Disentangling Attention and Utility Channels in Recommendations*

Paul H.Y. Cheung[†] Yusufcan Masatlioglu[‡]

Jan 5, 2024

Abstract

Recommendations play an undeniable role in decision-making. The empirical literature argues that recommendation can influence demand through two distinct channels: i) by enlarging awareness (attention channel), or ii) by altering preferences (utility channel). In this paper, we develop a framework to study these two channels using a parsimonious parametric model. We show that the model can produce various documented phenomena such as the Choice Effect, the Positive Effect of Negative Publicity, and the Spillover Effect. In addition, we offer simple and intuitive behavioral postulates characterizing our model so that one can test it. We offer unique identification under minimal data requirements. This enables us to measure the degree to which each channel affects choice behavior and to make out-of-sample predictions for counterfactual analysis for policy design purposes. Lastly, we apply our model in an auction setting to determine which alternatives to be recommended.

Keywords: Recommendation, Revealed Preference, Attention, Spillover

1 Introduction

Recommendation is one of the key determinants in decision-making nowadays. For instance, we constantly rely on recommendations from our friends, consumer reports, and mass media when select-

*We thank Jose Apesteguia, Miguel Ballester, Peter Caradonna, Chris Chambers, Kyle Hyndman, Ryo Kambayashi, Kohei Kawaguchi, Jay Lu, Paola Manzini, Marco Mariotti, Francesca Molinari, Hiroki Nishimura, Kota Saito, Ke Shaowei, Chris Turansick, Ece Yegane for their comments.

[†]Naveen Jindal School of Management, University of Texas at Dallas. Email: paul.cheung@utdallas.edu

[‡]Department of Economics, University of Maryland. Email: yusufcan@umd.edu

ing a movie to see; a book to read; a car to buy; or a school to send our children.[1] As digital platforms and their offerings continue to grow, consumers face an unprecedented number of online product and service options. As a result, recommendations have become indispensable for online shopping. Many internet sites such as Amazon, Netflix, Spotify, Tripadvisor, and Facebook incorporate recommendation tools to help customers with the burden of choice. "*Amazon's Choice*" of Amazon, "*Superhost*" of Airbnb, and "*Top 10*" of Netflix are just a few examples of these recommendation tools.

Individual choices are directly influenced by what is recommended to them.[2] The evidence on recommendations influencing choice behavior is conclusive across a wide spectrum of economic activities.[3] One of the main findings of this literature is called the choice effect, an increase in the sales of recommended products (Goodman et al. (2013), Häubl and Trifts (2000), Rowley (2000), Senecal and Nantel (2004), and Vijayasarathy and Jones (2000, 2001)). Several empirical studies have suggested that increased sales of recommended alternatives can be attributed to one of two channels: utility/preference or attention. While Chen et al. (2019), Goodman et al. (2013), Gupta and Harris (2010), and Mayzlin and Shin (2011) suggest that recommendations' primary role is to inform customers about the existence and the availability of the products, Adomavicius et al. (2013, 2018), Bairathi et al. (2022), Cosley et al. (2003), Kawaguchi et al. (2021), and Kumar and Benbasat (2006) state that recommendations might also affect choices by influencing consumers' preferences. One of the puzzles in this literature is that it is unclear how to distinguish between these two channels. In this paper, we ask whether it is possible to determine whether a recommendation operates through the attention channel or the utility channel from *observed choices*. Distinguishing these two channels is essential for policymakers/firms designing effective recommendations, especially when contextual factors may influence the effects of recommendations on utility and attention in a distinct way.

In addition, the literature on consumer behavior highlights several phenomena beyond the choice effect, including the spillover effect, the sales effect, and the intriguing positive impact of negative recommendations. The choice effect implies that sales of non-recommended products should decline. However, Bairathi et al. (2022) and Kawaguchi et al. (2021) observed a counterintuitive increase in demand for these products, a phenomenon termed *the spillover effect*. Additionally, research suggests that recommendations can broadly increase total sales, known as *the sales effect* (De et al. (2010) and Kawaguchi et al. (2021)). While these effects generally portray recommendations in a positive light,

[1] Some online service websites, such as Angi, HomeAdvisor, Houzz, Thumbtack recommend services for consumers' specific projects, which shows that people are willing to pay for recommendations.

[2] According to the 2013 data released by McKinsey & Company, recommendation systems drive 35% of purchases at Amazon. Similarly, 75% of what people watch on Netflix is initiated by their product recommendations.

[3] For instance, for the labor market, Horton (2017); for hospitality and tourism, Litvin et al. (2008); for music streaming service, Adomavicius et al. (2018) and Li et al. (2007); for e-commerce (of commodities or goods), Goodman et al. (2013), Häubl and Trifts (2000), Rowley (2000), Senecal and Nantel (2004), and Vijayasarathy and Jones (2000, 2001).

they can also be negative, like unfavorable product opinions. Contrary to the assumption that negative recommendations harm sales, studies by Allard et al. (2020), Berger et al. (2010), and Huang et al. (2023) reveal a paradox where negative publicity can positively affect consumer choices. This suggests that recommendations influence more than just product evaluations through the utility channel; they can also raise product awareness or accessibility, potentially leading to increased sales. Thus, it's vital to understand the dual pathways through which recommendations operate and to discern the significance of each channel in shaping consumer behavior.

While the practical importance of recommendations is evident, there is currently a lack of a comprehensive theoretical framework to analyze and comprehend their influence on consumer choices. This paper studies how recommendations influence the decision-making process. We employ revealed preference techniques to elucidate the choice process influenced by recommendations. Our goal is to explore the empirical ramifications of *aggregate* behavior within these models. More crucially, by perceiving choice behavior as an outcome of an unobservable cognitive process, we aim to identify the primitives of our model from the observed choices.

In this novel framework, we posit that recommendations are presented in their most basic form, where each decision-making scenario corresponds to a unique set of recommendations. In this context, the analyst lacks access to detailed qualitative information about the recommended products. This simplification is justified for two main reasons. Firstly, in real-life scenarios, generic recommendation indicators like "Best Seller," "Award Winner," and "Editor's Pick" are ubiquitous. Secondly, the analyst may not fully grasp how the decision-maker interprets these qualitative aspects of recommendations. Moreover, we assume that although different products may be recommended across various decision scenarios, the source and nature of these recommendations remain constant. This approach allows us to assess the decision-maker's subjective valuation of the recommended products.[4]

In our model, attention and utility play distinct roles in influencing choices. In the attention channel, recommending a particular product alters the likelihood that a customer is aware of the product. In the utility channel, recommendations affect the perceived valuation of a product even though she is aware of the product. We provide a parametric recommendation model in which both channels are present to illuminate the abovementioned puzzle. The parameters in our model have clear meanings, making it easier to understand and explain the relationships between them. In addition, our parametric model offers tractability in predicting consumer behavior and optimizing optimal recommendations. It helps to analyze and predict environmental changes based on both parameters. We also illustrate that our model is not only tractable but also uniquely identified.

[4]Furthermore, it enables us to contrast the effects of recommendations from different sources. See Section 5 for further discussion.

Here, we present a parametric recommendation model in which recommendations have the potential to impact both attention and utility. This particular feature is of paramount importance as it allows us to discern whether the influence of recommendations operates via the attention mechanism or the utility mechanism. Both the utility values and the degree of attention are contingent upon whether the product is recommended or not. In our model, without recommendation, each alternative enjoys a baseline utility value, denoted by u, and an attention parameter, denoted by γ. u represents a crude measure of the baseline utility value. $\gamma < 1$ measures the relative probability of being in the consideration set. The degree of attention is contingent upon whether the product is recommended or not.

When items are recommended, both the baseline utility and the attention parameters are allowed to change. However, the direction of change could be anything. To capture both the positive and negative effects of recommendation, we allow both positive and negative changes in the parameters' values. We denote the levels of attention and utility parameters of x receives with a recommendation as $\gamma'(x)$ and $u'(x)$, respectively. Hence, our utility and attention parameters are recommendation-dependent:

$$u_R(x) = \begin{cases} u'(x) & \text{if } x \in R \\ u(x) & \text{otherwise} \end{cases} \qquad \gamma_R(x) = \begin{cases} \gamma'(x) & \text{if } x \in R \\ \gamma(x) & \text{otherwise} \end{cases}$$

The choice probability of an alternative will be the share of the utility weight relative to the total utility weight in the consideration set. This is the same as Luce's formulation (Luce (1959)). The consideration set probabilities are calculated using the model of Manzini and Mariotti (2014), which is a parametric model of limited attention. Formally, the choice probability of x given recommendation R under the utility channel is given by[5]

$$\rho_S(x, R) = \sum_{x \in A \subseteq S} \underbrace{\Big[\prod_{y \in A} \gamma_R(y) \prod_{z \in S \setminus A} (1 - \gamma_R(z)) \Big]}_{\substack{\text{Prob. of } A \text{ being} \\ \text{the consideration set}}} \underbrace{\frac{u_R(x)}{u_R(A)}}_{\substack{\text{Prob. of } x \text{ being} \\ \text{chosen in } A}}$$

In section 3, we showcase how this simple model accommodates various observed behaviors, such as the choice effect, the spillover effect, the sales effect, and the positive effect of negative recommendation. First, in our model, the choice effect can be both positive or negative depending on the effect of the recommendation. If the recommendation increases both u and γ, then the choice effect

[5] Kawaguchi et al. (2021) illustrates that a similar model can be estimated and tested by using empirical data from beverage vending machines. Unlike our approach, their estimation is based on exclusion restrictions on menu-related variables.

is always positive. We show that both the choice effect and the spillover effect could exist in our model. In addition, while a negative review might decrease the perceived utility ($u'(x) < u(x)$), it might increase the awareness of the product ($\gamma'(x) > \gamma(x)$). Hence, our model predicts that negative publicity is more likely to harm products with high γ (Chevalier and Mayzlin (2006)). On the other hand, negative publicity could result in higher sales for products that are relatively unfamiliar (low γ).

While our model naturally accommodates observed behaviors documented in the literature, it also uncovers further implications that warrant exploration. In section 3, we bring attention to a particularly intriguing implication of our model. Specifically, we investigate some implications of positive recommendations, where both attention and utility exhibit enhancements ($u'(x) > u(x)$ and $\gamma'(x) > \gamma(x)$). We show that it is theoretically possible for the sales of a recommended product to fall compared to the case without any recommendations. We call such behavior as the crowding-out effect. To our knowledge, no empirical study has explored this behavior. We anticipate that our theoretical model will serve as a compass, guiding empirical researchers to explore these types of implications of our model in real-world scenarios.

In Section 4, we establish that three simple behavioral postulates, namely Positivity, R-IIA, and Consistency, fully characterize this model. These axioms are applied to a novel construct based on the choice data. Our characterization results do not require knowledge of the complete dataset. While augmenting parametric attention models (Brady and Rehbeck (2016) and Manzini and Mariotti (2014)) with parametric utility, we do not require any additional richness assumptions regarding set availability beyond those already assumed in the literature for the characterization to hold. Moreover, we also characterize two special cases of the model where the effects of the recommendation on attention and utility are positive. This sheds light on the relationship between recommendation *returns* and the size of the available alternatives.

In Section 5, we demonstrate how to *evaluate* recommendations. First, we illustrate how to uniquely identify the change in utility and attention from choice data. Once the parameters are identified, it can still be puzzling for the recommender to determine which channel — attention or utility — affects choices more. We demonstrate how to perform a counterfactual analysis by shutting down either channel. This analysis allows us to compare the marginal and conditional effects that each channel has on choices. Additionally, we show how one can compare across two different recommendation sources.

Lastly, in Section 6, to illustrate the tractability of our model, we study it in an auctioning environment. We analyze the second-price sealed bid auction and an auction akin to Vickrey-Clarke-Groves auction, deriving the equilibrium strategy in our settings. We illustrate the intricate relationship between the auctioneer's revenue and the underlying model's parameters through a comparative statics analysis, showing the distinct roles that attention and utility play in its effect on revenue. For instance, it shows the *complementary* and *substitute* relationships between the instrinsic attention level and the two channels of recommendation. The analysis also reveals the nature of revenue received by recommender — the recommender acts *as if* it takes up a share of the market by exploiting the extent to which recommendations can affect choices.

Literature Review

In our model, we utilize the idea that modelers can observe consumers' choice data as a function of their recommendation set. Coupled with the standard set-availability variation, we incorporate another layer of variation — recommendation, into the choice function. Therefore, one can regard the problem as a (stochastic) choice with frame (Salant and Rubinstein (2008)), where the frame is the subset of available alternatives. Indeed, our abstract setup even allows frame outside of recommendation, as long as they matter to both utility and attention.

We would like to mention a recent paper by Ke et al. (2021). While the focus of their paper is distinct from ours, their primitive enjoys a similar feature to our choice function. They propose a belief-updating model where a DM receives information from an unknown source. Their model differs from other updating rules where the posterior always belongs to the information set. Similar to our primitive, they allow posterior being outside of the information set (in our setting, recommendation set).

The classical Luce model attracts a variety of scholarly attention to developing different generalizations of it (e.g. Ahumada and Ülkü (2018), Echenique and Saito (2019), Echenique et al. (2018), Fudenberg et al. (2015), Gul et al. (2014), Kovach and Tserenjigmid (2022a, 2022b), and Tserenjigmid (2021)). All of these models involve different relaxations of the Luce IIA axiom. In our model, we impose the independence property onto a new object derived from choice data to provide a simple applicable, and tractable parametric model.

