)\big]\int\Big(\prod_{b'\in\mathscr{B}} F(v_b - v_{b'} + \epsilon_b)\Big)dF(\epsilon_b)\\ &= \sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\big[\mathbb{1}(b\ni j)+\mathbb{1}(b\not\ni j)\big]\int\Big(\prod_{b'\ni j} F(v_b - v_{b'} + \epsilon_b)\cdot\prod_{b'\not\ni j} F(v_b - v_{b'} + \epsilon_b)\Big)dF(\epsilon_b)\\ &= \sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\mathbb{1}(b\ni j)\int\Big(\prod_{b'\ni j} F(v_b - v_{b'} + \epsilon_b)\cdot\prod_{b'\not\ni j} F(v_b - v_{b'} + \epsilon_b)\Big)dF(\epsilon_b)\\ &\quad+\sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\mathbb{1}(b\not\ni j)\int\Big(\prod_{b'\ni j} F(v_b - v_{b'} + \epsilon_b)\cdot\prod_{b'\not\ni j} F(v_b - v_{b'} + \epsilon_b)\Big)dF(\epsilon_b) \qquad (11)\end{aligned}$$

In this case, given $b \ni i$ and if $b \ni j$, for each $b' \ni j$, $\frac{\partial F(v_b - v_{b'} + \epsilon_b)}{\partial \delta_i} = 0$, and for each $b' \not\ni j$, $\frac{\partial F(v_b - v_{b'} + \epsilon_b)}{\partial \delta_i} = f(v_b - v_{b'} + \epsilon_b) > 0$. However, if $b \not\ni j$, for each $b' \ni j$, $\frac{\partial F(v_b - v_{b'} + \epsilon_b)}{\partial \delta_i} = -f(v_b - v_{b'} + \epsilon_b) < 0$, and for each $b' \not\ni j$, $\frac{\partial F(v_b - v_{b'} + \epsilon_b)}{\partial \delta_i} = 0$. Therefore, the sign of $\frac{\partial \mathbb{P}(i)}{\partial \delta_j}$ can go either way.

Intuitively, for any bundle that contains i, when some other product j's intrinsic value, δ_j, increases, only the bundle that does not involve i and j together becomes less likely to be purchased; in particular, choice probability of a bundle that contains both i and j actually increases. The two opposing forces make us unable to sign the cross effect *ex ante*. In fact, we can give many examples for which (P2) does not hold.

The example below demonstrates even with "Logit" errors, (P2) may not hold.

Example 1 (Type-I extreme value distribution) *1. When $n = 2$, $i \in \{A, B, C\}$, $\mathscr{B} = \{A, B, AB, \emptyset\}$; e.g.,*

$$\frac{\partial \mathbb{P}(A)}{\partial \delta_B} = \frac{e^{\delta_A+\delta_B}(e^{-\Gamma_{AB}} - 1)}{\left(\sum_{b' \in \mathscr{B}} e^{v_{b'}}\right)^2} < 0 \quad \text{iff} \quad \Gamma_{AB} > 0. \tag{12}$$

2. When $n = 3$, $i \in \{A, B, C\}$, $\mathscr{B} = \{A, B, C, AB, CB, AC, ABC, \emptyset\}$; e.g.,

$$\frac{\partial \mathbb{P}(A)}{\partial \delta_B} = \frac{e^{\delta_A+\delta_B}(e^{-\Gamma_{AB}} - 1) + e^{\delta_C}[e^{-\Gamma_{ABC}} + e^{-\Gamma_{AB}} - e^{-\Gamma_{AC}} - e^{-\Gamma_{BC}} + e^{\delta_C}(e^{-\Gamma_{ABC}} - e^{-(\Gamma_{AC}+\Gamma_{BC})})]}{\left(\sum_{b' \in \mathscr{B}} e^{v_{b'}}\right)^2}. \tag{13}$$

In this case, It is easy to find some values for $\Gamma()$ such that the denominator is non-negative. For further discussion and examples of the Gumbel case, see Section ? below.

Claim 1 *In a bundle model with Gumbel error, P1 and P3 are always satisfied while P2 is not.*

Example 2 (Normal distribution) *Another widely used distribution is Normal. Without loss of generality, consider the case of Standard Normal with $\Gamma = 0$ and $n = 2$. It is straightforward to verify that $\frac{\partial \mathbb{P}(A)}{\partial \delta_B} > 0$ for the following pairs of values of (δ_A, δ_B): $(7, 9)$, $(8, 9)$, $(9, 8)$, ... for instance.*

In a Hendel-Fan model, we can write explicitly the expression for $\mathbb{P}(i)$:

$$\begin{aligned}
\mathbb{P}(i) =& \int_{-\infty}^{\infty} \Big\{ F(\delta_i + \epsilon_i) \cdot \prod_{\substack{j \in \{1,\cdots,n\} \\ j \neq i}} F(\delta_i + \epsilon_i - \delta_j) \Big\} dF(\epsilon_i) \\
&+ \sum_{\substack{j \in \{1,\cdots,n\} \\ j \neq i}} \int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot \prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma(2)) \Big\} dF(\epsilon_i) \\
&+ \sum_{\substack{j \in \{1,\cdots,n\} \\ k \in \{1,\cdots,n\} \\ l \neq j \neq i}} \int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot [1 - F(\delta_i + \epsilon_i - \delta_k)] \cdot \prod_{\substack{l \in \{1,\cdots,n\} \\ l \neq k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_l) \cdot F(\delta_i + \epsilon_i - \Gamma(3)) \Big\} \\
&\vdots \\
&+ \int_{-\infty}^{\infty} \Big\{ \prod_{\substack{j \in \{1,\cdots,n\} \\ j \neq i}} [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot F(\delta_i + \epsilon_i - \Gamma(n)) \Big\} dF(\epsilon_i). \qquad (14)
\end{aligned}$$

First, we notice that when $\Gamma(N) \to \infty$ for $N > 1$, only the first term in (8) survives, and all other terms go to zero; hence, $\mathbb{P}(i) = Pr(u_i \geq \max_{k \neq i}\{u_k\})$, and (**P3a**) holds. This property is well-known in the literature.

