Market Power with Limited Attentive Population

Yusufcan Demirkan*

August 1, 2022

Abstract

We study two market scenarios in which firms face a population of consumers with limited attention. A consumer considers each product with a certain probability, called the attention parameter, which determines the aggregate choice probabilities of the alternatives offered in the market. Firms know the distribution of this parameter in the population. We first consider the scenario of a single firm facing no competition, and then a market this firm becomes the default seller for all consumers but faces competition of another firm. We are interested in the consequences of limited attention of the consumers and initial market share on the market outcomes such as the market power of firms measured by their profits, the variety and quality of products offered, and the welfare of the consumers. We discuss ways to incorporate the pricing decision and different specifications of the attention parameter, the latter demonstrating that the market implications depend on the determinants of attention.

1 Introduction

This paper presents two related market scenarios in which consumers are cognitively limited and a firm has the monopoly power. First, we model the decision-making of a single firm, and later two firms one of which is the default provider for all consumers. While we keep the standard assumption that consumers have stable preferences, we assume that they have limited awareness of the choice environment, possibly with different levels of attention. Our goal is to analyze the market implications of these scenarios.

Consumers often choose sub-optimal alternatives with respect to their preferences, and usually the choices of even a single consumer are observed to be probabilistic. There is a well-established literature demonstrating that attention is a scarce resource, and people often consider a subset of available alternatives.[1] A firm that wants to sell a product needs to ensure that its product is considered, while the better products offered by other firms are not.

Our model consists of a measure one of limited attentive consumers. Consumers have identical preferences, but differ on the basis of their attention levels. In particular, we use the well-known model of Manzini and Mariotti [28] and its extension to a population of consumers by Dardanoni et al. [9]. Each consumer pays attention to a product with a certain probability $\theta \in [0,1]$, independent and identical across products. The distribution of the attention parameter in the

*Department of Economics, New York University. Email: yd1131@nyu.edu

[1]See Hauser and Wernfelt[19] and Shocker et al. [34] for some classic references.

population is represented by the cdf F. A consumer buys a product if she considers that product and does not consider any better products according to her preferences. When the consumer does not consider any alternatives, she goes with her default option. The interpretation of the default option will depend on the specific market setting we will consider. We discuss two market scenarios to see the effects of the initial market share and competition depending on the distribution of attention in the population: In the first one, there is a single firm choosing a menu to offer. The consumers in this scenario has the outside option of not buying anything. If this firm enters the market, then we assume that it gains the loyalty of all consumers. In the second scenario, we imagine that the firm from the first scenario enters the market and becomes the default brand for all consumers, meaning that a consumer not considering any new products purchases the default product offered by this firm. We introduce competition by assuming there is a competitor to the incumbent. Most importantly, in both settings there is a firm having monopoly power, and two models help us to see the impact of competition on the market outcomes.

Our results demonstrate that whether a firm prefers a more attentive population or not depends on the market structure. Furthermore, it provides another channel why a firm provides multiple products in equilibrium in addition to the classical approach of preference heterogeneity. A population is said to be more attentive compared to another if the distribution of attention in the former first-order stochastically dominates the distribution of the latter. When a firm does not face competition, it prefers a more attentive population as shown in Theorem 1. The intuition behind this result is related to the availability of the default option and its interpretation: A more attentive population considers each menu with a greater probability compared to a less attentive one, implying that the probability of considering nothing is less. This implies that if the consumers are more attentive, then the probability that their consideration set is empty is lower, and hence the probability of buying nothing. Furthermore, this holds for any finite level of stochastic dominance, and we provide an interpretation why this holds for the second-order stochastic dominance with an example. In the extreme case of a fully attentive population (meaning all consumers consider all products with probability 1), it offers the lowest quality product to the consumers as the only option. On the other hand, if the population is fully inattentive (meaning no consumer is considering a new product), it offers nothing and stays outside the market. In intermediate cases, the number and quality of products offered depend on the attention distribution, but we show that the number of products offered monotonically increase with respect to the attention of the population provided a certain condition is satisfied. In this case, the more attentive population is better off compared to a less attentive one, because the firm offers all products it offered to the latter and more.