Lastly, there are several different strands of research departing from choice-set variation in the standard model. For example, some studies utilize list variation to study choices (e.g. Guney (2014) and Ishii et al. (2021)) and approval rates (Manzini et al. (2021)). Natenzon (2019) and Guney et al. (2018) study how non-choosable phantom options affect choices. In a similar spirit to our model, these lines of research are also augmenting the standard choice environment to enhance our understanding of human behavior.

2 Model

Let X be a non-empty finite set representing the set of feasible alternatives. Given a set of alternatives S, the decision maker receives a recommendation in the form of a set of alternatives, say $R \subset S$. Any subset of S could constitute a decision problem, including the empty set (no recommendation). While a recommendation can influence choices, it does not constrain them. To capture this, we define a probabilistic choice rule ρ_S to be a function of the recommendation set, R, but allow for $\rho_S(x, R)$ being positive for $x \in S \setminus R$. This assumption aims to capture some real-world environments where the collection of recommended sets is just a fraction of the entire product space. We also assume that the decision makers can always choose the outside option and buy nothing, denoted by o^*. Many models in marketing employ such alternatives. Let $\mathcal{R}_S$ be a subset of 2^S denoting all implementable recommendation sets when S is the set of feasible alternatives. Let the domain $\mathcal{D}$ (menus) be a collection of non-empty subsets of X. Therefore, $(\mathcal{D}, \{\mathcal{R}_S\}_{S\in\mathcal{D}})$ represents the available data for the analyst. We assume that the domain $\mathcal{D}$ satisfies a standard richness assumption in the parametric attention literature (Brady and Rehbeck (2016) and Manzini and Mariotti (2014)): $S \in \mathcal{D}$ whenever $|S| \leq 3$ and $T \in \mathcal{D}$ if $S \in \mathcal{D}$ and $T \subseteq S$. Also, we assume i) recommending none or any one alternative are always implementable: $R \in \mathcal{R}_S$ whenever $|R| \leq 1$ and $R \subseteq S$ and ii) implementable recommendation remains implementable under limited availability: $R \in \mathcal{R}_S$ implies $R \cap T \in \mathcal{R}_T$ for $T \subseteq S$. The following definition captures the choice rule under this framework.

Definition 1. Given the domain $\mathcal{D}$, a probabilistic choice rule $\{\rho_S\}$ is a collection of mappings from $\mathcal{R}_S$ to $\Delta(S \cup \{o^*\})$ such that $\sum_{x\in S\cup\{o^*\}} \rho_S(x, R) = 1$.

Here, we provide a parametric model of recommendation wherein recommendations can influence both attention and utility. This feature is crucial to identifying to what extent the effect of recommendations operates through the attention channel and/or the utility channel. Indeed, we will be able to distinguish this at the product level. In other words, it is possible that for one product recommendation product works more through the attention channel, for another one works more through the utility channel.

The attention channel is captured through a parametric consideration set model introduced by Manzini and Mariotti (2014). Whether a product is included in the consideration set is determined by the level of attention the decision-maker pays to it. This level depends on whether it is recommended or not. $\gamma(x)$ and $\gamma'(x)$ denote the attention levels product x receives with and without being recommended, respectively. These parameters are assumed to be between 0 and 1. If $\gamma(x) = \gamma'(x)$, recommending x does not influence choices through the attention channel. Otherwise, the attention channel's effect could be positive ($\gamma(x) < \gamma'(x)$) or negative ($\gamma(x) > \gamma'(x)$). Let γ_R represent utility values when R is the recommendation set.

$$\gamma_R(x) = \begin{cases} \gamma'(x) & \text{if } x \in R \\ \gamma(x) & \text{otherwise} \end{cases}$$

The key ingredient in our model is probabilistic consideration sets. Given a choice problem S, each subset of S can be a consideration set with a certain probability. Given S and R, the probability of $A \subseteq S$ being the consideration set is given by

$$\prod_{y \in A} \gamma_R(y) \prod_{z \in S \setminus A} (1 - \gamma_R(z))$$

Note that the empty set is also a consideration set, with the probability of its occurrence being $\prod_{z \in S} (1 - \gamma_R(z))$. In this scenario, the decision maker will choose the outside option.

To capture the utility channel, we utilize the multinomial logistic model introduced by Luce (1959). Similar to the attention channel, each product x is represented by two parameters representing utilities: $u(x) > 0$ and $u'(x) > 0$ where u' and u are crude measures of the utility value, representing utility with and without being recommended, respectively. We do not assume $u'(x) \geq u(x)$ in order to capture the potential negative effect of a recommendation on utility. Let u_R represent utility values when R is the recommendation set.

$$u_R(x) = \begin{cases} u'(x) & \text{if } x \in R \\ u(x) & \text{otherwise} \end{cases}$$

The probability of choosing a recommended item is proportional to its utility weight relative to the total utility weight of the considered alternatives. Thus, an alternative with a high u_R will be chosen more frequently than an alternative with a low u_R. For notational simplicity, we denote $u_R(A)$ as the shorthand for $\sum_{x \in A} u_R(A)$. Hence, the probability of choosing x for a given consideration set A is $\frac{u_R(x)}{u_R(A)}$. We are now prepared to formally define our parametric model.

Definition 2. A probabilistic choice rule $\{\rho_S\}$ has a parametric recommendation representation if there exists functions $u_R : X \to \mathbb{R}_{++}$ and $\gamma_R : X \to (0, 1)$ such that for $x \in X$,

$$\rho_S(x, R) = \sum_{x \in A \subseteq S} \underbrace{\left[\prod_{y \in A} \gamma_R(y) \prod_{z \in S \setminus A} (1 - \gamma_R(z)) \right]}_{\substack{\text{Prob. of } A \text{ being} \\ \text{the consideration set}}} \underbrace{\frac{u_R(x)}{u_R(A)}}_{\substack{\text{Prob. of } x \text{ being} \\ \text{chosen in } A}}$$

Our model encompasses two notable special cases. When $\gamma(x) = \gamma'(x)$, the recommendation solely influences utility. On the other hand, if $u(x) = u'(x)$, then only the attention channel is active. It's important to note that we do not presume any specific relationship between the parameters. Thus, our model accommodates the possibility that recommendations positively affect both attention and utility, i.e., $\gamma'(x) \geq \gamma(x)$ and $u'(x) \geq u(x)$ for all x. It also allows for scenarios where one effect is positive and the other negative, as well as cases where both are negative.

3 Accommodating Observed Behaviors

In this section, we investigate whether our model can accommodate behaviors observed in the recommendation literature. These include the choice effect, the spillover effect, the sales effect, negative publicity, and the crowding-out effect, all of which have been documented in various studies such as Allard et al. (2020), Bairathi et al. (2022), Berger et al. (2010), Chevalier and Mayzlin (2006), De et al. (2010), Goodman et al. (2013), Häubl and Trifts (2000), Huang et al. (2023), Kawaguchi et al. (2021), Rowley (2000), Senecal and Nantel (2004), and Vijayasarathy and Jones (2000, 2001).

Choice Effect. The impact of product recommendation can generally be ascribed to the direct effect on the recommended product (Goodman et al. (2013), Häubl and Trifts (2000), Rowley (2000), Senecal and Nantel (2004), and Vijayasarathy and Jones (2000, 2001)). The product recommendations have the potential to affect the sales of recommended products by changing consumer attention or altering consumers' utility. This influence on the sales of recommended products is termed the "choice effect" coined by Kawaguchi et al. (2021). Formally, the choice effect of recommending x in S is defined as $\rho_S(x, R \cup x) - \rho_S(x, R)$. In our model, the choice effect can be both positive or negative depending on the effect of the recommendation. We first investigate what different values of attention parameter $\gamma'(x)$ and utility $u'(x)$ would exhibit the positive choice effect. The next figure illustrates regions for both positive and negative choice effects.

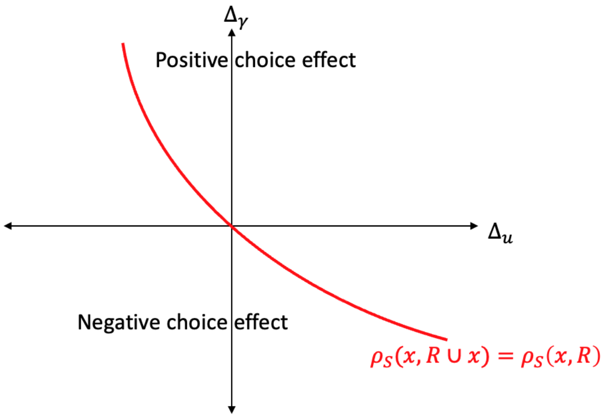


Figure 1. Choice Effect

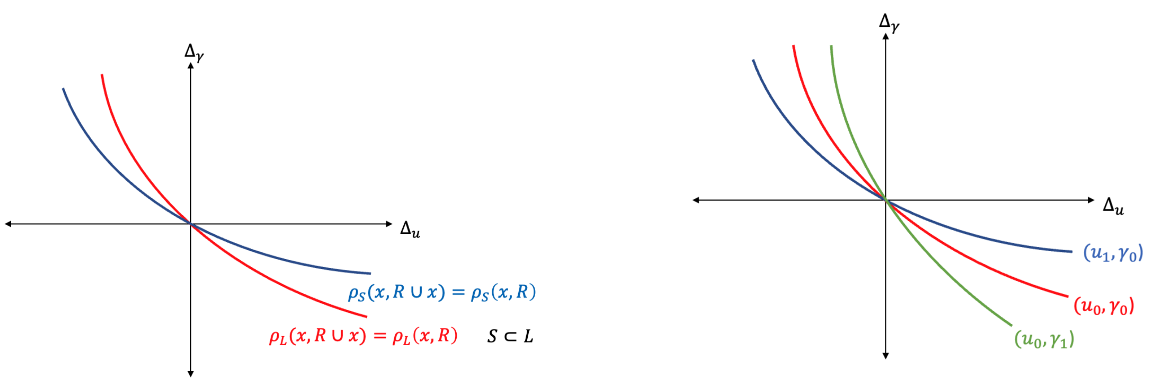
Refer to Figure 1, where the vertical axis denotes the change in the attention parameter ($\Delta_\gamma = \gamma'(x) - \gamma(x)$), and the horizontal axis represents the change in utility ($\Delta_u = u'(x) - u(x)$). At the origin, there is no recommendation effect, signifying that $\gamma'(x) = \gamma(x)$ and $u'(x) = u(x)$. The first (third) quadrant represents positive (negative) recommendations in both attention and utility, consistently belonging to the positive (negative) choice effect.

Conversely, the second quadrant (top-left) characterizes a region with a positive attention effect but a negative utility effect. Meanwhile, the fourth quadrant (bottom-right) denotes a region where the attention effect is negative, but the utility effect is positive. The boundary between the positive and negative choice effect regions necessarily lies in the second and fourth quadrants. This boundary is referred to as the zero-choice-effect curve, visually represented by the red line in Figure 1, specifically tailored for the choice set S and the recommendation set R excluding x.

Take a look at Figure 2a, where we explore the impact of varying choice sets. As the set of alternatives decreases, notice how the curve shifts towards the x-axis (depicted by the blue line). When faced with a negative Δ_u, a smaller Δ_γ is needed to counteract the adverse utility effect of recommendations. The attention effect is notably more potent in a smaller set, where product competition is less intense. To put it simply, as the choice set expands, the influence of attention weakens.

Now, turning our attention to Figure 2b, we observe the movement of the zero-choice-effect curve with respect to initial attention and utility parameters. Within this sub-figure, we maintain fixed choice and recommendation sets while altering the initial parameters for a specific product. Each zero-choice-effect curve is identified by its initial attention and utility parameter pair.

When the initial utility parameter is high ($u_1 > u_0$), the zero-choice-effect curve flattens. This occurs because a product with an exceptionally high initial utility experiences a less impactful boost from recommendations compared to a product with lower initial utility — a phenomenon akin to diminishing marginal utility. Similarly, if the initial attention parameter is high ($\gamma_1 > \gamma_0$), the zero-choice-effect curve steepens. A product with very high initial attention sees a diminished impact from an additional boost, reflecting the diminishing returns associated with heightened attention.



(a) Varying Choice set

(b) Varying Baseline Levels ($\gamma_0 < \gamma_1$, $u_0 < u_1$)

Figure 2. Comparative Choice Effect

Positive Effect of Negative Publicity. There is a stream of literature that looks into how negative publicity can have a positive effect on choice (e.g., Allard et al. (2020), Berger et al. (2010), and Huang et al. (2023)). If the recommendation operates solely through the utility channel, unfavorable information would typically lead to a decline in product evaluations. Therefore, the utility channel fails to account for instances where negative publicity could enhance sales. On the other hand, negative publicity might yield positive effects by amplifying product awareness or accessibility.

In our framework, we capture this effect by assuming that the recommendation generates a negative effect on u but a positive effect on γ. The negative publicity draws the DM's attention to the product, weakly increasing the probability that the alternative will be considered while also decreasing the valuation of the products. We illustrate that the impact of negative publicity is contingent on the initial product awareness. Our model predicts that negative publicity is more likely to harm products with high γ (Chevalier and Mayzlin (2006)) since there is not enough room to boost sales through the attention channel. On the other hand, negative publicity could result in higher sales for products that are relatively unfamiliar.

In the limited attention setting, as can be seen from Figure 1, for each choice set, there is a non-empty region in the top-left quadrant so that x can still have a higher probability of being chosen even if the utility effect is negative but the attention effect is positive. We put this observation in the following proposition.

Proposition 1 (Effect of Negative Publicity)**.** For every choice set S with recommendation R including x with $u(x) > u'(x)$, there exist two cut-offs k_1 and k_2 such that[6]

1) (Overall Negative) if $u'(x) - u(x) < k_1$, $\rho_S(x, R \setminus x) > \rho_S(x, R)$;

2) (Overall Positive) if $u'(x) - u(x) > k_1$, $\rho_S(x, R \setminus x) \leq \rho_S(x, R)$ if and only if $\gamma'(x) - \gamma(x) \geq k_2$.