On the hand other, when $\Gamma(N) = 0$ for all N, the last item in ?, $\mathbb{P}^n(i)$, which is the probability that i is chosen as the N^{th} choice is

$$\begin{aligned}
\mathbb{P}^n(i) = Pr\Big(\min_{j \neq i} u_j > u_i > u_\emptyset\Big) &= Pr\Big(u_i > u_\emptyset\Big) - \sum_{\substack{k=1 \\ k \neq i}}^{N} Pr\Big(\min_{\substack{j=\{1,\cdots,N\} \\ j \neq i \neq k}} \{u_j\} > u_i > u_k; u_i > u_\emptyset\Big) - \cdots \\
&\quad - \sum_{\substack{j=1 \\ j \neq i}}^{N} Pr\Big(u_j > u_i > \max_{\substack{k=\{1,\cdots,N\} \\ k \neq i,j}} \{u_k\}; u_i > u_\emptyset\Big) - Pr\Big(u_i > \max_{j \neq i}\{u_j\}; u_i > u \\
&= Pr(u_i > u_\emptyset) - \sum_{m=1}^{n-1} \mathbb{P}^m(i).
\end{aligned}$$

Therefore, $\mathbb{P}(i) = \sum_{m=1}^{n-1} \mathbb{P}^m(i) + \Big(Pr(u_i > u_\emptyset) - \sum_{m=1}^{n-1} \mathbb{P}^m(i)\Big) = Pr(u_i > u_\emptyset)$

Also for the first term in the expression, which is $\mathbb{P}^1(i)$, it is easy to see that $\frac{\partial \mathbb{P}^1(i)}{\partial \theta_i} > 0$. However, for each of the subsequent terms, the partial derivative with respect to θ_i can go either direction. The difficulty is to show that their sum is positive. For the special case of Logit, Fan (2013) shows this property holds when $m = 2$ or in language of this paper, she assumes that $\Gamma(m) \to \infty$ for $m > 2$. Shi (2017) proves the Logit case with $m = 4$. We prove these properties for a general distribution of the randomness.

Proposition 2 *In a Hendel model, P1 to P3 are always satisfied.*

Lemma 2 *In a Hendel model,* $\sum_{k=1}^{n} \frac{\partial \mathbb{P}(i)}{\partial \delta_k} > 0$.

Proof. We show Lemma and P2 first. Together they imply P1. Details in the Appendix. ■

4 Type I Extreme Value Distribution

In Logit cases with a small number of products, (P2) and (P3b) are often satisfied. This is one main reason that in empirical settings, problems with the bundle model are less obvious to researchers. Nonetheless, our study shows that the Gumbel distribution is very special.

the probability of making each "discrete choice" of bundle $b \in \mathscr{B}$ under the "Multinomial Logit" form is

$$Pr(u_b = \max_{b' \in \mathscr{B}} u_{b'}) = \frac{e^{v_b}}{\sum_{b' \in \mathscr{B}} e^{v_{b'}}} = \frac{e^{v_b}}{1 + \sum_{b' \in \mathscr{B}/\{0\}} e^{v_{b'}}}.$$

In many contexts, instead of a bundle, we care more about the probability that a product is ever chosen, which can be written as

$$\mathbb{P}(i) \equiv \sum_{b \in \mathscr{B}} \mathbb{1}(b \ni i) \cdot Pr(u_b = \max_{b' \in \mathscr{B}} u_{b'}) = \frac{\sum_{b \in \mathscr{B}} \mathbb{1}(b \ni i) \cdot e^{v_{b'}}}{1 + \sum_{b' \in \mathscr{B}/\{0\}} e^{v_{b'}}}, \tag{15}$$

where $\mathbb{1}(b \ni i)$ is an indicator function that returns 1 if i is an element of b and 0 otherwise.

Note that the parts of the denominator in (15) can be spitted into two parts with

$$\sum_{b' \in \mathscr{B}/\{0\}} e^{v_{b'}} = \sum_{b \in \mathscr{B}} \mathbb{1}(b \ni i) \cdot e^{v_b} + \sum_{b \in \mathscr{B}/\{0\}} \mathbb{1}(b \not\ni i) \cdot e^{v_b}.$$

For the numerator,

$$\sum_{b\in\mathscr{B}} \mathbb{1}(b \ni i)\cdot e^{v_b} = e^{\delta_i}\cdot\Big(\sum_{b\in\mathscr{B}/\{0\}} \mathbb{1}(b\ni i)\cdot e^{v_b-\delta_i}\Big).$$

Dividing both the numerator and the denominator by $\sum_{b\in\mathscr{B}/\{0\}} \mathbb{1}(b\ni i)\cdot e^{v_b-\delta_i}$ – which is independent of δ_i by construction – yields

$$\mathbb{P}(i) = \frac{e^{\delta_i}}{e^{\delta_i}+R}, \tag{16}$$

where $R = \frac{1+\sum_{b\in\mathscr{B}/\{0\}} \mathbb{1}(b\not\ni i)\cdot e^{v_b}}{\sum_{b\in\mathscr{B}/\{0\}} \mathbb{1}(b\ni i)\cdot e^{v_b-\delta_i}}$, which again does not depend on δ_i.

Using this notation, we note conveniently the following features of this model.

First, if $\Gamma(N)=0$, $\sum_{b\in\mathscr{B}/\{0\}} \mathbb{1}(b\ni i)\cdot e^{v_b-\delta_i} = 1+\sum_{b\in\mathscr{B}/\{0\}} \mathbb{1}(b\not\ni i)\cdot e^{v_b}$, so $R=1$, and

$$\mathbb{P}(i) = \frac{e^{\delta_i}}{e^{\delta_i}+1}\Big(=Pr(u_i>u_0)\Big).$$

Remark 1 *A bundel model with "logit" errors satisfies* **(P3b)**.

Second, the exact expression is analogous to that of a standard logit model.

Remark 2 (own effect) $\dfrac{\partial\mathbb{P}(i)}{\partial\delta_i} = \dfrac{e^{\delta_i}\cdot R}{(e^{\delta_i}+R)^2}\Big(=\mathbb{P}(i)\cdot(1-\mathbb{P}(i))\Big)>0.$

This property could be extend to general distributions (will be analyzed later??).

Remark 3 *No IIA at the product level.*

Remark 4 (Slutsky symmetry) $\dfrac{\partial\mathbb{P}(i)}{\partial\delta_j} = \dfrac{\partial\mathbb{P}(j)}{\partial\delta_i}$ *for any* $j\neq i$.

Proof. In particular,

$$\frac{\partial \mathbb{P}(i)}{\partial \delta_j} = \frac{\partial}{\partial \delta_j}\left(\frac{\sum\limits_{b\in\mathscr{B}} \mathbb{1}(b \ni i)\cdot e^{v_b}}{\sum\limits_{b'\in\mathscr{B}} e^{v_{b'}}}\right)$$

$$= \frac{\left(\sum\limits_{b\in\mathscr{B}} \mathbb{1}(b \ni i)\mathbb{1}(b \ni j)\cdot e^{v_b}\right)\cdot \sum\limits_{b'\in\mathscr{B}} e^{v_{b'}} - \left(\sum\limits_{b\in\mathscr{B}} \mathbb{1}(b \ni i)\cdot e^{v_b}\right)\cdot\left(\sum\limits_{b\in\mathscr{B}} \mathbb{1}(b \ni j)\cdot e^{v_b}\right)}{\left(\sum\limits_{b'\in\mathscr{B}} e^{v_{b'}}\right)^2},$$

and

$$\frac{\partial \mathbb{P}(j)}{\partial \delta_i} = \frac{\partial}{\partial \delta_i}\left(\frac{\sum\limits_{b\in\mathscr{B}} \mathbb{1}(b \ni j)\cdot e^{v_b}}{\sum\limits_{b'\in\mathscr{B}} e^{v_{b'}}}\right)$$

$$= \frac{\left(\sum\limits_{b\in\mathscr{B}} \mathbb{1}(b \ni j)\mathbb{1}(b \ni i)\cdot e^{v_b}\right)\cdot \sum\limits_{b'\in\mathscr{B}} e^{v_{b'}} - \left(\sum\limits_{b\in\mathscr{B}} \mathbb{1}(b \ni j)\cdot e^{v_b}\right)\cdot\left(\sum\limits_{b\in\mathscr{B}} \mathbb{1}(b \ni i)\cdot e^{v_b}\right)}{\left(\sum\limits_{b'\in\mathscr{B}} e^{v_{b'}}\right)^2}.$$