If the firm faces competition as in the second scenario, the results change. First of all, we do not have a monotonic relation as in the case of a single firm: A firm does not necessarily prefer a more attentive population. The incumbent prefers to face a fully inattentive population in this setting, while its profits decrease as a result of introducing competition if the population is fully attentive. When the population is highly attentive or inattentive, the welfare of the consumers increase compared to the benchmark the incumbent faces no competition. We show that under a condition guaranteeing the distribution of the attention is not extreme, the unique pure strategy equilibrium of the game is the default provider only offering the default product, and the competitor tries to penetrate into the market by offering the full set of products. In this equilibrium, the incumbent achieves a higher market power and market share only by offering the default product, provided that the expected level of attention of the population is bounded above by a threshold (related to the cost and the number of all products in the market). Note that the default provider loses some of its market share (because its initial share is one), but manages to keep a higher share compared to its competitor by simply offering what it previously offers. This has empirical support in some industries: Hortacsu et al. [25] shows that the consumers whose default is the incumbent in the electricity market do not switch to the competitors even if they offer lower prices. In fact, the incumbent (which has a share of one) manages to keep a share of sixty percent after four years of entrance by several competitors.

We extend the result in pure strategies to any equilibrium of the game including mixed strategies. Put precisely, we show that under certain conditions, the competitor offers all products in equilibrium, although not necessarily at the same time. Moreover, the default firm offers no new products, i.e. it only offers the default product it produces, with strictly positive probability. To better grasp the intuition of the model, we simplify the attention distribution in the population by assuming consumers are either fully attentive or fully inattentive, in which case the unique equilibrium of the game is in mixed strategies. A fully attentive consumer chooses always the best option offered to her, while a fully inattentive consumer always goes with the default. While the former can be interpreted as a rational agent, the latter can be interpreted as an agent who is satisficed with her default and does not search the market any more. In equilibrium, the incumbent offers the best product and the default, while the competitor offers the best and worst products. Performing a comparative static becomes simple in this setting. Letting λ denote the share of fully attentive consumers in the population, an increase in λ decreases the equilibrium payoff of the default seller, while it increases the equilibrium payoff of the competitor.

The model we present has certain limitations. First of all, we abstract away from the pricing decision of firms, because our main goal is to investigate the impact of initial market share and the attention distribution of the population on the number and quality of the products offered to the consumers, market power and market share of competing firms. There are several industries in which prices do not play a role such as media industry and search engines (see the comment in Eliaz and Spiegler [15]), and there is evidence that in some industries prices are uniform (see DellaVigna and Gentzkow [10]). We discuss several ways to incorporate prices in both frameworks, but we focus on price competition in a companion paper.

The second limitation is that we assume the attention parameter to be independent of the products and the menu it is presented in. Reutskaya et al. [33] provides support for this: They show that there is no correlation between quality of a product and attention towards it. We also discuss ways of incorporating this in our framework. First, we show that in the monopolistic market, the result that the firm prefers a more attentive population holds when the attention to a product depends on its position with respect to its rank in preferences, under an independence condition. We discuss two ways to relax the independent attention parameter in the second setting: the preference-dependent attention and menu-dependent attention. These show that the equilibrium outcomes depend on how the attention parameter is specified, so one should be careful about the determinants of the attention. For instance, menu-dependent attention incorporates the effects of complexity on the attention parameter by assuming that the consumer does not pay attention to any products if it exceeds her attention limit, in which case she goes with her default option. When this is the case, the probability that the population chooses the default can increase as the variety in the market increases, which is not the case when the attention parameter is menu-independent. In this case, the incumbent offers more products compared to the entrant in equilibrium.