The first part of the proposition states that if the effect of negative publicity is too high, the overall effect is always negative. In other words, the attention channel cannot compensate for the utility effect. The second part says that when negative publicity is not strong enough in utility, one can find a cutoff for the attention effect such that the overall effect is positive. Furthermore, as shown in Figure 2b, the negative effect on valuations can be easily canceled out if $\gamma(x)$ is relatively smaller. In other words, negative publicity can more easily generate a positive choice effect if the initial exposure of the product is low to begin with. All in all, we demonstrate that the effect of negative publicity depends on existing product awareness. Although negative publicity hurts products that already have broad awareness, it might help products that were relatively unknown.

[6]The cutoffs depend on the choice set, the recommendation set, and $u_R(x)$, $\gamma_R(x)$.

Spillover Effect. Empirical literature (e.g., Bairathi et al. (2022) and Kawaguchi et al. (2021)) has documented spillover effects — recommending certain alternatives might lead to an increase in sales of non-recommended products. One might think our model cannot capture such spillovers since recommendations only affect the attention and utility of the recommended product — an independence assumption embedded in the model. Nonetheless, it can provide one possible channel for spillover effects even if such independence is assumed. We illustrate this in the following figure.

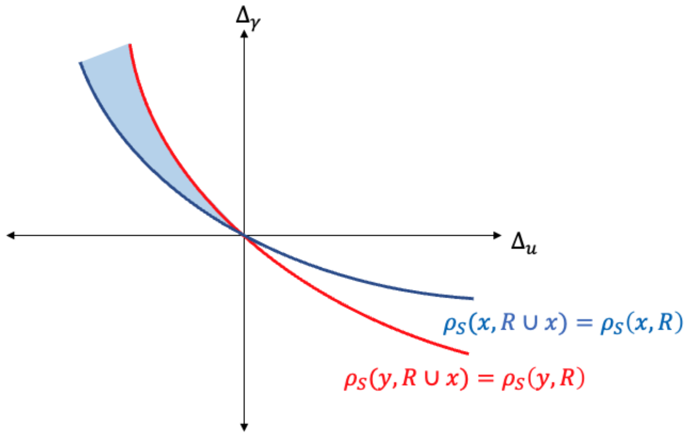


Figure 3. Spillover Effect

Figure 3 shows the zero-choice-effect curves for x and y. Both curves are drawn to capture the effect of recommending x. As before, the blue line represents the zero-choice-effect curves for x. The right side of this curve implies that the sales of x increased. On the other hand, the red line represents the zero-choice-effect curve for y when x is recommended. Since this curve represents the choice effect for an alternative other than x, the left side of this curve indicates that the choice probability of y increases as we recommend x. For example, if both attention and utility effects for x are negative, the choice probability of y must go up. Hence, contrary to the zero-choice-effect curve for x, everything to the right of the red curve will deliver a negative choice effect for y. This is because the more positively an alternative x is impacted by the recommendation, the more likely it is to steal market share from other alternatives in the set. Hence, everything to the left of the red curve will give a positive choice effect for y. Intriguingly, there exists a non-empty overlapping region for a positive choice effect on both x and y. This implies that recommending x has a positive spillover effect on both x and y.

Sales Effect. In our model, recommendations may prompt consumers to explore alternatives more frequently, potentially leading to an overall increase in total sales (see De et al. (2010) and Kawaguchi et al. (2021)). This impact of recommendations is termed the "sales effect," as coined by Kawaguchi et al. (2021). The sales effect occurs as a result of encouraging consumers to buy products they might not have otherwise considered. Formally, define the total sales as $1 - \rho_S(o^*, R) = 1 - \prod_{z \in S}(1 - \gamma_R(z))$.

One can show that $1 - \rho_S(o^*, R)$ is increasing if the recommendation has a positive effect on attention even when the additional recommended alternative lowers the perceived utility.

Crowding-out Effect. In this section, we would like to highlight an intriguing implication of our model. Even when focusing on positive recommendations, it is conceivable that the sales of a recommended product may fall compared to the case without any recommendations. To our knowledge, no empirical study has explored this behavior. We anticipate that our theoretical model will serve as a compass, guiding empirical researchers to explore this implication in real-world scenarios.

We now focus on positive recommendations where both attention and utility have positive effects. In this case, a product is always weakly more likely to be chosen when it is present in the recommendation set than when it is absent, i.e., for $x \in R$, $\rho_S(x, R) \geq \rho_S(x, R \setminus x)$. Having said that, it is not always better for an alternative to be included in the recommendation compared to no recommendation. This is called the crowding-out effect, $\rho_S(x, \emptyset) \geq \rho_S(x, R)$ and $\rho_S(x, T) \geq \rho_S(x, T \setminus x)$ for all $T \subset R$.

Our model can accommodate the crowding-out effect. To see this, consider two products x, y. We compare the choice probabilities for x under the recommendation set $\{x, y\}$ versus no recommendation in a choice set S. Imagine in some choice set S where only y is recommended. In this case, y will steal choice probabilities from *all* alternatives, including x, so that $\rho_S(x, \{y\}) < \rho_S(x, \emptyset)$. Consider that both $\gamma'(x) - \gamma(x)$ and $u'(x) - u(x)$ are almost close to zero, so that $\rho_S(x, \{x, y\}) \approx \rho_S(x, \{y\})$. Therefore, we can have $\rho_S(x, \{x, y\}) < \rho_S(x, \emptyset)$.

4 Characterization

We now explicitly investigate how one can identify whether the choice data align with our model. In other words, for a given ρ_S, what specific characteristics of these behaviors guarantee that the market behaves as if it is making choices in accordance with our model?

We first introduce a novel construct derived from the observed choice data. We then discuss the properties of this construction and relate them to the characterization of our model in the next section. For every $R \subseteq S$ and $x \in S$, we let

$$F_S(x, R) := \sum_{A: x \in A \subseteq S} (-1)^{|S \setminus A|} \rho_A(x, R \cap A) \frac{\prod_{z \in S \setminus A} (1 - \rho_z(z, R \cap z))}{\prod_{z \in S} \rho_z(z, R \cap z)}$$

Note that F_S is solely constructed from the choice data and, hence, is observable. We first investigate one key essential property of this function. In particular, we can show that the choice data can be written in terms of F_A's.

$$\rho_S(x,R) = \sum_{x\in A\subseteq S}\Big[\prod_{z\in A}\rho_z(z,R\cap z)\prod_{z\in S\setminus A}(1-\rho_z(z,R\cap z))\Big]F_A(x,R\cap A) \tag{1}$$

This result is due to the following lemma, a generalization of Mobius Inversion on Boolean algebra.

Lemma 1. *For real-valued functions f,g and h on 2^X,*

$$f(S) := \sum_{T\subseteq A\subseteq S}(-1)^{|S\setminus A|}g(A)h(S\setminus A) \text{ implies } g(S) = \sum_{T\subseteq A\subseteq S} f(A)h(S\setminus A)$$

where $h(A\cup B)=h(A)h(B)$ whenever $A\cap B=\emptyset$ and $h(\emptyset)=1$.

Finally, to see why (1) is true, let $T=\{x\}$, $f(S)=F_S(x,R)\prod_{z\in S}\rho_z(z,R\cap z)$, $g(A)=\rho_A(x,R\cap A)$ and $h(S\setminus A)=\prod_{z\in S\setminus A}(1-\rho_z(z,R\cap z))$. Note that for $A\cap B=\emptyset$, $h(A)h(B)=\prod_{s\in A}(1-\rho_z(z,R\cap z))\prod_{s\in B}(1-\rho_z(z,R\cap z))=\prod_{s\in A\cup B}(1-\rho_z(z,R\cap z))=h(A\cup B)$ and $h(\emptyset)=1$. Therefore, by applying Lemma 1 and rearrangement, the result follows.

Relationship to the model: Notice that (1) shares a surprisingly similar structure to that in the parametric recommendation representation. Notice that there is a summation over $x\in A\subseteq S$ and products over $z\in A$ and $z\in S\setminus A$. Indeed, if we assume the model is correct, the F function represents the choice probabilities conditional on full attention. Since our model becomes the Luce model conditional on full attention, F will deliver exactly the Luce utility ratio. Take a simple example where $R=\emptyset$, $S=\{x,y\}$ and x, by assuming that the data is represented by the model, we have

$$\begin{aligned} F_{\{x,y\}}(x,\emptyset) &= \rho_{\{x,y\}}(x,\emptyset)\frac{1}{\rho_{\{x\}}(x,\emptyset)\rho_{\{y\}}(y,\emptyset)} - \rho_{\{x\}}(x,\emptyset)\frac{1-\rho_{\{y\}}(y,\emptyset)}{\rho_{\{x\}}(x,\emptyset)\rho_{\{y\}}(y,\emptyset)} \\ &= \frac{\gamma(x)\gamma(y)\frac{u(x)}{u(x)+u(y)}+\gamma(x)(1-\gamma(y))}{\gamma(x)\gamma(y)} - \frac{1-\gamma(y)}{\gamma(y)} = \frac{u(x)}{u(x)+u(y)} \end{aligned}$$

Therefore, $F_{\{x,y\}}(x,\emptyset)$ captures the choice probability under full attention when S is a two-element set and no recommendation. Similarly, when x is recommended, $F_{\{x,y\}}(x,\{x\})$ is the choice probability of x conditional on the consideration set being $\{x.y\}$. Hence, we have

$$F_{\{x,y\}}(x,\{x\}) = \frac{u'(x)}{u'(x)+u(y)}$$

In general, for any $x\in R\subseteq S$, one can show that [7]

$$F_S(x,R) = \frac{u_R(x)}{u_R(S)} \tag{2}$$

[7] The proof is in Claim 1 in the Appendix.

The novelty of this approach is that i) we could separate the choice probabilities conditional on full attention (F's), and ii) we were able to define F's directly in terms of choice probabilities, rendering it an observable object in our framework.

4.1 Behavioral Postulates

With the novel construct of F, the behavioral postulates for the model become surprisingly simple, which in some ways reminisce the well-known characterization of the Luce model with respect to the choice probability. Here, we place the corresponding behavioral postulates on the function F instead of on ρ itself. First, since F captures the choice probability under full attention, our first axiom states that F must be positive.

Axiom 1. (Positivity) For $x \in S \supseteq R$, $F_S(x, R) > 0$.

The second axiom imposes an independence structure on F. In particular, the following states that whenever x and y do not change their status with respect to the recommendation sets R and R' while being available across choice sets T and S, then the ratio of F should not change.

Axiom 2. (R-IIA) For any $R \cup R' \subseteq S \cap T$ and $x, y \in S \cap T \setminus R \triangle R'$,

$$\frac{F_T(x, R)}{F_T(y, R)} = \frac{F_S(x, R')}{F_S(y, R')}$$

where $\triangle$ is the symmetric difference operator.

While the axiom resembles the Luce IIA, it encompasses two variations simultaneously. If one replaces F with the actual choice probability, fixing the recommendation set, this becomes the standard Luce IIA axiom; fixing the choice set, this becomes the Strong Luce-IIA axiom from the companion paper Cheung and Masatlioglu (2023). Lastly, since F is the choice probability under full attention, the total choice probability is 1.

Axiom 3. (Consistency) For $R \subseteq S$, $\sum_{x \in S} F_S(x, R) = 1$.

Axioms 1-3 directly pertain to the behavior of F in our model, focusing solely on utility values without explicitly addressing the attention channel. One might question whether additional axioms related to the attention channel are needed. Despite appearances that the axioms concern only utility values, the following theorem asserts that these axioms are both necessary and sufficient for the models to remain valid.

Theorem 1 (Characterization). A probabilistic choice rule $\{\rho_S\}$ satisfies Axioms 1-3 if and only if $\{\rho_S\}$ has a parametric recommendation representation.

As discussed above, the necessity of the proof can be easily shown by the fact that the function F satisfies (2) if the choice probabilities are represented by the model. Therefore, it boils down to showing that (2) holds in this case. On the other hand, for the sufficiency of the proof, we use a ratio of F to construct the utilities u_R by normalizing with respect to an arbitrary alternative and use singleton choice probabilities to define the attention parameters γ_R, where the latter ones are between 0 and 1 given the corresponding assumption on choice probabilities. Note that Axiom 1 ensures that the utilities u_R are also positive. Axiom 2 ensures that the constructed utilities u_R satisfy a ratio property with respect to F. Then, Axiom 3 helps us to restore (2) for the constructed utilities u_R and F. Finally, by applying (1), one can show that the choice probability can be represented by the model with the constructed utilities u_R and attention parameters γ_R.

4.2 Positive Attention and Utility

In this section, we investigate more closely the relationship between u' and u (and γ' and γ) and their implications on choice data. When there is positive attention and positive utility, the positive choice effect will always be true. As shown in Section 3, the positive choice effect may not hold if the effect on attention and utility goes in the opposite direction. Suppose we sometimes observe a positive choice effect in some cases but at the same time also observe a negative choice effect in some other cases for a particular product, we would like to ask whether the positive choice effect is driven by positive attention or positive utility (and vice-versa for the negative choice effect).

We first focus on utility. As shown in the previous discussion, we learn that F represents the conditional choice probability given full attention (i.e., equation (2)). Therefore, one can use F to derive the behavioral postulate for a positive utility effect. Nevertheless, intriguingly, the model enjoys an even simpler and more meaningful postulate for a positive utility effect without invoking F and focusing on choice data ρ alone. We first state the axiom below.