■

For any general distribution of ϵ_b, $F(\cdot)$, consider a simple two-good case, $\{A, B\}$, then the probability that good A is involved, for simplicity, we denote $\Gamma(|2|) = \Gamma$, then

$$\begin{aligned} Pr(A) = \int [&F(\delta_A + x)F(\delta_A - \delta_B + x)F(-\delta_B + \Gamma + x) \\ &+ F(\delta_B - \Gamma + x)F(\delta_A - \Gamma + x)F(\delta_A + \delta_B - \Gamma + x)]dF(x) \end{aligned}$$

by Definition 1. Hence the cross effect is

$$\frac{\partial Pr(A)}{\partial \delta_B} = \int [\underbrace{-F(\delta_A + x) f(\delta_A - \delta_B + x) F(-\delta_B + \Gamma + x)}_{\text{term 1}} \underbrace{-F(\delta_A + x) F(\delta_A - \delta_B + x) f(-\delta_B + \Gamma + x)}_{\text{term 2}} \underbrace{+f(\delta_B - \Gamma + x) F(\delta_A - \Gamma + x) F(\delta_A + \delta_B - \Gamma + x)}_{\text{term 3}} \underbrace{+F(\delta_B - \Gamma + x) F(\delta_A - \Gamma + x) f(\delta_A + \delta_B - \Gamma + x)}_{\text{term 4}}] dF(x). \quad (17)$$

5 Discussion

Bundle with product-specific shocks

Does a bundle model with product-specific shocks solves the problem? No.

Without loss of generality, we look at the case with $n = 2$; in particular,

$$u_{i,j} = \delta_i + \epsilon_j + \delta_i + \epsilon_j - \Gamma;$$

and for convenience, we call it ''type-2 bundle" model.

Proposition 3 *Under "type-2 bundle" assumption, (P1) and (P2) doe not always hold for general distributions.*

Proof. We start with the two-good case, A and B, then type-2 bundle option means,

$$u^2_{AB} = \delta_A + \epsilon_A + \delta_B + \epsilon_B - \Gamma.$$

For $\Gamma = 0$,

$$\begin{aligned}
Pr\Big(\text{include } A\Big) &= Pr\Big(u_A \geq u_B, u_A \geq u_{AB}^2, u_A \geq u_\emptyset\Big) + Pr\Big(u_{AB}^2 \geq u_A, u_{AB}^2 \geq u_B, u_{AB}^2 \geq u_\emptyset\Big) \\
&= Pr\begin{pmatrix} \delta_A + \epsilon_A \geq \delta_B + \epsilon_B \\ \delta_A + \epsilon_A \geq \delta_A + \epsilon_A + \delta_B + \epsilon_B \\ \delta_A + \epsilon_A \geq 0 + \epsilon_\emptyset \end{pmatrix} + Pr\begin{pmatrix} \delta_A + \epsilon_A + \delta_B + \epsilon_B \geq \delta_A + \epsilon_A \\ \delta_A + \epsilon_A + \delta_B + \epsilon_B \geq \delta_B + \epsilon_B \\ \delta_A + \epsilon_A + \delta_B + \epsilon_B \geq 0 + \epsilon_\emptyset \end{pmatrix} \\
&= Pr\begin{pmatrix} \epsilon_B \leq \delta_A - \delta_B + \epsilon_A \\ \epsilon_B \leq -\delta_B \\ \epsilon_\emptyset \leq \delta_A + \epsilon_A \end{pmatrix} + Pr\begin{pmatrix} \epsilon_B \geq -\delta_B \\ \epsilon_A \geq -\delta_A \\ \epsilon_\emptyset \leq \delta_A + \epsilon_A + \delta_B + \epsilon_B \end{pmatrix} \\
&= \underbrace{\int_{-\infty}^{-\delta_A} \Big(F(\delta_A - \delta_B + x)F(\delta_A + x)\Big) \cdot f(x)dx}_{1^{st}\text{ term}} + \underbrace{\int_{-\delta_A}^{\infty} \Big(F(\delta_A + x)F(-\delta_B)\Big) \cdot f(x)dx}_{2^{nd}\text{ term}} \\
&\quad + \underbrace{\int_{-\delta_B}^{\infty} \int_{-\delta_A}^{\infty} F(\delta_A + \delta_B + x + y) \cdot f(x)f(y)dxdy}_{3^{rd}\text{ term}}.
\end{aligned}$$

We denote the third term in above equation by,

$$\psi(\delta_A, \delta_B) \triangleq \int_{-\delta_B}^{\infty} \int_{-\delta_A}^{\infty} F(\delta_A + \delta_B + x + y) \cdot f(x)f(y)dxdy.$$

Noted that $\psi(\delta_A, \delta_B) \geq 0$ and

$$\begin{aligned}
\lim_{\delta_B \to -\infty} \psi(\delta_A, \delta_B) &= \lim_{\delta_B \to -\infty} \int_{-\delta_B}^{\infty} \int_{-\delta_A}^{\infty} F(\delta_A + \delta_B + x + y) \cdot f(x)f(y)dxdy \\
&\leq \lim_{\delta_B \to -\infty} \int_{-\delta_B}^{\infty} \int_{-\infty}^{\infty} F(\delta_A + \delta_B + x + y) \cdot f(x)f(y)dxdy \\
&\leq \lim_{\delta_B \to -\infty} \int_{-\delta_B}^{\infty} \int_{-\infty}^{\infty} f(x)f(y)dxdy \\
&= \lim_{\delta_B \to -\infty} \int_{-\delta_B}^{\infty} f(y)dy = 0,
\end{aligned}$$

which implies $\psi(\delta_A, \delta_B) \to 0$ as $\delta_B \to -\infty$. Meanwhile, we have,