The organization of the paper is as follows: In Section 2, we discuss the related literature. Then, in Section 3, we explain our model. In Section 4, we discuss the case of a single firm. Then, we define a game in which the incumbent faces competition in Section 5. In both scenarios, we consider ways of incorporating the simplifications we make. All proofs not stated in the main body are in the Appendix.

2 Related Literature

This paper is an example of a growing body of research in the literature of bounded rationality in industrial organization. For a textbook treatment of the subject, one can refer to Spiegler [36]. There are several reviews of the subject such as Ellison [17], Heidhues and Koszegi [23], and Spiegler [37]. Eliaz and Spiegler [15],[16] discuss models in which the formation of the consideration set is deterministic. The latter paper analyzes a model in which the consumer has a deterministic consideration set and firms can use marketing techniques in order to affect the consideration set of the consumer, while the former one discusses the strategic use of products which are only used as attention grabbers. Firstly, this paper differs from these because the consideration is here probabilistic. Furthermore, we analyze multiproduct firms, which is also the case in Eliaz and Spiegler [15], but the products in their setting are pure attention grabbers. de Clippel et al. [12] looks at a population of consumers having a unit demand and a common default seller in separate markets, and each consumer allocates attention to different markets to investigate further opportunities. Matejka [30] applies rational inattention framework on the consumer side to analyze price rigidities in the presence of a monopoly. Matejka and McKay [31] extends this to a setting with multiple firms. These papers have the same conceptual background: The consumer is limited attentive, but the allocation of this attention is rational. On the other hand, the model used in this paper is silent on this issue, even though the consideration probabilities may result from such a rational allocation process. Armstrong and Vickers [4] analyze consumers who consider a subset of firms, and a consumer is captive for a firm if she considers only this firm. A consumer chooses the product of the firm offering the lowest price among the ones she considers. In this model, the attention is not brand-specific (but across products), but we discuss a closer model when the incumbent faces competition in which firms compete by choosing prices and the attention parameter is brand-specific. Furthermore, the product offered is homogeneous and firms compete in prices, while we look at the variety and quality of products offered in the market. The result that a single firm facing no competition prefers a more attentive population is also reached by Dahremoller and Fells [8]. Armstrong and Chen [3] discusses a variant of the model developed by Varian [38] in which some consumers are inattentive and only consider the price dimension of products, while not paying attention to the quality.

There are other papers which use probabilistic demand functions in an IO setting. Basov and Danilkina [5] look at the Bertrand competition with consumers choosing probabilistically according to Luce rule. Anderson and de Palma [13] uses one of the most widely used probabilistic demand procedures, multinomial logit, in an IO setting. Manzini and Mariotti [27] defines and studies attention games where the attention paid to a product depends on the salience profile of all products offered by firms. Using their choice procedure in Manzini and Mariotti [28], they investigate the correlation between salience and quality of the product, establishing that under some conditions competitive sources imply a positive correlation. Although we discuss several ways in which the attention parameter is determined by some other factor, we mainly focus on the attention parameter without giving it a structure. Furthermore, our analysis rests on outlining the equilibrium behavior of multiproduct firms w.r.t the number and quality of products, while their paper focuses on the relation between salience and quality in equilibrium.

Both the model and the results achieved in our model has empirical support from the literature. There is a growing literature identifying the preference and attention parameters in "first consider, then choose" models such as Dardanoni et al. [9], Abaluck and Adams [1], and Aguiar et al. [2]. Martinovici et al. [29] uses eye-movements as a predictor of the amount of attention paid by the consumer, and show that the attention trajectories are successful in predicting the final brand choice. Goeree [21] shows that differences in attention is another contributor of sales variation other than taste differences, the classical explanation in the literature. Furthermore, it shows that the mark-ups in the computer industry is dramatically different when one includes limited attention and differences in information in the analysis. Honka et al. [24] presents a recent overview of the econometric literature on search and marketing-based consideration set models. Gallego and Li [20] uses the model of Manzini and Mariotti and shows how to recover the preferences and attention probabilities from the data. Furthermore, they provide evidence for the overperformance of the random consideration model over the multinomial logit.