Axiom 4 (Increasing Recommendation Return to Size). For $x \in R \subseteq S \subseteq L$,

$$\frac{\rho_L(x, R)}{\rho_L(x, R \setminus x)} \geq \frac{\rho_S(x, R)}{\rho_S(x, R \setminus x)}$$

There are two (equivalent) ways to interpret this axiom. Firstly, for any S, one can define the percentage change in choice probability due to recommending x:

$$\frac{\rho_S(x, R) - \rho_S(x, R \setminus x)}{\rho_S(x, R \setminus x)}$$

This can be seen as the return of recommendation. Notice that the return is not necessarily positive; it can also be negative so that x suffers from a negative choice effect from recommendation. Therefore, Axiom 4 essentially compares the returns of recommendation across a smaller set S and a larger set L and states that the return must be weakly better as the set size increases.

Secondly, one can also re-arrange the condition in the axiom so that it requires

$$\frac{\rho_L(x,R)}{\rho_S(x,R)} \geq \frac{\rho_L(x,R\setminus x)}{\rho_S(x,R\setminus x)}$$

While not explicitly stated, the model respects regularity. That is, as the set size increases, each alternative is less likely to be chosen, and it is true regardless of whether it is recommended or not. That gives us $\frac{\rho_L(x,R)}{\rho_S(x,R)} \leq 1$ and $\frac{\rho_L(x,R\setminus x)}{\rho_S(x,R\setminus x)} \leq 1$. Therefore, this axiom essentially says that a product will be less affected by an increasing number of competitors (as measured by the rate of decrease in choice probabilities) when it is recommended than when it is excluded from the recommendation. Surprisingly, this axiom alone implies the positive utility effect. We state this result below.

Theorem 2 (Positive Utility Effect)**.** Let (u_R, γ_R) be a parametric recommendation representation of $\{\rho_S\}$. Then, $u'(x) \geq u(x)$ for all x if and only if $\{\rho_S\}$ satisfies Axiom 4.

The proof of this theorem relies on the observation that for $x \in R \subseteq S \subseteq L$ and $L \setminus S = \{z\}$,

$$\rho_L(x,R) = \gamma(z)\rho_S^*(x,R) + (1-\gamma(z))\rho_S(x,R)$$

where $\rho_S^*(x,R)$ are the same as $\rho_S(x,R)$ except in every denominator that involves the summation of u's we also add in $u(z)$. This expression allows us to connect $\frac{\rho_S^*(x,R)}{\rho_S^*(x,R\setminus x)}$ to $\frac{\rho_S(x,R)}{\rho_S(x,R\setminus x)}$ so that one can further deduce that the Axiom 4 holds if and only if $\frac{\rho_S^*(x,R)}{\rho_S^*(x,R\setminus x)} \geq \frac{\rho_S(x,R)}{\rho_S(x,R\setminus x)}$ if and only if $u'(x) \geq u(x)$.

Secondly, as is transparent in the model, in order to capture the positive attention effect, one can utilize the singleton choice probability. The next axiom states that x must enjoy the positive choice effect when it is the sole alternative in the choice set.

Axiom 5 (Positive Choice Effect at Singleton)**.** For all x, $\rho_x(x,x) \geq \rho_x(x,\emptyset)$.

With this axiom, we can state the following characterization.

Theorem 3 (Positive Attention Effect)**.** Let (u_R, γ_R) be a parametric recommendation representation of $\{\rho_S\}$. Then, $\gamma'(x) \geq \gamma(x)$ for all x if and only if $\{\rho_S\}$ satisfies Axiom 5.

Within the model, one can easily see that positive attention and positive utility will lead to a positive choice effect. Without assuming the model, interestingly, the behavioral postulates in this section are already sufficient for the positive choice effect. To see why it is true, we let $T = \{x\}$ in Axiom 4 and utilize both axioms, we get

$$\frac{\rho_S(x,R)}{\rho_S(x,R\setminus x)} \geq \frac{\rho_x(x,x)}{\rho_x(x,\emptyset)} \geq 1$$

so that $\rho_S(x,R) \geq \rho_S(x,R\setminus x)$. We put this observation in the following.

Observation 1 (Positive Choice Effect)**.** For any choice probabilities $\{\rho_S\}$, Axiom 4 and Axiom 5 imply a positive choice effect.

Lastly, one might be tempted to think that Axiom 4 and Axiom 5 are only reserved for positive utility and positive attention, respectively. Nevertheless, in fact, one can utilize Axiom 4 to reveal the recommendation effect on attention (and use Axiom 5 to reveal the recommendation effect on utility) in some cases. In particular, suppose that choice data occasionally satisfies positive choice effects in some choice sets but also violates Axiom 4 for some other choice sets, it immediately reveals that the positive choice effect is driven by positive attention but not positive utility. The same reasoning can be applied across three other scenarios: positive choice effect but violate Axiom 5; negative choice effect but (partially) satisfies Axiom 4; and negative choice effect but (partially) satisfies Axiom 5. Therefore, we put this observation in the following corollary.

Corollary 1 (Complementary Revelation)**.** Let (u_R, γ_R) be a parametric recommendation representation of $\{\rho_S\}$. Then, if for some S, L, S', R and R' such that $x \in R \subseteq S \subseteq L$ and $x \in R' \subseteq S'$
1) if $\rho_{S'}(x, R') > (<) \rho_{S'}(x, R' \setminus x)$ and $\frac{\rho_L(x,R)}{\rho_L(x,R \setminus x)} \leq (\geq) \frac{\rho_S(x,R)}{\rho_S(x,R \setminus x)}$, then $\gamma'(x) > (<) \gamma(x)$; and
2) if $\rho_{S'}(x, R') > (<) \rho_{S'}(x, R' \setminus x)$ and $\rho_x(x, x) \leq (\geq) \rho_x(x, \emptyset)$, then $u'(x) > (<) u(x)$.

5 Recommendation Evaluation

While the model is agnostic about the specific form of recommendation, when given choice data, one can utilize the model to identify how exactly their recommendation affects people's choices. It is of practical importance because it helps make an out-of-sample prediction and also informs the recommender of the nature of their recommendation through the lens of the model. Some forms of recommendation can be more attention-grabbing (e.g. recommending an alternative with shiny and enlarged text on a product webpage), and some forms of recommendation can be more persuasive (e.g. a detailed editorial recommendation). The model can help assess the recommender's presumption regarding how their recommendation affects choices.

Identification of parameters. The model enjoys a strong identification power. In particular, the attention parameters are uniquely identified, and the utility parameter is identified up to a scaling factor. We state this result in the following Proposition.

Proposition 2 (Uniqueness)**.** Suppose (u_R^1, γ_R^1) and (u_R^2, γ_R^2) are two parametric recommendation representations of the same choice probabilities, then $\gamma_R^1 = \gamma_R^2$ and $u_R^1 = \alpha u_R^2$ for some scalar $\alpha > 0$.

To identify the underlying attention parameter $\gamma_R(x)$, one can simply utilize the choice data, which can be identified from singleton choice probabilities. In particular, $\rho_x(x, R \cap x) = \gamma_R(x)$. On the other hand, one can utilize the F function to identify the underlying u_R. As in the standard Luce model, $u_R(x)$ can be identified up to a positive scaling transformation. Therefore, one can normalize a $u(x)$ and infer the rest from the F function. One caveat is that $u(x)$ and $u'(x)$ will never appear at the same time in F. Therefore, one needs to be cautious when performing the normalization. The proof for Theorem 1 provides a detailed construction for one set of u_R from choice data.

Relative Strength in Recommendation. Given the identified parameters, it can still be puzzling to the recommender *which* channel affects the choice probability *more*. While the two effects are essentially intertwined when entering into the parametric model, one can construct counterfactual choice data to infer and compare attention and utility for the degree to which they affect choice probabilities. In this section, we focus on the case that recommendation has a positive effect on attention and utility.

To achieve this, we construct two hypothetical (counterfactual) choice data. One represents the choice probability where x only receives attentional influence from recommendation, denoted by $\rho_S^A(x, R)$, and the other represents the choice probability where x only receives utility influence from recommendation, denoted by $\rho_S^U(x, R)$. In particular, we define, for $x \in R \subseteq S$,

$$\rho_S^A(x, R) := \sum_{x \in A \subseteq S} \Big[\prod_{y \in A} \gamma_R(y) \prod_{z \in S \setminus A} (1 - \gamma_R(z))\Big] \frac{u(x)}{u(x) + u_R(A \setminus x)}$$

$$\rho_S^U(x, R) := \sum_{x \in A \subseteq S} \Big[\gamma(x) \prod_{y \in A \setminus x} \gamma_R(y) \prod_{z \in S \setminus A} (1 - \gamma_R(z))\Big] \frac{u_R(x)}{u_R(A)}$$

Note that the assumption of positive attention and utility dictates that $\rho_S^A(x, R), \rho_S^U(x, R) \in [\rho_S(x, R \setminus x), \rho_S(x, R)]$. Therefore, these two counterfactual data allow for two different decompositions of the effects of attention and utility – marginal effect vs conditional effect on choice. The marginal effect compares the choice probability before the recommendation to the hypothetical scenario where x only receives one of the effects. On the other hand, the conditional effect compares the choice probability after the recommendation to the hypothetical scenario where x only receives the other effects. In other words, conditional on the scenario that the marginal effect from the other channel is in place, it looks at how much more increase in choice probability the target channel will bring. These are displayed in Table 1 for the attention and utility channels, respectively.

	Attention	Utility
Marginal Effect	$\rho_S^A(x, R) - \rho_S(x, R \setminus x)$	$\rho_S^U(x, R) - \rho_S(x, R \setminus x)$
Conditional Effect	$\rho_S(x, R) - \rho_S^U(x, R)$	$\rho_S(x, R) - \rho_S^A(x, R)$

Table 1. Marginal and Conditional Effect on Choice from Attention vs Utility

To compare attention and utility, we pair up the effects in the respective domains (marginal or conditional). Interestingly, it is easy to see that the marginal utility effect is greater than the marginal attention effect if and only if $\rho_S^U(x,R) \geq \rho_S^A(x,R)$ if and only if the conditional utility effect is greater than the conditional attention effect. Therefore, it suffices to only compare across $\rho_S^A(x,R)$ and $\rho_S^U(x,R)$. Moreover, one can show that this holds if and only if

$$\frac{\rho_S(x,R)}{\rho_S(x,R\setminus x)} \geq \left(\frac{\rho_x(x,x)}{\rho_x(x,\emptyset)}\right)^2$$

We put this observation in the following proposition.

Proposition 3 (Relative Strength). Let (u_R, γ_R) be a parametric recommendation representation of $\{\rho_S\}$. Then, the utility effects are greater than the attention effects for alternative x at set S under recommendation set R if and only if $\frac{\rho_S(x,R)}{\rho_S(x,R\setminus x)} \geq \left(\frac{\rho_x(x,x)}{\rho_x(x,\emptyset)}\right)^2$.

Proposition 3 also speaks to Axiom 4 and Axiom 5 in a very specific way. Notice that from Axiom 4, we know that $\frac{\rho_S(x,R)}{\rho_S(x,R\setminus x)} \geq \frac{\rho_x(x,x)}{\rho_x(x,\emptyset)}$ due to the increasing recommendation return to set size. On the other hand, by Axiom 5, $\frac{\rho_x(x,x)}{\rho_x(x,\emptyset)} \geq 1$ so that it must be $(\frac{\rho_x(x,x)}{\rho_x(x,\emptyset)})^2 \geq \frac{\rho_x(x,x)}{\rho_x(x,\emptyset)}$. Therefore, there is no deterministic answer from the model to tell which one gives a greater effect– it is left to be determined in actual choice data.

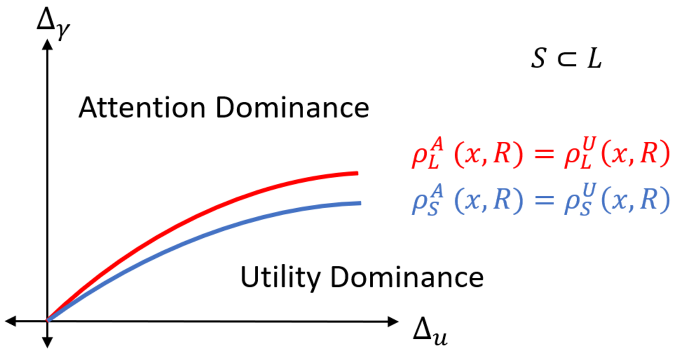


Figure 4. Marginal and Conditional Effect on Choice from Attention vs Utility Channels

One can visualize the dominating effects in Figure 4. Similar to previous figures, the vertical axis represents the change in attention parameter, and the horizontal axis denotes the change in utility for the alternative x. The red line shows the values of Δ_y and Δ_x so that the attention effect and utility effect are equal. There are two things to note. Firstly, the curves are upward-sloping so that an increase in utility can be compensated by an increase in attention. Secondly, as the utility effect dominates the attention effect in the smaller set S (below the blue line), it will continue to be so in the bigger set L (below the red line). Therefore, it implies a single-crossing property with respect to the subset relation. This condition is also true in the other direction. We put this observation in the following corollary.

Corollary 2 (Single-Crossing)**.** Let (u_R, γ_R) be a parametric recommendation representation of $\{\rho_S\}$. If the utility effects are greater than the attention effects for alternative x under recommendation set R at set S, then the same holds for any set $L \supseteq S$; If the attention effects are greater than the utility effects for alternative x under recommendation set R at set L, then the same holds for any set $S \subseteq L$ provided that $R \subseteq S$.