$$\begin{aligned}
\lim_{\delta_B \to \infty} \psi(\delta_A, \delta_B) &= \lim_{\delta_B \to \infty} \int_{-\delta_B}^{\infty} \int_{-\delta_A}^{\infty} F(\delta_A + \delta_B + x + y) \cdot f(x) f(y) dx dy \\
&= \lim_{\delta_B \to \infty} \int_{-\infty}^{\infty} \int_{-\delta_A}^{\infty} F(\delta_A + \delta_B + x + y) \cdot f(x) f(y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\delta_A}^{\infty} \lim_{\delta_B \to \infty} F(\delta_A + \delta_B + x + y) \cdot f(x) f(y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\delta_A}^{\infty} f(x) f(y) dx dy = 1 - F(-\delta_A),
\end{aligned}$$

where the third the equality requires some sufficient conditions for the exchange of "integration" and "limit", for example, $f(\cdot)$ has a large rate of decay, i.e., $f(x) \sim e^{-|x|}$, which is true for a lot of widely used distributions (e.g. "logit" case and Exponential distribution, etc.). As a result, for different δ_B, $\psi(\delta_A, \delta_B)$ are different (since $\lim_{\delta_B \to -\infty} \psi(\delta_A, \delta_B) \neq \lim_{\delta_B \to \infty} \psi(\delta_A, \delta_B)$). As a result, $Pr(\text{include } A)$ is a multiple-variable function of (δ_A, δ_B).

On the other hand, once we move the monotonicity part,

$$\begin{aligned}
\frac{\partial}{\partial \delta_A} Pr\Big(\text{include } A\Big) &= \frac{\partial}{\partial \delta_A} \int_{-\infty}^{-\delta_A} \Big(F(\delta_A - \delta_B + x) F(\delta_A + x)\Big) \cdot f(x) dx + \frac{\partial}{\partial \delta_A} \int_{-\delta_A}^{\infty} \Big(F(\delta_A + x) F(-\delta_B)\Big) \cdot f \\
&\quad + \frac{\partial}{\partial \delta_A} \int_{-\delta_B}^{\infty} \int_{-\delta_A}^{\infty} F(\delta_A + \delta_B + x + y) \cdot f(x) f(y) dx dy \\
&= -\int_{-\infty}^{0} F(z - \delta_B) F(z) f'(z - \delta_A) dz - \int_{0}^{\infty} F(z) f'(z - \delta_A) dz \\
&\quad - \int_{0}^{\infty} \int_{0}^{\infty} F(z + \hat{z}) \cdot f'(z - \delta_A) f(\hat{z} - \delta_B) dz d\hat{z},
\end{aligned}$$

where $z = x + \delta_A$ and $\hat{z} = y + \delta_B$.[2] We claim that the sign of (18) is distribution-specific, for example,

[2] By "changing of variable", set $z = x + \delta_A$ and $\hat{z} = y + \delta_B$, we have,

$$\frac{\partial}{\partial \delta_A} \int_{-\infty}^{-\delta_A} \Big(F(\delta_A - \delta_B + x) F(\delta_A + x)\Big) \cdot f(x) dx = \frac{\partial}{\partial \delta_A} \int_{-\infty}^{0} F(z - \delta_B) F(z) f(z - \delta_A) dz = -\int_{-\infty}^{0} F(z - \delta_B) F(z) f'(z - \delta_A) dz;$$

$$\frac{\partial}{\partial \delta_A} \int_{-\delta_A}^{\infty} \Big(F(\delta_A + x) F(-\delta_B)\Big) \cdot f(x) dx = \frac{\partial}{\partial \delta_A} \int_{0}^{\infty} F(z) f(z - \delta_A) dz = -\int_{0}^{\infty} F(z) f'(z - \delta_A) dz;$$

Example 3 *Consider the exponential distribution,*

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0; \\ 0 & else. \end{cases}$$

Thus we have, for $\delta_A, \delta_B > 0$,

$$\frac{\partial}{\partial \delta_A} Pr\Big(include\ A\Big) = \frac{1}{2} e^{\delta_A \cdot \lambda} + \frac{1}{4} e^{-(\delta_A + \delta_B)\cdot \lambda} - 2. \tag{19}$$

Obviously, the sign of (19) *could be either positive or negative, which is depend on* δ_A, δ_B *and* λ.

Similarly, above results also hold for the case with $\Gamma > 0$. To sum up, for the two-good case, the monotonic increasing and δ_A-dependency properties not always hold for all general distributions, neither does the general n-good case. ■

Invertibility of demand

Corollary 5 *The demand system arising from a Hendel-Fan model is invertible.*

Following Berry *et al.* (2013)'s main theorem, the underlying products in a Hendel-Fan model satisfy the definition of "connected substitutes," so the system of demand is invertible.

---

and

$$\begin{aligned} \frac{\partial}{\partial \delta_A} \int_{-\delta_B}^{\infty} \int_{-\delta_A}^{\infty} F(\delta_A + \delta_B + x + y) \cdot f(x) f(y) dx dy &= \frac{\partial}{\partial \delta_A} \int_{-\delta_B}^{\infty} \int_{0}^{\infty} F(z + \delta_B + y) \cdot f(z - \delta_A) f(y) dz dy \\ &= \frac{\partial}{\partial \delta_A} \int_{0}^{\infty} \int_{0}^{\infty} F(z + \hat{z}) \cdot f(z - \delta_A) f(\hat{z} - \delta_B) dz d\hat{z} \\ &= -\int_{0}^{\infty} \int_{0}^{\infty} F(z + \hat{z}) \cdot f'(z - \delta_A) f(\hat{z} - \delta_B) dz d\hat{z}. \end{aligned}$$

To sum up,

$$\begin{aligned} \frac{\partial}{\partial \delta_A} Pr\Big(\text{include } A\Big) &= -\int_{-\infty}^{0} F(z - \delta_B) F(z) f'(z - \delta_A) dz - \int_{0}^{\infty} F(z) f'(z - \delta_A) dz \\ &\quad - \int_{0}^{\infty} \int_{0}^{\infty} F(z + \hat{z}) \cdot f'(z - \delta_A) f(\hat{z} - \delta_B) dz d\hat{z}. \end{aligned}$$

On the other hand, products in a bundle model may not be gross substitutes. Its demand invertibility therefore remains a question.

6 Conclusion

(To be written.)

References

Ambrus, A., E. Calvano, and M. Reisinger (2017): "Either or Both Competition: A Two-sided Theory of Advertising with Overlapping Viewerships," *American Economic Journal: Microeconomics*, forthcoming.

Anderson, S. P., O. Foros, and H. J. Kind (2017): "Competition for Advertisers and for Viewers in Media Markets," *Economic Journal*, forthcoming.

Augereau, A., S. Greenstein, and M. Rysman (2006): "Coordination versus differentiation in a standards war: 56k modems," *RAND Journal of Economics*, 37(4), 887–909.

Berry, S., A. Gandhi, and P. Haile (2013): "Connected Substitutes and Invertibility of Demand," *Econometrica*, 81(5), 2087–2111.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

Deza, M. (2015): "Is there a stepping stone effect in drug use? Separating state dependence from unobserved heterogeneity within and between illicit drugs," *Journal of Econometrics*, 184(1), 193–207.

Dube, J.-P. (2004): "Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks," *Marketing Science*, 23(1), 66–81.

Fan, Y. (2013): "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market," *American Economic Review*, 103(5), 1598–1628.

Gentzkow, M. (2007): "Valuing new goods in a model with complementarity: Online newspapers," *American Economic Review*, 97(3), 713–744.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2014): "Competition and Ideological Diversity: Historical Evidence from US Newspapers," *American Economic Review*, 104(10).

Hendel, I. (1999): "Estimating multiple-discrete choice models: An application to computerization returns." *Review of Economic Studies*, 66, 423–446.

Hiedemann, B., M. Sovinsky, and S. Stern (2013): "Will You Still Want Me Tomorrow? The Dynamics of Families' Long-Term Care Arrangements." *Working Paper.*

Liu, H., P. K. Chintagunta, and T. Zhu (2010): “Complementarities and the Demand for Home Broadband Internet Services,” *Marketing Science*, 29(4), 701–720.

Pereira, P., T. Ribeiro, and J. Vareda (2013): “Delineating markets for bundles with consumer level data: The case of triple-play,” *International Journal of Industrial Organization*, 31(6), 760–773.

Perry, E. D., G. Moschini, and D. A. Hennessy (2016): “Testing for Complementarity: Glyphosate Tolerant Soybeans and Conservation Tillage,” *American Journal of Agricultural Economics*, 98(3), 765–784.

Shi, C. M. (2017): “Catching (Exclusive) Eyeballs: Multi-Homing and Platform Competition in the Magazine industry,” *Working paper.*

Song, I. and P. K. Chintagunta (2006): “Measuring Cross-Category Price Effects with Aggregate Store Data,” *Management Science*, 52(10), 1594–1609.

Appendix

Omitted Proofs

Lemma 1

Let’s simply consider a two-good case, denoted by $\{A, B\}$. According to the definition (1), we have

$$Pr(A) = \int_{-\infty}^{\infty} \Big(\underbrace{F(\delta_A + x)}_{u_A > u_\emptyset} \cdot \underbrace{F(\delta_A - \delta_B + x)}_{u_A > u_B} \cdot \underbrace{F(-\delta_B + x)}_{u_A > u_{AB}} + \underbrace{F(\delta_A + x)}_{u_{AB} > u_B} \cdot \underbrace{F(\delta_B + x)}_{u_{AB} > u_A} \cdot \underbrace{F(\delta_A + \delta_B + x)}_{u_{AB} > u_\emptyset} \Big) dF(x).$$

We denote $Pr(A) \triangleq g(\delta_A, \delta_B)$. Note that

$$\lim_{\delta_B \to \infty} g(\delta_A, \delta_B) = \int_{-\infty}^{\infty} F(\delta_A + x) dF(x),$$

which is function of δ_A. To prove that $Pr(A)$ depends only on δ_A is sufficient to show $g(\delta_A, \delta_B) = \int_{-\infty}^{\infty} F(\delta_A + x) dF(x)$ for all δ_B. However, for $\delta_B = 0$,

$$Pr(\text{includes } A) = \int_{-\infty}^{\infty} 2F^2(\delta_A + x) \cdot F(x) dF(x)$$

The independent of δ_B requires that

$$\int_{-\infty}^{\infty} 2F^2(\delta_A + x) \cdot F(x) dF(x) = \int_{-\infty}^{\infty} F(\delta_A + x) dF(x). \tag{20}$$

Obviously, above equality is not held for all general distribution $F(x)$. For example,

Example 4 *Consider the following exponential distribution,*

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & else. \end{cases}$$

For $\delta_A \geq 0$, we have:

$$\phi(\delta_A) = \int_{-\infty}^{\infty} F(\delta_A + x) \cdot f(x)dx = \frac{1}{2}e^{-\delta_A \lambda} \cdot (-1 + 2e^{\delta_A \lambda}). \tag{21}$$

Meanwhile,

$$g(\delta_A, \delta_B)\Big|_{\delta_B = 0} = \frac{1}{6}e^{-2\delta_A \lambda} \cdot (1 - 4e^{\delta_A \lambda} + 6e^{2\delta \lambda}). \tag{22}$$

Obviously, (21)≠(22), which means $g(\delta_A, \delta_B) \neq \phi(\delta_A)$ for $\delta_B = 0$.

Therefore, Example 4 above implies that δ_i-dependent-only property fails to hold for $\Gamma(|b|) = 0$.

Remark 6 *We have proved that $Pr(A) = \frac{e^{\delta_A}}{e^{\delta_A + 1}}$ for the Type I extreme value distribution case in Observation 1, which is δ_A-dependent-only. Thus, we claim that such distribution $F(x) = e^{-e^{-(x+\gamma)}}$ satisfies (20). In particular,*

$$\begin{aligned} LHS &= \int_{-\infty}^{\infty} 2F^2(\delta_A + x) \cdot F(x)dF(x) = \int_{-\infty}^{\infty} F^2(\delta_A + x)dF^2(x) = \int_{-\infty}^{\infty} e^{-2e^{-(\delta_A + x + \gamma)}} dF^2(x) \\ &= \int_{-\infty}^{\infty} e^{-2e^{-\delta_A} \cdot e^{-(x+\gamma)}} dF^2(x) = \int_{-\infty}^{\infty} [F^2(x)]^{e^{\delta_A}} dF^2(x) = \frac{1}{e^{\delta_A} + 1} [F^2(x)]^{e^{\delta_A} + 1}\Big|_{-\infty}^{\infty} = \frac{1}{e^{\delta_A} + 1}. \end{aligned}$$

Meanwhile, consider the right-hand side of (20),

$$\begin{aligned} RHS &= \int_{-\infty}^{\infty} F(\delta_A + x)dF(x) = \int_{-\infty}^{\infty} e^{-e^{-(\delta_A + x + \gamma)}} dF(x) = \int_{-\infty}^{\infty} e^{-e^{-\delta_A} \cdot e^{-(x+\gamma)}} dF(x) \\ &= \int_{-\infty}^{\infty} [F(x)]^{e^{\delta_A}} dF(x) = \frac{1}{e^{\delta_A} + 1} [F(x)]^{e^{\delta_A} + 1}\Big|_{-\infty}^{\infty} = \frac{1}{e^{\delta_A} + 1} = LHS. \end{aligned}$$

Therefore we have proved that (20) holds for the special case, $F(x) = e^{-e^{-(x+\gamma)}}$, which coincides with our previous results.