This result is essentially driven by Proposition 3 and Axiom 4. One can easily see that if ever it is true that $\frac{\rho_S(x,R)}{\rho_S(x,R\backslash x)} \geq \left(\frac{\rho_x(x,x)}{\rho_x(x,\emptyset)}\right)^2$, then it must be that $\frac{\rho_L(x,R)}{\rho_L(x,R\backslash x)} \geq \frac{\rho_S(x,R)}{\rho_S(x,R\backslash x)} \geq \left(\frac{\rho_x(x,x)}{\rho_x(x,\emptyset)}\right)^2$. The other direction is also immediately implied.

Comparing Two Recommendation Sources. We have thus far operated under the assumption that the underlying data-generating process adheres to the same recommendation source. While the preceding sections have elucidated the methods for evaluating a single recommendation source, a natural question arises: how can we effectively compare and contrast disparate recommendation sources? In this section, we investigate the choice behaviors for the same group of individuals when exposed two different recommendation sources.

To accomplish this, we assume that there are two choice data sets, denoted by ρ^i for $i = 1, 2$, which are represented by two parametric recommendation models with (u^i_R, γ^i_R). In particular, we denote, for $i = 1, 2$,

$$u^i_R(x) = \begin{cases} u'_i(x) & \text{if } x \in R \\ u_i(x) & \text{otherwise} \end{cases} \qquad \gamma^i_R(x) = \begin{cases} \gamma'_i(x) & \text{if } x \in R \\ \gamma_i(x) & \text{otherwise} \end{cases}$$

where the superscript i becomes subscript to avoid conflict in notation. Since the choice data are generated through the same group of individuals, we assume that the underlying utility before recommendation is the same across the two groups. Therefore, we let $u_1 = u_2 =: u$.[8] The discrepancy in u'_1 and u'_2 will reflect the differences in persuasiveness for the two sources. Moreover, we impose no restrictions on the relationships between γ^1_R and γ^2_R, allowing for the possibility that distinct sources may present products in different manners, resulting in diverse attention parameters.

We say that the source i is more persuasive than source j if $u'_i \geq u'_j$. It is noteworthy that the distinctions in persuasiveness can be revealed by employing the novel constructs F as defined in

[8]While the u's are unique to a scaling factor, the assumption here is behaviorally equivalent to assuming that the choice probability under full attention without recommendation is the same across two different sources. We then further impose the same normalization.

Section 4. This stems from the fact that the F function encapsulates the conditional choice probability under full attention. We denote F^i as the corresponding function F for source i for $i = 1, 2$. We put this observation in the following corollary.

Corollary 3 (Persuasiveness Comparison)**.** The source i is more persuasive than the source j if and only if $F^i_S(x, \{x\}) \geq F^j_S(x, \{x\})$ for all i and some sets S.

Since the underlying utilities are the same across the two different sources, recommending one alternative will deliver the choice probabilities where only one term differs. To see this, note that $F^i_S(x, \{x\}) = \frac{u'_i(x)}{u'_i(x)+u(S \backslash x)}$, and the result immediately follows. Note again that the F function is based entirely on choice probabilities, so that this provides a test of persuasiveness only by using choice data ρ^i and ρ^j.

6 Application: Auctioning Recommendation

In this section, we apply our representation in an auction problem for a recommendation. There is a wealth of literature studying position auctions (e.g. Aggarwal et al. (2006), Edelman et al. (2007), and Varian (2007)), which well-known search engines use for sponsored search results. Our problem is different from that framework in several perspectives. Firstly, our problem focuses on recommendation rather than listing alternatives, so that it can be applied to scenarios where a recommendation does not involve a rank or order among alternatives. For example, a company may wish to bid for a detailed editorial recommendation on a platform or strive to be recognized with certain recommendation signage (e.g. Amazon Choice, Best Seller, and so on). Secondly, our model deals with the intricate issues of the effect of recommendation on both attention and utility rather than assuming an order effect for click-through rate. Thirdly, our model speaks directly to the market share of a product as a result of recommendations rather than narrowing down to a fixed value for each click-through. The last distinction has a technical consequence on our problem so that existing techniques on position auction, to the best of our knowledge, are inapplicable in our setup.

Suppose that there are one recommender and N single-product sellers so that $|S| = N$ and Seller x sells product $x \in S$ to passive consumers. Suppose the recommender can choose to auction $K < N$ slots of recommendations. With recommendation $R \subseteq S$, the market share for a seller x, $\rho_S(x, R)$, is determined by an underlying parametric recommendation model with positive effects on attention and utility. Since we fixed the choice set S, we omit S in our notation and write $\rho(x, R)$ for brevity.

We parameterize the effect on attention and utility from the recommendation. For attention, we assume that

$$\gamma'(x) = (1 - \alpha)\gamma(x) + \alpha$$

Here, $\alpha \in [0, 1]$ represents the effectiveness of the recommendation platform on the attention channel. If $\alpha = 1$, then the recommendation is fully effective independent of the initial attention and $\gamma'(x) = 1$. If $\alpha = 0$, then the recommendation does not affect the initial attention parameter, $\gamma'(x) = \gamma(x)$. Therefore, $\gamma'(x) \in [\gamma(x), 1]$. α is fixed and affects all alternatives monotonically, i.e., $\gamma(x) \geq \gamma(y)$ then $\gamma'(x) \geq \gamma'(y)$. Note that α also has a probabilistic interpretation: Suppose the population will (independently) consider a recommended alternative with probability α, then the probability that x will be considered is exactly $\gamma'(x)$ if x is recommended. To see this, suppose x is recommended. Then, α fraction of the population will consider it immediately. Given the event that x is not considered when it is recommended (with probability $(1 - \alpha)$), $\gamma(x)$ fraction of the people will consider it, delivering $\gamma'(x) = (1 - \alpha)\gamma(x) + \alpha$.

On the other hand, for utility, we assume that $u'(x) = (1 + \beta)u(x)$. Likewise, $\beta \in [0, \infty)$ can be regarded as the measure of effectiveness of the recommendation on the utility channel. If $\beta = 0$, then the recommendation has no utility effect, i.e., $u'(x) = u(x)$. Therefore, $u'(x) \in [u(x), \infty)$. Here, the recommendation effect on utility depends on the baseline utility: The higher the baseline utility, the greater the effect in the utility channel. Therefore, it assumes a percentage effect in utility.

In the following, we investigate the recommendation auction problem using two different classes of games: one is the classic second-price auction, and the other is the game of recommendation bid schedule. The first one is applied to a case with two players, while the second one is applied to more general cases with N players and multiple winners. It is worth noting that even in the position auction, practitioners have not consistently used the same auction formats. For example, Google has transitioned from a version of a (generalized) second-price auction to a first-price auction (Paes Leme et al. (2020)). Given that our problem differs from the standard framework, the present study aims to examine the benchmark cases to illuminate the structure of the problem and demonstrate how the auctioneer can extract revenue from its power in redistributing market share through recommendations.

Two alternatives with one slot. Let $|S|= 2$ with generic element x, y. We first analyze the recommendation problem with a second-price sealed bid auction: Each seller $x \in S$ submits a bid $b(x)$. Then, the top bids win the auction, and the winner pays the second-highest bid. In the event of a tie, the recommender randomly and evenly selects winners. Conditional on $R = \{x\}$ or $\{y\}$ winning the auction, the seller x payoff can be written as

$$\pi(x, R) := \begin{cases} \rho(x, R) - b(y) & \text{if } R = \{x\} \\ \rho(x, R) & \text{if } R = \{y\} \end{cases}$$

On the other hand, the recommender revenue is simply the second-highest bid. In the following, for illustrative purposes, we focus on a game of complete information where the market shares under each realization of recommendation are common knowledge.

One can show that the bidding strategy $b^{SP}(x) := \rho(x,x) - \rho(x,y)$ is a weakly dominant strategy, similar to the classic reasoning for a second-price auction for a single item. The same reasoning also applies to y due to symmetry. Therefore, $b^{SP}(y) := \rho(y,y) - \rho(y,x)$ is also a weakly dominant strategy. In our setting, one can show that $b^{SP}(x) = b^{SP}(y)$. The reasoning is that the choice probability for the outside option is the same across recommendation set $\{x\}$ and $\{y\}$, due to the independent assumption embedded in α. To see this,

$$\rho(x,x) + \rho(y,x) = 1 - (1-\alpha)(1-\gamma(x))(1-\gamma(y)) = \rho(x,y) + \rho(y,y)$$

Therefore, by rearrangement, we have $b^{SP}(x) = b^{SP}(y)$. The two sellers will have an equal chance of winning and the recommendation slot will be allocated randomly across x and y. In either case, the (equilibrium) recommender revenue will be $b^{SP}(x)$. In the following, we distinguish between the two channels of recommendation and their corresponding impact on revenue.

Proposition 4 (Comparative Statics on Revenue)**.** For a second-price sealed bid recommendation auction with two options and one slot,

a) (only utility channel) suppose that $\alpha = 0$ and $\beta > 0$,
    (i) the revenue is increasing in $\gamma(x)$,
    (ii) the revenue is increasing in $u(x)$ when $u(x) < u(y)$ and is decreasing in $u(x)$ when $u(x) > u(y)$.
b) (only attention channel) suppose that $\alpha > 0$ and $\beta = 0$,
    (i) the revenue is decreasing in $\gamma(x)$,
    (ii) the revenue is increasing in $u(x)$ when $\gamma(x) < \gamma(y)$ and is decreasing in $u(x)$ when $\gamma(x) > \gamma(y)$.

There are several things to highlight. Note that the baseline attention level, $\gamma(x)$, has a different impact on revenue depending on whether the recommendation affects attention or utility. Firstly, the recommendation through the utility channel ($\alpha = 0$ and $\beta > 0$) has a *complementary* relationship with baseline attention level (i.e. Proposition 4a(i))– The higher the baseline attention level is, the more valuable the recommendation is to the sellers, which in turn becomes the revenue for the recommender. Secondly, the recommendation through the attention channel ($\alpha > 0$ and $\beta = 0$) has a *substitute* relationship with baseline attention level (i.e. Proposition 4b(i)), where the recommendation becomes less valuable, the higher the initial attention level. Indeed, consider if $\gamma(x) = 1$, compared to no recommendation, there is no gain in market share from recommendation, so the seller will have less incentive to bid for recommendation.[9]

[9]Nevertheless, it does not mean that seller x will not bid anything: it will still bid a positive amount to try to prevent y

Moreover, the baseline utility level has a more intriguing relationship with the two channels of recommendation– It reflects the *combative* nature in terms of utility level in the market share model. Firstly, as the recommendation affects only through the utility channel ($\alpha = 0$ and $\beta > 0$), the revenue essentially decreases in the absolute difference between the baseline utility level of x and y (i.e. $|u(x) - u(y)|$) (Proposition 4a(ii)).[10] In other words, the closer the two baseline utility levels get, the higher the incentive for the sellers to bid for (utility-based) recommendations to beat their component to get a higher market share. Secondly, considering only the attention channel ($\alpha > 0$ and $\beta = 0$), Proposition 4b(ii) says that the marginal gain in revenue from an increase in baseline utility of x will always be positive (resp. negative) as long as the attention level for x is less (resp. greater) than y. The effect of that is zero when the two baseline attention levels are equal.

N alternatives with K slots. Let $|S|= N$ with generic element x, y. In this section, we focus on the payment received by the recommender by proposing a game that induces a strategy similar to the Vickrey-Clarke-Groves mechanism. As we will show later, in our setting with $N = 2$ and $K = 1$, the second-price seal-bid auction essentially implements the same outcome as in this game. However, this is no longer true for $N > 2$. We investigate this game as an ideal scenario for the VCG bidding strategies to hold, serving as a benchmark for illuminating the structure of the problem and how the underlying parameters change the revenue for the recommender.

Consider the games as follows: Each seller submits a bid schedule for each realization of winnings. Therefore, a seller might need to pay even if they do not "win" the auction. This can capture scenarios such as the lesser of the evil so that a seller can pay to support other sellers to win to prevent a bigger steal of market share from some other sellers. After submitting the bid, the recommendar picks the winning R with the highest combined submitted bid. Denote the bid of a seller x for bidding outcome R as $b(x, R)$. Conditional on R winning the auction, the seller x payoff can be written as

$$\pi(x, R) = \rho(x, R) - b(x, R)$$

Therefore, the recommender revenue, conditional on a winning bid, denoted by $\Pi(K, R)$, is simply

$$\Pi(K, R) = \sum_{x \in S} b(x, R)$$

We call this game as the game of recommendation bid schedule. In the following, similar to the last section, for illustrative purposes, we will assume complete information so that the market shares under each realization of recommendation are common knowledge.

from getting the recommendation slot.

[10]Here, we hold $u(y)$ fixed.

There is one bidding strategy that maps out the classic payment transfer under the VCG mechanism. In particular, it is

$$b^{VCG}(x,R) := \max_{B\in S_K} \sum_{y\in S\setminus x} \rho(y,B) - \sum_{y\in S\setminus x} \rho(y,R)$$

where S_K is the collections of subsets of S with the size K. We call this the VCG bidding strategy. In the language of VCG mechanism, one can understand that this amounts to the (maximum) total loss in payoff for all other sellers (i.e. $S \setminus x$) as a result of the winning of R. Notice that this game's VCG bidding strategy mimics the equilibrium outcome in the second-price sealed bid auction in the last section. To see this, let $S = \{x, y\}$. First, the seller x pays 0 when they lose, which matches $b^{VCG}(x,y) = \rho(y,y) - \rho(y,y) = 0$. Second, the seller x pays the second highest bid $b^{SP}(y) = \rho(y,y) - \rho(y,x)$ when they win, which is exactly $b^{VCG}(x,x)$.

For an arbitrary set S with K slots, the VCG bidding strategies constitute an equilibrium for this game of recommendation bid schedule.