For monotonicity, using equation (10),

$$\begin{aligned} \frac{\partial \mathbb{P}(i)}{\partial \delta_i} &= \sum_{b \in \mathscr{B}} \mathbb{1}(b \ni i) \cdot \int \left(\prod_{b' \ni i} F(v_b - v_{b'} + \epsilon_b) \cdot \frac{\partial}{\partial \delta_i} \Big(\prod_{b' \not\ni i} F(v_b - v_{b'} + \epsilon_b) \Big) \right) dF(\epsilon_b) \\ &= \sum_{b \in \mathscr{B}} \mathbb{1}(b \ni i) \cdot \int \left(\prod_{b' \ni i} F(v_b - v_{b'} + \epsilon_b) \cdot \prod_{b' \not\ni i} \Big(f(v_b - v_{b'} + \epsilon_b) \prod_{\substack{b'' \neq b, b' \\ b'' \not\ni i}} F(v_b - v_{b''} + \epsilon_b) \Big) \right) dF(\epsilon_b) \\ &> 0. \end{aligned}$$

For the cross effect, using equation (11),

$$\begin{aligned}
\frac{\partial \mathbb{P}(i)}{\partial \delta_j} &= \sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\mathbb{1}(b\ni j)\int \left(\prod_{b'\ni j} F(v_b - v_{b'}+\epsilon_b)\cdot \frac{\partial}{\partial \delta_j}\Big(\prod_{b'\not\ni j} F(v_b - v_{b'}+\epsilon_b)\Big)\right) dF(\epsilon_b) \\
&\quad + \sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\mathbb{1}(b\not\ni j)\int \left(\frac{\partial}{\partial \delta_j}\Big(\prod_{b'\ni j} F(v_b - v_{b'}+\epsilon_b)\Big)\cdot \prod_{b'\not\ni j} F(v_b - v_{b'}+\epsilon_b)\right) dF(\epsilon_b) \\
&= \sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\mathbb{1}(b\ni j)\int \left(\prod_{b'\ni j} F(v_b - v_{b'}+\epsilon_b)\cdot \prod_{b'\not\ni j}\Big(f(v_b - v_{b'}+\epsilon_b)\prod_{\substack{b''\neq b,b' \\ b''\not\ni j}} F(v_b - v_{b''}+\epsilon_b)\Big)\right) dF(\epsilon_b) \\
&\quad + \sum_{b\in\mathscr{B}} \mathbb{1}(b\ni i)\mathbb{1}(b\not\ni j)\int \left(\prod_{b'\ni j}\Big(-f(v_b - v_{b'}+\epsilon_b)\prod_{\substack{b''\neq b,b' \\ b''\ni j}} F(v_b - v_{b''}+\epsilon_b)\Big)\cdot \prod_{b'\not\ni j} F(v_b - v_{b'}+\epsilon_b)\right) dF(\epsilon_b)
\end{aligned}$$

Proof of Proposition 2

Proof. Noted that, we have proved

$$\frac{\partial Pr(\text{include } A)}{\partial \delta_B} < 0,$$

which means the change of δ_B also results in the changes in $Pr(\text{include } A)$.

Further, the result of above 3-good example in Lemma **??** can be extended to the more general n-good case. Specifically, for any $i \in \{1, \cdots, n\}$,

$$\begin{aligned}
&Pr(\text{include } i) \\
=&\int_{-\infty}^{\infty} \Big\{F(\delta_i+\epsilon_i)\cdot \prod_{\substack{j\in\{1,\cdots,n\} \\ j\neq i}} F(\delta_i+\epsilon_i-\delta_j)\Big\} dF(\epsilon_i) \\
&+ \sum_{\substack{j\in\{1,\cdots,n\} \\ j\neq i}} \int_{-\infty}^{\infty} \Big\{[1-F(\delta_i+\epsilon_i-\delta_j)]\cdot \prod_{\substack{k\in\{1,\cdots,n\} \\ k\neq j\neq i}} F(\delta_i+\epsilon_i-\delta_k)\cdot F(\delta_i+\epsilon_i-\Gamma)\Big\} dF(\epsilon_i) \\
&+ \sum_{\substack{j\in\{1,\cdots,n\} \\ k\in\{1,\cdots,n\} \\ l\neq j\neq i}} \int_{-\infty}^{\infty} \Big\{[1-F(\delta_i+\epsilon_i-\delta_j)]\cdot[1-F(\delta_i+\epsilon_i-\delta_k)]\cdot \prod_{\substack{l\in\{1,\cdots,n\} \\ l\neq k\neq j\neq i}} F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-2\Gamma)\Big\} dF(\epsilon) \\
&\vdots \\
&+ \int_{-\infty}^{\infty} \Big\{\prod_{\substack{j\in\{1,\cdots,n\} \\ j\neq i}} [1-F(\delta_i+\epsilon_i-\delta_j)]\cdot F(\delta_i+\epsilon_i-(n-1)\Gamma)\Big\} dF(\epsilon_i). \qquad (23)
\end{aligned}$$

On one hand, for simplicity, we denote $Pr(\text{include } i)$ as Pr_i. Considering the following variation,

$$\Phi(\triangle) = Pr_i(\delta_i + \triangle, \delta_{-i} + \triangle),$$

given other goods' evaluations δ_{-i}, where $-i \in \{1, \cdots, n\}$ and $-i \neq i$, obviously, $\Phi(\triangle)|_{\triangle=0} = Pr(\text{include } i)$. Noted that for each separate integration in (23), the only $\triangle$-dependent term occurs in its last term, i.e.,

$$F(\delta_i + \epsilon_i + \triangle - (m-1)\Gamma),$$

for $1 \leq m \leq n$, which yields

$$\frac{\partial F(\delta_i + \epsilon_i + \triangle - (m-1)\Gamma)}{\partial \triangle}\Big|_{\triangle=0} = f(\delta_i + \epsilon_i - (m-1)\Gamma) > 0.$$

Therefore, to sum them up, we have $\frac{\partial \Phi(\triangle)}{\partial \triangle}\Big|_{\triangle=0} > 0$.

On the other hand, since we have proved

$$\frac{\partial \Phi(\triangle)}{\partial \triangle}\Big|_{\triangle=0} = \frac{\partial Pr_i}{\partial \delta_i} + \frac{\partial Pr_i}{\partial \delta_{-i}} > 0$$

for all $-i \neq i$. To prove the desired result, it remains to show $\frac{\partial Pr_i}{\partial \delta_{-i}} < 0$. In particular, for some $\delta_j \in \{\delta_{-i}\}$, starting from the first line in (23), i.e.,

$$\int_{-\infty}^{\infty} \Big\{ F(\delta_i + \epsilon_i) \cdot \prod_{\substack{j \in \{1, \cdot, n\} \\ j \neq i}} F(\delta_i + \epsilon_i - \delta_j) \Big\} dF(\epsilon_i),$$

which represents good i as individual's first choice. Then we take the partial derivative w.r.t. δ_j, yields that,

$$\begin{aligned} & \frac{\partial}{\partial \delta_j} \int_{-\infty}^{\infty} \Big\{ F(\delta_i + \epsilon_i) \cdot \prod_{\substack{j \in \{1, \cdots, n\} \\ j \neq i}} F(\delta_i + \epsilon_i - \delta_j) \Big\} dF(\epsilon_i) \\ = & \int_{-\infty}^{\infty} \Big\{ F(\delta_i + \epsilon_i) \cdot F(\delta_i + \epsilon_i - \delta_j) \cdot \prod_{\substack{k \in \{1, \cdots, n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \Big\} dF(\epsilon_i) \\ = & -\int_{-\infty}^{\infty} \Big\{ F(\delta_i + \epsilon_i) \cdot f(\delta_i + \epsilon_i - \delta_j) \cdot \prod_{\substack{k \in \{1, \cdots, n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \Big\} dF(\epsilon_i) \end{aligned} \quad (24)$$