Proposition 5 (VCG as Nash)**.** In the game of recommendation bid schedule, the VCG bidding strategies profile is a Nash Equilibrium.

We first focus on the following two results to see why this is true.

Proposition 6 (Recommendation-independent Payoff)**.** In the game of recommendation bid schedule with K slots and VCG bidding strategies,
1) The seller payoff is independent of R.
2) The recommendar revenue is independent of R.

Note that it is straightforward to see that 1) in Proposition 6 holds. One reason is similar to the last section: The choice probability for the outside option is the same across recommendation sets with the same size K. Notice that by a little rearrangement, the seller x payoff $\pi(x,R)$ is equal to

$$\begin{aligned}
\rho(x,R) - b^{VCG}(x,R) &= \sum_{y\in S} \rho(y,R) - \max_{B\in S_K} \sum_{y\in S\setminus x} \rho(y,B) \\
&= 1-(1-\alpha)\prod_{y\in S}(1-\gamma(y)) - \max_{B\in S_K}\Big[1-(1-\alpha)\prod_{y\in S}(1-\gamma(y)) - \rho(x,B)\Big] \\
&= \min_{B\in S_K} \rho(x,B)
\end{aligned}$$

Notice that $\min_{B\in S_K} \rho(x,B)$ is independent of R, which proves $1)$. Note that this result also demonstrates that the VCG strategies essentially push every seller to get the lowest possible payoff when some alternatives are recommended. With this discovery, the second result is also straightforward.

Note that the recommender revenue $\Pi(K, R)$ is equal to

$$\sum_{x \in S} b^{VCG}(x, R) = \sum_{x \in S} \rho(x, R) + \sum_{x \in S} -(\rho(x, R) - b^{VCG}(x, R))$$
$$= \underbrace{1 - (1 - \alpha) \prod_{y \in S} (1 - \gamma(y))}_{\text{Available Surplus}} - \underbrace{\sum_{x \in S} \min_{B \in S_K} \rho(x, B)}_{\text{Redistribution to Sellers}}$$

Both terms in the expression are independent of R, which proves 2). Note that this expression also spells out the nature of the payment received by the recommender, which is the difference between the total available surplus and the redistribution of payoff to sellers. Note that if the recommendation has no effect on choices (i.e. $\alpha = \beta = 0$), the revenue will be exactly zero. To see this, the available surplus will become the (joint) choice probabilities for all alternatives in S as if there is no recommendation, which exactly equals the second terms where the (joint) choice probabilities are the same across recommendation sets. Therefore, the bigger impact the recommendation has on choice probabilities, the more revenue the recommender can extract from the sellers. One can re-arrange equation which further accentuates the role that recommender play in this game

$$\underbrace{\pi(S, R)}_{\text{Sellers' Payoff}} + \underbrace{\Pi(K, R)}_{\text{Recommender's Revenue}} + \underbrace{\overbrace{\rho(o^*, \emptyset)}^{\text{Original Uncaptured}} - \overbrace{\alpha\rho(o^*, \emptyset)}^{\text{Expansion in Shares}}}_{\text{New Uncaptured}} = 1$$

In the equation, the four terms sum to one, which represents the total share in the market. The recommender, by auctioning recommendation, acts *as if* he own shares of the market. When recommendation has a positive effect on attention (where α is positive), the recommendation expands the total available market shares (a.k.a the *Sales Effect* in Section 3) so that the uncaptured market share decreases. In this game, the recommender revenue is always greater than the expansion in market share. To see this, note that

$$\Pi(K, R) - \alpha\rho(o^*, \emptyset) = \sum_{x \in S} \Big(\rho(x, \emptyset) - \min_{B \in S_K} \rho(x, B) \Big)$$

which is greater than zero as long as α or β is non-zero. Therefore, the recommender not only claims the expansion in market share as its own revenue but also reaps benefits from its power over the redistribution of market share among sellers through recommendation.

We then discuss why the VCG strategy profile is a Nash Equilibrium in this game (Proposition 5). Notice that all recommendations set R have an equal chance of winning since the total bid for each recommendation set will be the same. Suppose a seller deviates and bids slightly higher for one recommendation set R. It is for sure that R will win. However, the payoff will be strictly lower than the expected payoff from a random realization of any recommendation set. On the other hand, if a seller shades any one bid for an arbitrary recommendation set R, although it is for sure that R will not win, the seller still ends up with the same expected payoff with the remaining randomly realized recommendation set. Lastly, one may wonder whether a seller can uniformly shade all bids by a tiny amount so that they can enjoy a higher expected payoff. However, in VCG bidding strategies, for every x, there is always one R such that $b^{VCG}(x, R) = 0$. Therefore, this already imposes a binding floor constraint, making uniform shading impossible.

Lastly, we discuss its implication on the optimal recommendation. Notice that for any given number of slots for the recommendation, Proposition 6 informs us that the winning set of alternatives does not matter to the revenue. Yet, we have not introduced the cost side of the recommendation. There are several reasons why recommendations can be costly. Besides potential materialistic costs associated with a recommendation, reputation costs are an issue. If recommendations increase the (perceived) instantaneous consumption utility (utility channel) for an alternative, but later on the consumers regret doing so, there can be future repercussions to the recommender. Likewise, the same reasoning applies if the consumers are led towards (attention channel) a lower-quality alternative due to recommendations but later find out better options are also available. One can suspect there can be a heterogeneous cost function in products, denoted by $c(x)$. Therefore, the aim of the recommender will be to design an auction mechanism that can maximize the total profit by subtracting the cost in recommendation $\sum_{x \in R} c(x)$. For example, the recommender can have a reserve price for winning the recommendation for different alternatives. Indeed, even in position auction such as Google Ads, Merchants are not only ranked by based on bid, but also "relevance" (Aggarwal et al. (2006)). This section serves to demonstrate the nature of the revenue that can be generated by recommendation and show how different channels can affect the revenue. Designing an optimal mechanism is beyond the scope of this paper but future research is greatly encouraged in this direction.

7 Conclusion

In our daily lives, recommendations play an integral and pervasive role. This study delves into the nuanced dynamics of recommendation influence on decision-making, examining two distinct channels: attention and utility. The framework introduces a probabilistic choice model, interpretable through the lens of aggregated choice data. To ensure practicality and broad applicability, we opt for a parametric model. This characterization facilitates a comprehensive differentiation between the attention and utility effects of recommendations.

While our models adeptly explore two key channels through which recommendations shape choices, we acknowledge the potential for refinement or generalization based on specific needs and circumstances. The versatility of our framework accommodates the examination of additional operational channels for recommendations. Future research prospects encompass the development of more intricate models, inclusive of strategic recommendations, limited consideration, status quo, behavioral search, satisficing, and temptation. This paper not only lays the foundation for substantial future research but also provides a bridge, extending economic wisdom from conventional models to this nuanced setting.

References

Adomavicius, G., Bockstedt, J. C., Curley, S. P., & Zhang, J. (2013). Do recommender systems manipulate consumer preferences? A study of anchoring effects. *Information Systems Research*, *24*(4).

Adomavicius, G., Bockstedt, J. C., Curley, S. P., & Zhang, J. (2018). Effects of Online Recommendations on Consumersâ Willingness to Pay. *Information Systems Research*, *29*(1), 84–102.

Aggarwal, G., Goel, A., & Motwani, R. (2006). Truthful auctions for pricing search keywords. *Proceedings of the 7th ACM Conference on Electronic Commerce*, 1â7.

Ahumada, A., & Ülkü, L. (2018). Luce rule with limited consideration. *Mathematical Social Sciences*, *93*, 52–56.

Allard, T., Dunn, L. H., & White, K. (2020). Negative reviews, positive impact: Consumer empathetic responding to unfair word of mouth. *Journal of Marketing*, *84*(4), 86–108.

Bairathi, M., Zhang, X., & Lambrecht, A. (2022). The value of platform endorsement. *Working Paper*.

Berger, J., Sorensen, A. T., & Rasmussen, S. J. (2010). Positive effects of negative publicity: When negative reviews increase sales. *Marketing Science*, *29*(5), 815–827.

Brady, R. L., & Rehbeck, J. (2016). Menu-Dependent Stochastic Feasibility. *Econometrica*, *84*(3), 1203–1223.

Chen, Y., Lu, Y., Wang, B., & Pan, Z. (2019). How do product recommendations affect impulse buying? An empirical study on WeChat social commerce. *Information and Management*, *56*(2).

Cheung, P., & Masatlioglu, Y. (2023). Decision making with recommendation. *working paper*.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, *43*(3), 345–354.

Cosley, D., Lam, S. K., Albert, I., Konstan, J. A., & Riedl, J. (2003). Is seeing believing? How recommender interfaces affect users' opinions. *Conference on Human Factors in Computing Systems - Proceedings*.

De, P., Hu, Y., & Rahman, M. S. (2010). Technology usage and online sales: An empirical study. *Management Science*, *56*(11), 1930–1945.

Echenique, F., & Saito, K. (2019). General Luce model. *Economic Theory*, *68*(4), 811–826.

Echenique, F., Saito, K., & Tserenjigmid, G. (2018). The perception-adjusted Luce model. *Mathematical Social Sciences*, *93*, 67–76.

Edelman, B., Ostrovsky, M., & Schwarz, M. (2007). Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *American Economic Review*, *97*(1), 242–259.

Fudenberg, D., Iijima, R., & Strzalecki, T. (2015). Stochastic Choice and Revealed Perturbed Utility. *Econometrica*, *83*(6), 2371–2409.

Goodman, J. K., Broniarczyk, S. M., Griffin, J. G., & McAlister, L. (2013). Help or hinder? When recommendation signage expands consideration sets and heightens decision difficulty. *Journal of Consumer Psychology*, *23*(2), 165–174.

Gul, F., Natenzon, P., & Pesendorfer, W. (2014). Random Choice as Behavioral Optimization. *Econometrica*, *82*(5), 1873–1912.

Guney, B. (2014). A theory of iterative choice in lists. *Journal of Mathematical Economics*, *53*, 26–32.

Guney, B., Richter, M., & Tsur, M. (2018). Aspiration-based choice. *Journal of Economic Theory*, *176*.

Gupta, P., & Harris, J. (2010). How e-WOM recommendations influence product consideration and quality of choice: A motivation to process information perspective. *Journal of Business Research*, *63*(9-10), 1041–1049.

Häubl, G., & Trifts, V. (2000). Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Science*, *19*(1), 4–21.

Horton, J. J. (2017). The Effects of Algorithmic Labor Market Recommendations: Evidence from a Field Experiment. *Journal of Labor Economics*, *35*(2), 345–385.

Huang, J., Xie, R., Cao, Q., Shen, H., Zhang, S., Xia, F., & Cheng, X. (2023). Negative can be positive: Signed graph neural networks for recommendation. *Information Processing Management*, *60*(4), 103403.

Ishii, Y., Kovach, M., & Ülkü, L. (2021). A model of stochastic choice from lists. *Journal of Mathematical Economics*, 102509.

Kawaguchi, K., Uetake, K., & Watanabe, Y. (2021). Designing Context-Based Marketing: Product Recommendations Under Time Pressure. *Management Science, 67*(9).

Ke, S., Wu, B., & Zhao, C. (2021). Learning from a Black Box. *Working Paper*.

Kovach, M., & Tserenjigmid, G. (2022a). Behavioral foundations of nested stochastic choice and nested logit. *Journal of Political Economy*, *130*(9), 2411–2461.

Kovach, M., & Tserenjigmid, G. (2022b). The focal luce model. *American Economic Journal: Microeconomics*, *14*(3), 378–413.

Kumar, N., & Benbasat, I. (2006). Research note: The influence of recommendations and consumer reviews on evaluations of websites. *Information Systems Research*, *17*(4), 425–439.

Li, Q., Myaeng, S. H., & Kim, B. M. (2007). A probabilistic music recommender considering user opinions and audio features. *Information Processing and Management*, *43*(2), 473–487.

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management*, *29*(3), 458–468.

Luce, R. D. (1959). *Individual choice behavior*. Wiley, New York.

Manzini, P., & Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica, 82*(3), 1153–1176.

Manzini, P., Mariotti, M., & Ulku, L. (2021). Sequential Approval: A Model of âLikesâ, Paper Downloads and Other Forms of Click Behaviour.

Mayzlin, D., & Shin, J. (2011). Uninformative advertising as an invitation to search. *Marketing Science*, *30*(4), 666–685.

Natenzon, P. (2019). Random choice and learning. *Journal of Political Economy*, *127*(1), 419–457.

Paes Leme, R., Sivan, B., & Teng, Y. (2020). Why do competitive markets converge to first-price auctions? *Proceedings of The Web Conference 2020*, 596â605.

Rowley, J. (2000). Product searching with shopping bots. *Internet Research*, *10*(3), 203–214.

Salant, Y., & Rubinstein, A. (2008). (A, f): Choice with Frames1. *The Review of Economic Studies*, *75*(4), 1287–1296.

Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*, *80*(2), 159–169.

Tserenjigmid, G. (2021). The order-dependent luce model. *Management Science*, *67*(11), 6915–6933.

Varian, H. R. (2007). Position auctions. *International Journal of Industrial Organization*, *25*(6), 1163–1178.

Vijayasarathy, L. R., & Jones, J. M. (2000). Print and Internet catalog shopping: Assessing attitudes and intentions. *Internet Research*, *10*(3), 191–202.

Vijayasarathy, L. R., & Jones, J. M. (2001). Do internet shopping aids make a difference? an empirical investigation. *Electronic Markets*, *11*(1), 75–83.

Appendix

Proof for Proposition 1

Proof. We let $f : [0, u(x)) \times [0, 1 - \gamma(x)) \to \mathbb{R}$ such that

$$f(a,b) := (\gamma(x) + b) \sum_{x \in A \subseteq S} \prod_{y \in A \setminus x} \gamma_R(y) \prod_{z \in S \setminus A} (1 - \gamma_R(z)) \frac{u(x) - a}{u(x) - a + u_R(A \setminus x)}$$

Therefore, $f\Big(u(x) - u'(x), \gamma'(x) - \gamma(x)\Big) = \rho_S(x, R)$ and $f(0,0) = \rho_S(x, R \setminus x)$.

Note that it is easy to see that f is decreasing in a and increasing in b. By continuity of the function f, we solve a implicit function $f(a,b) = f(0,0)$ for b in terms of a so that it defines a function B

$$B(a) = \frac{f(0,0)}{\sum_{x \in A \subseteq S} \prod_{y \in A \setminus x} \gamma_R(y) \prod_{z \in S \setminus A} (1 - \gamma_R(z)) \frac{u(x) - a}{u(x) - a + u_R(A \setminus x)}} - \gamma(x)$$

Note that as a tends to $u(x)$, $B(a)$ tends to infinity, which falls outside of the domain of b. Therefore, combined with the fact that $f(a,b)$ is decreasing in a, there exists a cutoff l_1 so that if $a > k_1$, we must have $f(a,b) < f(0,0)$. In other words, there exists cutoff $k_1 = -l_1$ if $u(x) - u'(x) < k_1$, $\rho_S(x, R \setminus x) > \rho_S(x, R)$.

Secondly, note that if $a < k_1$, we have $f(a,b) \geq f(0,0)$ if and only if $b \geq B(a)$, since the function f is increasing in b. In other words, there exists a cutoff $k_2 = B(u(x) - u'(x))$ such that $\rho_S(x, R \setminus x) \leq \rho_S(x, R)$ if and only $\gamma'(x) - \gamma(x) \geq k_2$. The proof is complete. ■

Proof for Lemma 1

Proof.

$$\begin{aligned}
\sum_{T \subseteq A \subseteq S} f(A) h(S \setminus A) &= \sum_{T \subseteq A \subseteq S} h(S \setminus A) \Bigg(\sum_{T \subseteq B \subseteq A} (-1)^{|A \setminus B|} g(B) h(A \setminus B) \Bigg) \\
&= \sum_{T \subseteq A \subseteq S} \sum_{T \subseteq B \subseteq A} (-1)^{|A \setminus B|} g(B) h(S \setminus A) h(A \setminus B) \\
&= \sum_{T \subseteq A \subseteq S} \sum_{T \subseteq B \subseteq A} (-1)^{|A \setminus B|} g(B) h(S \setminus B) \\
&= g(S) h(S \setminus S) = g(S)
\end{aligned}$$

The last line is by the implications of $h(\emptyset) = 1$ and by the fact that $\sum_{T \subseteq A \subseteq S} \sum_{T \subseteq B \subseteq A} (-1)^{|A \setminus B|} \varphi(B) = \varphi(S)$ for any real-valued function φ on 2^X. ■

Proof for Theorem 1

Proof. We first show Necessity. We first show that F represents the choice probability under full attention.

Claim 1. *For any $R \subseteq S$ and $x \in S$,*

$$F_S(x, R) = \frac{u_R(x)}{u_R(S)}$$

Proof. We prove by induction on the size of S. For $|S|= 1$ and $R \subseteq S$, it is easy to see that $LHS = F_S(x, R) = \frac{\rho_x(x,R)}{\rho_x(x,R\cap x)} = 1 = RHS$. Assume it is true for any size less than $n = |S|$. Note that by utilizing Equation (1), we know that

$$\rho_S(x, R) = \sum_{x\in A\subseteq S} F_A(x, R \cap A) \prod_{z\in S\setminus A} (1 - \rho_z(z, R \cap z)) \prod_{z\in A} \rho_z(z, R \cap z)$$

By re-arrangement, we have

$$\begin{aligned}
F_S(x, R) \prod_{z\in S} \rho_z(z, R \cap z) &= \rho_S(x, R) - \sum_{x\in A\subset S} F_A(x, R \cap A) \prod_{z\in S\setminus A} (1 - \rho_z(z, R \cap z)) \prod_{z\in A} \rho_z(z, R \cap z) \\
&= \frac{u_R(x)}{u_R(S)} \prod_{z\in S} \gamma_R(z) + \sum_{x\in A\subset S} \Big\{ \prod_{z\in A} \gamma_R(z) \prod_{z\in S\setminus A} (1 - \gamma_R(z))\Big\} \frac{u_R(x)}{u_R(A)} \\
&- \sum_{x\in A\subset S} F_A(x, R \cap A)\Big\{ \prod_{z\in A} \gamma_R(z) \prod_{z\in S\setminus A} (1 - \gamma_R(z))\Big\} \\
&= \frac{u_R(x)}{u_R(S)} \prod_{z\in S} \gamma_R(z) + \sum_{x\in A\subset S} \prod_{y\in S\setminus A} (1 - \gamma_R(y)) \prod_{y\in A} \gamma_R(y)(\frac{u_R(x)}{u_R(A)} - F_A(x, R \cap A)) \\
&= \frac{u_R(x)}{u_R(S)} \prod_{z\in S} \gamma_R(z)
\end{aligned}$$

The second term before the last equality vanishes because of the induction hypothesis. Therefore,

$$F_S(x, R) = \frac{u_R(x)}{u_R(S)} \frac{\prod_{y\in S} \gamma_R(y)}{\prod_{z\in S} \rho_z(z, R \cap z)} = \frac{u_R(x)}{u_R(S)}$$

The proof is complete.

■

Hence, we have, for $R \cup R' \subseteq S \cap T$ where x, y do not change its status across R, R'

$$\frac{F_S(x, R)}{F_S(y, R)} = \frac{u_R(x)}{u_R(y)} = \frac{u_{R'}(x)}{u_{R'}(y)} = \frac{F_T(x, R')}{F_T(y, R')}$$

Axiom 2 is proven. Lastly, for Axiom 3, by Claim 1, we know that

$$\sum_{x\in S} F_S(x, R) = \sum_{x\in S} \frac{u_R(x)}{u_R(S)} = 1$$

Therefore, Axiom 3 is proven.

For sufficiency, firstly, we define $\gamma(x) := \rho_x(x, \emptyset)$ and $\gamma'(x) := \rho_x(x, x)$. We designate an arbitrary $z_0 \in X$ where we define $u(z_0) := 1$. Then we define for $x \neq z_0$,

$$u(x) := \frac{F_{x,z_0}(x, \emptyset)}{F_{x,z_0}(z_0, \emptyset)} \text{ and } u'(x) := \frac{F_{x,z_0}(x, x)}{F_{x,z_0}(z_0, x)}$$

Note that $u(x) > 0$ and $u'(x) > 0$ by Axiom 1. To define $u'(z_0)$, we designate an arbitrary $w_0 \neq z_0$ so that $u'(z_0) := u(w_0)\frac{F_{z_0,w_0}(z_0,z_0)}{F_{z_0,w_0}(w_0,z_0)}$. The next claim shows that the anchors can indeed be arbitrary.

Claim 2. *For any R and $x, y \in S$,*

$$\frac{u_R(x)}{u_R(y)} = \frac{F_S(x, R)}{F_S(y, R)}$$

Proof. In the following, we utilize Axiom 2. We first consider $x, y \in S \setminus z_0$. There are three cases. Firstly, if $x, y \notin R$, then

$$\frac{u_R(x)}{u_R(y)} = \frac{F_{x,z_0}(x, \emptyset)/F_{x,z_0}(z_0, \emptyset)}{F_{y,z_0}(y, \emptyset)/F_{y,z_0}(z_0, \emptyset)} = \frac{F_{x,y,z_0}(x, \emptyset)/F_{x,y,z_0}(z_0, \emptyset)}{F_{x,y,z_0}(y, \emptyset)/F_{x,y,z_0}(z_0, \emptyset)} = \frac{F_S(x, R)}{F_S(y, R)}$$

If $x \in R, y \notin R$, then

$$\frac{u_R(x)}{u_R(y)} = \frac{F_{x,z_0}(x, x)/F_{x,z_0}(z_0, x)}{F_{y,z_0}(y, \emptyset)/F_{y,z_0}(z_0, \emptyset)} = \frac{F_{x,y,z_0}(x, x)/F_{x,y,z_0}(z_0, x)}{F_{x,y,z_0}(y, x)/F_{x,y,z_0}(z_0, x)} = \frac{F_S(x, R)}{F_S(y, R)}$$

If $x, y \in R$, then

$$\frac{u_R(x)}{u_R(y)} = \frac{F_{x,z_0}(x, x)/F_{x,z_0}(z_0, x)}{F_{y,z_0}(y, y)/F_{y,z_0}(z_0, y)} = \frac{F_{x,y,z_0}(x, \{x, y\})/F_{x,y,z_0}(z_0, \{x, y\})}{F_{x,y,z_0}(y, \{x, y\})/F_{x,y,z_0}(z_0, \{x, y\})} = \frac{F_S(x, R)}{F_S(y, R)}$$

Then, consider $z_0 = x \notin R$. There are two cases. If $y \notin R$, then

$$\frac{u_R(x)}{u_R(y)} = \frac{1}{F_{y,z_0}(y, \emptyset)/F_{y,z_0}(z_0, \emptyset)} = \frac{F_S(z_0, R)}{F_S(y, R)}$$

If $y \in R$, then

$$\frac{u_R(x)}{u_R(y)} = \frac{1}{F_{y,z_0}(y, y)/F_{y,z_0}(z_0, y)} = \frac{F_S(z_0, R)}{F_S(y, R)}$$

Lastly, consider $z_0 = x \in R$. We have

$$\begin{aligned}\frac{u_R(x)}{u_R(y)} = \frac{u'(z_0)}{u_{R\setminus w_0}(w_0)} \frac{u_{R\setminus w_0}(w_0)}{u_{R\setminus w_0}(y)} &= \frac{F_{z_0,w_0}(z_0, z_0)}{F_{z_0,w_0}(w_0, z_0)} \frac{u_{R\setminus w_0}(w_0)}{u_{R\setminus w_0}(y)} \\ &= \frac{F_{S\cup w_0}(z_0, R \setminus w_0)}{F_{S\cup w_0}(w_0, R \setminus w_0)} \frac{F_{S\cup w_0}(w_0, R \setminus w_0)}{F_{S\cup w_0}(y, R \setminus w_0)} = \frac{F_S(z_0, R)}{F_S(y, R)}\end{aligned}$$

The second fraction after the third equality sign is given by the "partially proven" Claim 2 where w_0 is not z_0 and is not in the recommended set and y can be either recommended or not recommended (and it is not z_0). Therefore, the claim is proved. ■

The next claim is that the construction matches one key property of F.

Claim 3. *For any S, $F_S(x,R) = \frac{u_R(x)}{u_R(S)}$.*

Proof. Note that by Claim 2 and Axiom 3, $\frac{u_R(x)}{u_R(S)} = \frac{1}{\sum_{y\in S}\frac{u_R(y)}{u_R(x)}} = \frac{1}{\sum_{y\in A}\frac{F_S(y,R)}{F_S(x,R)}} = F_S(x,R)$. ■

Lastly, we consider the formula for the choice probability generated with the defined primitives above, by Equation 1 and Claim 3, we have

$$\begin{aligned}
&\sum_{x\in A\subseteq S}\Big\{\prod_{y\in A}\gamma_R(y)\prod_{z\in S\setminus A}(1-\gamma_R(z))\Big\}\frac{u_R(x)}{u_R(A)}\\
=&\sum_{x\in A\subseteq S}F_A(x,R\cap A)\prod_{z\in S\setminus A}(1-\rho_z(z,R\cap z))\prod_{z\in A}\rho_z(z,R\cap z)\\
=&\rho_S(x,R)
\end{aligned}$$

Hence, the proof is complete. ■

Proof for Theorem 2

Proof. We first prove the following lemma.

Lemma 2. *Let $\{A_n\}$, $\{B_n\}$ be sequences of positive real numbers. Let k, k' be positive real numbers. Then, for $x > 0$, the function*

$$L_n(x) := \frac{\sum_{n=1}^N B_n\frac{1}{A_n+k'+x}}{\sum_{n=1}^N B_n\frac{1}{A_n+k+x}}$$

is increasing in x if and only if $k' \geq k$.