Next, we move on to the second line in (23), i.e.,

$$\sum_{\substack{j \in \{1, \cdots, n\} \\ j \neq i}} \int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot \prod_{\substack{k \in \{1, \cdots, n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma) \Big\} dF(\epsilon_i),$$

which represents good i as individual's second choice. Noted that good j has been chosen either before good i (first choice) or after i. If δ_j occurs at "$1 - F(\delta_i + \epsilon_i - \delta_j)$", we have

$$\begin{aligned} & \frac{\partial}{\partial \delta_j} \int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot \prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma) \Big\} dF(\epsilon_i) \\ = & \int_{-\infty}^{\infty} \Big\{ f(\delta_i + \epsilon_i - \delta_j)] \cdot \prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma) \Big\} dF(\epsilon_i). \end{aligned}$$

Alternatively, if δ_j occurs at "$\prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma)$", we have,

$$\begin{aligned} & \frac{\partial}{\partial \delta_j} \int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot \prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma) \Big\} dF(\epsilon_i) \\ = & \frac{\partial}{\partial \delta_j} \int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_l)] \cdot F(\delta_i + \epsilon_i - \delta_j) \cdot \prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq l \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma) \Big\} dF(\epsilon_i) \\ = & -\int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_l)] \cdot f(\delta_i + \epsilon_i - \delta_j) \cdot \prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq l \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma) \Big\} dF(\epsilon_i). \end{aligned}$$

To sum up,

$$\begin{aligned} & \frac{\partial}{\partial \delta_j} \Big(\sum_{\substack{j \in \{1,\cdots,n\} \\ j \neq i}} \int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot \prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma) \Big\} dF(\epsilon_i) \Big) \qquad (25) \\ = & \underbrace{\int_{-\infty}^{\infty} \Big\{ f(\delta_i + \epsilon_i - \delta_j)] \cdot \prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma) \Big\} dF(\epsilon_i)}_{\text{term 1}} \\ & \underbrace{- \sum_{\substack{l \in \{1,\cdots,n\} \\ l \neq k \neq j \neq i}} \int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_l)] \cdot f(\delta_i + \epsilon_i - \delta_j) \cdot \prod_{\substack{k \in \{1,\cdots,n\} \\ k \neq l \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_k) \cdot F(\delta_i + \epsilon_i - \Gamma) \Big\} dF(\epsilon_i)}_{\text{term 2}} \end{aligned}$$

Similarly, we continue to the third line in (23), i.e.,

$$\sum_{\substack{j \in \{1,\cdots,n\} \\ k \in \{1,\cdots,n\} \\ l \neq j \neq i}} \int_{-\infty}^{\infty} \Big\{ [1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot [1 - F(\delta_i + \epsilon_i - \delta_k)] \cdot \prod_{\substack{l \in \{1,\cdots,n\} \\ l \neq k \neq j \neq i}} F(\delta_i + \epsilon_i - \delta_l) \cdot F(\delta_i + \epsilon_i - 2\Gamma) \Big\} dF(\epsilon_i),$$

which represents good i as individual's third choice. Noted that good j has been chosen either before good i (first or second choice) or after i. If δ_j occurs at "$[1 - F(\delta_i + \epsilon_i - \delta_j)] \cdot [1 - F(\delta_i + \epsilon_i - \delta_k)]$", we have

$$\frac{\partial}{\partial \delta_j}\int_{-\infty}^{\infty}\Big\{[1-F(\delta_i+\epsilon_i-\delta_j)]\cdot[1-F(\delta_i+\epsilon_i-\delta_k)]\cdot\prod_{\substack{l\in\{1,\cdots,n\}\\ l\neq k\neq j\neq i}}F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-2\Gamma)\Big\}dF(\epsilon_i)$$

$$= \sum_{\substack{h\in\{1,\cdots,n\}\\ h\neq l\neq j\neq i}}\int_{-\infty}^{\infty}\Big\{[1-F(\delta_i+\epsilon_i-\delta_h)]\cdot f(\delta_i+\epsilon_i-\delta_j)\cdot\prod_{\substack{l\in\{1,\cdots,n\}\\ h\neq l\neq j\neq i}}F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-2\Gamma)\Big\}dF(\epsilon_i).$$

Alternatively, if δ_j occurs at "$\prod_{\substack{l\in\{1,\cdots,n\}\\ l\neq k\neq j\neq i}}F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-2\Gamma)$", we have,

$$\frac{\partial}{\partial \delta_j}\int_{-\infty}^{\infty}\Big\{[1-F(\delta_i+\epsilon_i-\delta_j)]\cdot[1-F(\delta_i+\epsilon_i-\delta_k)]\cdot\prod_{\substack{l\in\{1,\cdots,n\}\\ l\neq k\neq j\neq i}}F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-2\Gamma)\Big\}dF(\epsilon_i)$$

$$= \frac{\partial}{\partial \delta_j}\int_{-\infty}^{\infty}\Big\{[1-F(\delta_i+\epsilon_i-\delta_h)]\cdot[1-F(\delta_i+\epsilon_i-\delta_k)]\cdot F(\delta_i+\epsilon_i-\delta_j)$$
$$\cdot\prod_{\substack{l\in\{1,\cdots,n\}\\ l\neq h\neq k\neq j\neq i}}F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-2\Gamma)\Big\}dF(\epsilon_i)$$

$$= -\int_{-\infty}^{\infty}\Big\{[1-F(\delta_i+\epsilon_i-\delta_h)]\cdot[1-F(\delta_i+\epsilon_i-\delta_k)]\cdot f(\delta_i+\epsilon_i-\delta_j)\cdot\prod_{\substack{l\in\{1,\cdots,n\}\\ l\neq h\neq k\neq j\neq i}}F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-$$

To sum up,

$$\frac{\partial}{\partial \delta_j}\Big(\sum_{\substack{j\in\{1,\cdots,n\}\\ k\in\{1,\cdots,n\}\\ l\neq j\neq i}}\int_{-\infty}^{\infty}\Big\{[1-F(\delta_i+\epsilon_i-\delta_j)]\cdot[1-F(\delta_i+\epsilon_i-\delta_k)]\cdot\prod_{\substack{l\in\{1,\cdots,n\}\\ l\neq k\neq j\neq i}}F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-2\Gamma)$$

$$= \underbrace{\sum_{\substack{h\in\{1,\cdots,n\}\\ h\neq l\neq j\neq i}}\int_{-\infty}^{\infty}\Big\{[1-F(\delta_i+\epsilon_i-\delta_h)]\cdot f(\delta_i+\epsilon_i-\delta_j)\cdot\prod_{\substack{l\in\{1,\cdots,n\}\\ h\neq l\neq j\neq i}}F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-2\Gamma)\Big\}dF(\epsilon_i)}_{\text{term 1}}$$
$$\underbrace{-\sum_{\substack{h\in\{1,\cdots,n\}\\ k\in\{1,\cdots,n\}\\ l\neq h\neq k\neq j\neq i}}\int_{-\infty}^{\infty}\Big\{[1-F(\delta_i+\epsilon_i-\delta_h)]\cdot[1-F(\delta_i+\epsilon_i-\delta_k)]\cdot f(\delta_i+\epsilon_i-\delta_j)}_{\text{term 2}}$$
$$\underbrace{\cdot\prod_{\substack{l\in\{1,\cdots,n\}\\ l\neq h\neq k\neq j\neq i}}F(\delta_i+\epsilon_i-\delta_l)\cdot F(\delta_i+\epsilon_i-2\Gamma)\Big\}dF(\epsilon_i)}_{\text{term 2}}$$

We consider the summation of (24), (25) and (26). Noted that the term 1 in (25) could be merged to (24), i.e.,

$$
\begin{aligned}
& -\int_{-\infty}^{\infty} \Big\{ F(\delta_i+\epsilon_i) \cdot f(\delta_i+\epsilon_i-\delta_j) \cdot \prod_{\substack{k\in\{1,\cdots,n\} \\ k\neq j\neq i}} F(\delta_i+\epsilon_i-\delta_k) \Big\} dF(\epsilon_i) \\
& + \int_{-\infty}^{\infty} \Big\{ f(\delta_i+\epsilon_i-\delta_j)] \cdot \prod_{\substack{k\in\{1,\cdots,n\} \\ k\neq j\neq i}} F(\delta_i+\epsilon_i-\delta_k) \cdot F(\delta_i+\epsilon_i-\Gamma) \Big\} dF(\epsilon_i) \\
= & \int_{-\infty}^{\infty} \Big\{ f(\delta_i+\epsilon_i-\delta_j)] \cdot \prod_{\substack{k\in\{1,\cdots,n\} \\ k\neq j\neq i}} F(\delta_i+\epsilon_i-\delta_k) \cdot \Big[F(\delta_i+\epsilon_i-\Gamma) - F(\delta_i+\epsilon_i) \Big] \Big\} dF(\epsilon_i) \\
< & \; 0
\end{aligned}
$$

since $F(\delta_i+\epsilon_i-\Gamma) < F(\delta_i+\epsilon_i)$. Similarly, the term 1 in (26) could be merged to the term 2 in (25), i.e.,

$$
\begin{aligned}
& -\sum_{\substack{l\in\{1,\cdots,n\} \\ l\neq k\neq j\neq i}} \int_{-\infty}^{\infty} \Big\{ [1-F(\delta_i+\epsilon_i-\delta_l)] \cdot f(\delta_i+\epsilon_i-\delta_j) \cdot \prod_{\substack{k\in\{1,\cdots,n\} \\ k\neq l\neq j\neq i}} F(\delta_i+\epsilon_i-\delta_k) \cdot F(\delta_i+\epsilon_i-\Gamma) \Big\} dF(\epsilon_i) \\
& + \sum_{\substack{h\in\{1,\cdots,n\} \\ h\neq l\neq j\neq i}} \int_{-\infty}^{\infty} \Big\{ [1-F(\delta_i+\epsilon_i-\delta_h)] \cdot f(\delta_i+\epsilon_i-\delta_j) \cdot \prod_{\substack{l\in\{1,\cdots,n\} \\ h\neq l\neq j\neq i}} F(\delta_i+\epsilon_i-\delta_l) \cdot F(\delta_i+\epsilon_i-2\Gamma) \Big\} dF(\epsilon_i) \\
= & \sum_{\substack{h\in\{1,\cdots,n\} \\ h\neq l\neq j\neq i}} \int_{-\infty}^{\infty} \Big\{ [1-F(\delta_i+\epsilon_i-\delta_h)] \cdot f(\delta_i+\epsilon_i-\delta_j) \cdot \prod_{\substack{l\in\{1,\cdots,n\} \\ h\neq l\neq j\neq i}} F(\delta_i+\epsilon_i-\delta_l) \\
& \cdot [F(\delta_i+\epsilon_i-2\Gamma) - F(\delta_i+\epsilon_i-\Gamma)] \Big\} dF(\epsilon_i) \\
< & \; 0,
\end{aligned}
$$

since $F(\delta_i+\epsilon_i-2\Gamma) < F(\delta_i+\epsilon_i-\Gamma)$.

Following the same direction, the partial derivative of each line in (23) could be divided into two parts, namely "term 1" and "term 2" as (24), (25) and (26). Hence, starting from the first line of (23), the summation of each "term 2" in previous line and "term 1" in the following line is always negative, untiling the partial derivative of the last line in (23), i.e.,

$$
\begin{aligned}
& \frac{\partial}{\partial \delta_j} \int_{-\infty}^{\infty} \Big\{ \prod_{\substack{j\in\{1,\cdots,n\} \\ j\neq i}} [1-F(\delta_i+\epsilon_i-\delta_j)] \cdot F(\delta_i+\epsilon_i-(n-1)\Gamma) \Big\} dF(\epsilon_i) \\
= & \frac{\partial}{\partial \delta_j} \int_{-\infty}^{\infty} \Big\{ [1-F(\delta_i+\epsilon_i-\delta_j)] \cdot \prod_{\substack{k\in\{1,\cdots,n\} \\ k\neq j\neq i}} [1-F(\delta_i+\epsilon_i-\delta_k)] \cdot F(\delta_i+\epsilon_i-(n-1)\Gamma) \Big\} dF(\epsilon_i) \\
= & \int_{-\infty}^{\infty} \Big\{ f(\delta_i+\epsilon_i-\delta_j) \cdot \prod_{\substack{k\in\{1,\cdots,n\} \\ k\neq j\neq i}} [1-F(\delta_i+\epsilon_i-\delta_k)] \cdot F(\delta_i+\epsilon_i-(n-1)\Gamma) \Big\} dF(\epsilon_i),
\end{aligned}
$$

which could be regard as another "term 1" representation. As a result, adding up all the partial derivative lines ((24), (25), etc.) together yields that $\frac{\partial Pr_i}{\partial \delta_j} < 0$.

Similarly, applying the same arguments, we finally prove that for all $-i \neq i \in \{1, \cdots, n\}$, we have,

$$\frac{\partial Pr_i}{\partial \delta_{-i}} < 0.$$

Combining with $\frac{\partial \Phi(\triangle)}{\partial \triangle}\Big|_{\triangle=0} > 0$, we have proved the desired result. ■