Proof. We take a derivative with respect to x. Then, we get (we skip some subscript and superscript in the summation when confusion is unlikely)

$$\begin{aligned}
\frac{\partial L_n(x)}{\partial x} &= \frac{1}{(\sum\frac{Bn}{A_n+k+x})^2}\Big[-\sum\frac{B_n}{A_n+k+x}(\sum\frac{B_n}{(A_n+k'+x)^2})+\sum\frac{B_n}{(A_n+k+x)^2}(\sum\frac{B_n}{(A_n+k'+x)})\Big]\\
&= \frac{1}{(\sum\frac{Bn}{A_n+k+x})^2}\sum_{i=1}^N\sum_{j=1}^N\Big[\frac{B_i}{(A_i+k+x)^2}\frac{B_j}{A_j+k'+x}-\frac{B_i}{A_i+k+x}\frac{B_j}{(A_j+k'+x)^2}\Big]\\
&= \frac{1}{(\sum\frac{Bn}{A_n+k+x})^2}\sum_{i=1}^N\sum_{j=1}^N\frac{B_iB_j}{\Big((A_i+k+x)(A_j+k'+x)\Big)^2}\Big[(k'-k)+(A_j-A_i)\Big]
\end{aligned}$$

It remains to show that $k' > k$ if and only if

$$\sum_{i=1}^N\sum_{j=1}^N\frac{B_iB_j}{\Big((A_i+k+x)(A_j+k'+x)\Big)^2}\Big[(k'-k)+(A_j-A_i)\Big] > 0$$

Let $C_{ij} := \frac{B_iB_j}{\Big((A_i+k+x)(A_j+k'+x)\Big)^2}\Big[(k'-k)+(A_j-A_i)\Big]$. Note that it is easy to see that $C_{ii} > 0$ if and only if $k' > k$. For $i \neq j$, we consider

$$\begin{aligned}C_{ij} + C_{ji} =& B_iB_j(k'-k)\Big[\frac{1}{\Big((A_i+k+x)(A_j+k'+x)\Big)^2} + \frac{1}{\Big((A_j+k+x)(A_i+k'+x)\Big)^2}\Big] \\ &+ B_iB_j(A_j-A_i)\Big[\frac{1}{\Big((A_i+k+x)(A_j+k'+x)\Big)^2} - \frac{1}{\Big((A_j+k+x)(A_i+k'+x)\Big)^2}\Big]\end{aligned}$$

Note that the first part is greater than zero if and only if $k' > k$. On the other hand, the second part is equal to

$$\begin{aligned}=& B_iB_j(A_j-A_i)\Big[\frac{1}{\Big((A_i+k+x)(A_j+k'+x)\Big)} + \frac{1}{\Big((A_j+k+x)(A_i+k'+x)\Big)}\Big]\times \\ &\Big[\frac{1}{\Big((A_i+k+x)(A_j+k'+x)\Big)} - \frac{1}{\Big((A_j+k+x)(A_i+k'+x)\Big)}\Big] \\ =& B_iB_j(A_j-A_i)^2\Big[\frac{1}{\Big((A_i+k+x)(A_j+k'+x)\Big)} + \frac{1}{\Big((A_j+k+x)(A_i+k'+x)\Big)}\Big]\times \\ &\Big[\frac{k'-k}{\Big((A_i+k+x)(A_j+k'+x)\Big)\Big((A_j+k+x)(A_i+k'+x)\Big)}\Big]\end{aligned}$$

which is greater than zero if and only if $k' > k$. Therefore, we have shown that $\sum_{i=1}^N\sum_{j=1}^N C_{ij} > 0$ if and only if $k' > k$. The proof is complete. ■

Let (u_R, γ_R) be a parametric recommendation representation of $\{\rho_S\}$. Without loss of generality, we consider $T = S \setminus t$ for some $t \in X$. Then, we let $x \in R \subseteq T \subset S$, $\frac{\rho_S(x,R)}{\rho_S(x,R\setminus x)} \geq \frac{\rho_T(x,R)}{\rho_T(x,R\setminus x)}$ where $T = S \setminus t$. We define

$$\rho_T^t(x,R) := \sum_{x\in A\subseteq T}\Big[\prod_{y\in A}\gamma_R(y)\prod_{z\in T\setminus A}(1-\gamma_R(z))\Big]\frac{u_R(x)}{u_R(A)+u_t}$$

Therefore, $\frac{\rho_S(x,R)}{\rho_S(x,R\setminus x)} \geq \frac{\rho_T(x,R)}{\rho_T(x,R\setminus x)}$ if and only if

$$\frac{\gamma(t)\rho_T^t(x,R)+(1-\gamma(t))\rho_T(x,R)}{\gamma(t)\rho_T^t(x,R\setminus x)+(1-\gamma(t))\rho_T(x,R\setminus x)} \geq \frac{\rho_T(x,R)}{\rho_T(x,R\setminus x)}$$

One can show that the above holds if and only if

$$\frac{\rho_T^t(x,R)}{\rho_T^t(x,R\setminus x)} \geq \frac{\rho_T(x,R)}{\rho_T(x,R\setminus x)}$$

By canceling some $\gamma'(x)$, $\gamma(x)$, $u'(x)$ and $u(x)$ from both sides we get

$$\frac{\sum_{x\in A\subseteq T}\frac{\Big[\prod_{y\in A\setminus x}\gamma_R(y)\prod_{z\in T\setminus A}(1-\gamma_R(z))\Big]}{u_R(A\setminus x)+u'(x)+u(t)}}{\sum_{x\in A\subseteq T}\frac{\Big[\prod_{y\in A\setminus x}\gamma_{R\setminus x}(y)\prod_{z\in T\setminus A}(1-\gamma_{R\setminus x}(z))\Big]}{u_{R\setminus x}(A\setminus x)+u(x)+u(t)}} \geq \frac{\sum_{x\in A\subseteq T}\frac{\Big[\prod_{y\in A\setminus x}\gamma_R(y)\prod_{z\in T\setminus A}(1-\gamma_R(z))\Big]}{u_R(A\setminus x)+u'(x)}}{\sum_{x\in A\subseteq T}\frac{\Big[\prod_{y\in A\setminus x}\gamma_{R\setminus x}(y)\prod_{z\in T\setminus A}(1-\gamma_{R\setminus x}(z))\Big]}{u_{R\setminus x}(A\setminus x)+u(x)}}$$

Since $u(t)$ is arbitrary, the above holds if and only if the function

$$L(c) := \frac{\sum\limits_{x\in D\subseteq T} \frac{\left[\prod\limits_{y\in D\setminus x}\gamma_R(y)\prod\limits_{z\in T\setminus D}(1-\gamma_R(z))\right]}{u_R(D\setminus x)+u'(x)+c}}{\sum\limits_{x\in D\subseteq T} \frac{\left[\prod\limits_{y\in D\setminus x}\gamma_{R\setminus x}(y)\prod\limits_{z\in T\setminus D}(1-\gamma_{R\setminus x}(z))\right]}{u_{R\setminus x}(D\setminus x)+u(x)+c}}$$

is increasing in c for $c > 0$. Note that $\left[\prod\limits_{y\in D\setminus x}\gamma_R(y)\prod\limits_{z\in T\setminus D}(1-\gamma_R(z))\right] = \left[\prod\limits_{y\in D\setminus x}\gamma_{R\setminus x}(y)\prod\limits_{z\in T\setminus D}(1-\gamma_{R\setminus x}(z))\right]$ and $u_R(D\setminus x) = u_{R\setminus x}(D\setminus x)$ since x is extracted from these expressions.

Finally, by choosing $N = |\{D : x \in D \subseteq T\}|$ and assigning a distinct number $n_D \in \{1, ..N\}$ to each D in the set $\{D : x \in D \subseteq T\}$, we let $k' = u'(x)$, $k = u(x)$, $A_{n_D} = u_R(D\setminus x)$ and $B_{n_D} = \left[\prod\limits_{y\in D\setminus x}\gamma_R(y)\prod\limits_{z\in T\setminus D}(1-\gamma_R(z))\right]$. Then, applying Lemma 2, the above holds if and only if $u'(x) \geq u(x)$. ■

Proof for Theorem 3

Proof. It is immediate to see this is true by the fact that $\rho_x(x, x) = \gamma'(x)$ and $\rho_x(x, \emptyset) = \gamma(x)$. ■

Proof for Proposition 2

Proof. In the following, for notational consistency, we write for $i = 1, 2$

$$u_R^i(x) = \begin{cases} u_i'(x) & \text{if } x \in R \\ u_i(x) & \text{otherwise} \end{cases}$$

The attention parameter part is proven in the main text. For utility, notice that we have for any $x \in S$

$$\frac{u_1(x)}{u_1(S)} = F_S(x, \emptyset) = \frac{u_2(x)}{u_2(S)} \text{ and } \frac{u_1'(x)}{u_1'(S)} = F_S(x, S) = \frac{u_2'(x)}{u_2'(S)}$$

Therefore, by the standard argument, we know that there exists $\alpha, \beta > 0$ such that $u_1 = \alpha u_2$ and $u_1' = \beta u_2'$. Finally, consider $F_{\{x,y\}}(x, x)$, we have

$$\frac{u_1'(x)}{u_1'(x) + u_1(x)} = \frac{\beta u_2'(x)}{\beta u_2'(x) + \alpha u_2(x)} = F_{\{x,y\}}(x, x) = \frac{u_2'(x)}{u_2'(x) + u_2(x)}$$

Therefore, it must be that $\alpha = \beta$. The proof is complete. ■

Proof for Proposition 3

Proof. Note that that we have, for $x \in R \subseteq S$

$$\rho_S^A(x,R) := \sum_{x\in A\subseteq S} \Big[\prod_{y\in A} \gamma_R(y) \prod_{z\in S\setminus A} (1-\gamma_R(z))\Big] \frac{u(x)}{u(x)+u_R(A\setminus x)} = \rho_S(x, R\setminus x)\frac{\gamma'(x)}{\gamma(x)}$$

And also, for $x \in R \subseteq S$)

$$\rho_S^U(x,R) := \sum_{x\in A\subseteq S} \Big[\gamma(x) \prod_{y\in A\setminus x} \gamma_R(y) \prod_{z\in S\setminus A} (1-\gamma_R(z))\Big] \frac{u_R(x)}{u_R(A)} = \rho_S(x,R)\frac{\gamma(x)}{\gamma'(x)}$$

Therefore, we have $\rho_S^U(x,R) \geq \rho_S^A(x,R) \iff$

$$\rho_S(x,R)\frac{\gamma(x)}{\gamma'(x)} \geq \rho_S(x,R\setminus x)\frac{\gamma'(x)}{\gamma(x)} \iff \frac{\rho_S(x,R)}{\rho_S(x,R\setminus x)} \geq (\frac{\gamma'(x)}{\gamma(x)})^2 = (\frac{\rho_x(x,x)}{\rho_x(x,\emptyset)})^2$$

The proof is complete. ■

Proof for Proposition 4

Proof. Let the revenue be denoted by Π. In the following, we denote $u(x)$ and $\gamma(y)$ by u_x and u_y for notational simplicity. Note that $\Pi = b^{SP}(x) = \rho(x,x) - \rho(x,y)$, so that we have

$$\Pi = (\alpha+(1-\alpha)\gamma_x)\Big[\gamma_y\frac{u_x(1+\beta)}{u_x(1+\beta)+u_y}+(1-\gamma_y)\Big]-\gamma_x\Big[(\alpha+(1-\alpha)\gamma_y)\frac{u_x}{u_x+(1+\beta)u_y}+(1-\alpha)(1-\gamma_y)\Big] \tag{3}$$

Also, note that $b(x) = b(y) = \rho_S(y,y) - \rho_S(y,x)$, so that we have also

$$\Pi = (\alpha+(1-\alpha)\gamma_y)\Big[\gamma_x\frac{u_y(1+\beta)}{u_y(1+\beta)+u_x}+(1-\gamma_x)\Big]-\gamma_y\Big[(\alpha+(1-\alpha)\gamma_x)\frac{u_y}{u_y+(1+\beta)u_x}+(1-\alpha)(1-\gamma_x)\Big] \tag{4}$$

Firstly, we assume $\alpha = 0$ and $\beta > 0$. By taking derivative for (3) with respect to γ_x, we have

$$\begin{aligned}\frac{\partial\Pi}{\partial\gamma_x} &= \Big[\gamma_y\frac{u_x(1+\beta)}{u_x(1+\beta)+u_y}+(1-\gamma_y)\Big] - \Big[\gamma_y\frac{u_x}{u_x+(1+\beta)u_y}+(1-\gamma_y)\Big]\\ &= \gamma_y\Big[\frac{1}{1+\frac{u_y}{(1+\beta)u_x}} - \frac{1}{1+\frac{u_y(1+\beta)}{u_x}}\Big] > 0\end{aligned}$$

which proves Proposition 4 a)(i). Also, by taking derivative for (4) with respect to u_x, we have

$$\begin{aligned}\frac{\partial\Pi}{\partial u_x} &= -\gamma_y\gamma_x u_y(1+\beta)\frac{1}{(u_y(1+\beta)+u_x)^2} + \gamma_y\gamma_x u_y(1+\beta)\frac{1}{(u_y+u_x(1+\beta))^2}\\ &= \frac{\gamma_y\gamma_x u_y(1+\beta)}{(u_y(1+\beta)+u_x)^2(u_y+u_x(1+\beta))^2}(u_x+u_y)(2+\beta)\beta(u_y-u_x)\end{aligned}$$

which equals 0 when $u_x = u_y$, and it is positive when $u_x < u_y$ and it is positive when $u_x > u_y$. It proves Proposition 4 a)(ii).

Secondly, we assume $\alpha > 0$ and $\beta = 0$. By taking derivative for (3) with respect to γ_x, we have

$$\begin{aligned}\frac{\partial \Pi}{\partial \gamma_x} &= (1-\alpha)\Big[\gamma_y \frac{u_x}{u_x+u_y} + (1-\gamma_y)\Big] - \Big[(\alpha+(1-\alpha)\gamma_y)\frac{u_x}{u_x+u_y} + (1-\alpha)(1-\gamma_y)\Big] \\ &= -\alpha\frac{u_x}{u_x+u_y} < 0\end{aligned}$$

which proves Proposition 4 b)(i). Also, by taking derivative for (4) with respect to u_x, we have

$$\frac{\partial \Pi}{\partial u_x} = -(\alpha+(1-\alpha)\gamma_y)\gamma_x\frac{u_y}{(u_y+u_x)^2} + \gamma_y(\alpha+(1-\alpha)\gamma_x)\frac{u_y}{(u_y+u_x)^2} = \alpha(\gamma_y-\gamma_x)\frac{u_y}{(u_y+u_x)^2}$$

which equals 0 when $\gamma_x = \gamma_y$ and is positive when $\gamma_x < \gamma_y$ and is negative when $\gamma_x > \gamma_y$. It proves Proposition 4 b)(ii). ■