3 A Model

Let X be the set of all alternatives which is finite s.t. $|X| = n$ and $n \geq 2$. $\mathbb{X}$ denotes the set of all subsets of X, and a menu S is an element of $\mathbb{X}$. To each menu S, we associate the same default option $d \notin S$. The cost function is a mapping $c : \mathbb{X} \to [0,1]$. Firm incurs the fixed cost of the menu S it offers, $c(S)$, independent of the choice of the consumer. For simplicity, we denote $c(\{x\})$ as c_x. Each product has a strictly positive cost, but the default is costless. c is assumed to have an additive structure, that is, $c(S) = \sum_{x \in S} c_x$ for any $S \in \mathbb{X}$. Throughout the paper, we assume that producing a $\succsim$-better product is more costly. For simplicity, let c^* and c_* denote the cost of the products with the greatest and least cost, respectively.

Assume that there is a measure 1 of consumers. The choices of a single consumer is probabilistic. To model the probabilistic choice of the consumers, we use the well-known model of Manzini and Mariotti [28] and its extension to a population of consumers by Dardanoni et al. [9].[2] Each consumer has an attention parameter, denoted θ, which represents the probability that the consumer considers a particular product, so $\theta \in [0,1]$. All consumers are endowed with the same preference relation $\succsim$, a linear order on X. We will denote the $\succsim$-best product in a menu S as $b(S)$, and $\succsim$-worst product as $w(S)$. A single consumer with attention parameter θ chooses a product x offered in S with probability $p_\theta(x,S)$ s.t.:

$$p_\theta(x,S) = \theta[1-\theta]^{|B(x,S)|} \tag{1}$$

where $B(x,S)$ denotes the products that are strictly better than x in S w.r.t $\succsim$. The probability that a consumer of type θ chooses the default option is given by:

$$p_\theta(d,S) = [1-\theta]^{|S|} \tag{2}$$

which is the probability that the consumer considers no products offered in S. Assume that θ is distributed according to a cumulative distribution function F. So, $F(\theta)$ denotes the probability that a consumer in the population has an attention parameter less than or equal to θ. The

[2] In their model, each consumer is associated with a cognitive type $\theta \in \Theta \subset \mathscr{R}$, drawn independently across agents from a cdf F. We focus on *consideration capacity model*, an adaptation of Manzini and Mariotti [28].

aggregate choice probabilities in the population can be written as follows:

$$\begin{aligned} p_F(x,S) &= \int_0^1 p_\theta(x,S)dF \\ p_F(d,S) &= \int_0^1 p_\theta(d,S)dF \end{aligned} \tag{3}$$

Because we mostly focus on a single population, we drop the subscript F, and do not use it unless it leads to confusion.

4 A Monopolistic Market

Our first inquiry is to understand the behavior of a single firm who considers to enter a market in which there is no competitor. We interpret the firm as the creator of a new market (either by offering a new type of product or an already existing product to a region where it is not present), and therefore it has the monopoly power in the market once it offers a product. Another interpretation is to think of this firm as a natural monopoly considering to enter into a market. Firm chooses a menu $S \in \mathbb{X}$ to offer, which can be empty, in which case the market is not created. The default option of the consumers in single firm case is not to buy anything. We can write the monopolist's problem as follows:

$$\underset{S\in\mathbb{X}}{\text{argmax}}[\sum_{x\in S} p(x,S) - c(S)]$$

By definition, $\sum_{x\in S} p(x,S) = 1 - p(d,S)$. Furthermore, the monopolist will offer the cheapest $|S|$ products when it offers a menu of this size, because the cost is increasing as the quality of a product increases and the aggregate demand for a product is constant w.r.t the rank of a product.[3]Therefore, the monopolist chooses how many products to offer, starting from the worst-quality product. Let $X = \{x_1, \ldots, x_n\}$ and the quality of a product increases as the subscript increases. Abusing the notation, let us denote $p(d,k)$ as the default probability when the monopolist offers k products. The problem monopolist needs to solve becomes:

$$argmax_{k\in\{1,\ldots,n\}}[1 - p(d,k) - \sum_{i\leq k} c_{x_i}]$$

which can be equivalently written as:

$$argmin_{k\in\{1,\ldots,n\}}[p(d,k) + \sum_{i\leq k} c_{x_i}]$$

The monopolist decides to add k^{th} least costly product iff the gain in market share exceeds the cost of offering this product. For simplicity, we denote the cost of this product as c_k, while the optimal number of products under the distribution F is denoted as n^*_F. The next lemma provides a formula for the optimal number of products.

Lemma 1. $n^*_F = max\{k : \mathbb{E}(\theta[1-\theta]^{k-1}) > c_k\}$.

[3]We mean the following from the latter. For instance, consider the best product offered in two different menus S and S'. The aggregate probability of purchasing the best product from S and S' does not depend on the identity of the best product.

From this lemma, we can derive some simple conclusions regarding the number of products provided by the firm. For example, if the firm knows that any product offered will be purchased with a sufficiently high probability, then it will offer all products. A condition that is sufficient for this is $\mathbb{E}(\theta[1-\theta]^{n-1}) > c^*$. On the other hand, if the population is sufficiently inattentive, then the firm would not extend the product line. If for example $\mathbb{E}(\theta) < c_*$, then $n_F^* = 0$.

Single Attention Type

Let us denote a distribution F with a probability mass of 1 at a single point $\theta \in [0,1]$ as δ_θ. A population is said to be *fully attentive* if $F = \delta_1$, and it is *fully inattentive* if $F = \delta_0$. When the distribution of the attention in the population is as such, the aggregate choice probabilities becomes identical with the benchmark model of Manzini and Mariotti [28] (with identical attention parameters). Using Lemma 1, we can provide an expression for the optimal number of products. Indeed, we can express a more exact expression provided that the cost of all products are equal to c:

$$n_F^* = \frac{log(c/\theta)}{log(1-\theta)}$$

assuming the expression is an integer (otherwise, take the greatest integer below this number). As c increases, $log(c)$ increases because it is in $[0,1]$, which implies a decrease in the optimal number of products. Furthermore, n_F^* is monotonic w.r.t θ provided that c is sufficiently low.

The following proposition is about the relation between the optimal number of products and the variety offered conditional on the attention level of the population:

Proposition 4.1.

- *The monopolist offers the singleton menu $\{w(X)\}$ to a fully attentive population, while it offers no goods to a fully inattentive population.*
- *Assume that $\theta \in (0,1)$. If $|\hat{\theta} - \frac{1}{n^*_{\delta_\theta}}| \leq |\theta - \frac{1}{n^*_{\delta_\theta}}|$, then $n^*_{\delta_{\hat{\theta}}} \geq n^*_{\delta_\theta}$.*

The first result describes the behavior of the firm in two extreme cases. When all consumers in the population are fully attentive, they consider every product offered by the firm and select the $\succsim$-best product among them with probability one. Equipped with this knowledge, the firm would offer the cheapest product to maximize its profits and get a payoff of $1-c_*$. On the other hand, if the firm knows that no product it offers enter into the consideration set of the consumer, then the firm does not enter the market. For example, if consumers are extremely lazy so that they never look for any product they do not need (so what is offered by the monopolist is not a bottle of water or a medication), then the monopolist would prefer to offer no good at all. Observe that the monopolist achieves the maximum payoff possible in the fully attentive case, while it achieves the minimum in the fully attentive case.

The second result shows that the optimal number of products in two different populations can be compared provided a certain condition is satisfied. Assume that we have a population with an attention parameter of θ. The optimal number of products for this population is given by $n^*_{\delta_\theta}$. Note that when the monopolist adds a product, the rank of a product w.r.t $\succsim$ can change, but the choice probability for the same rank stays the same. Thus, the only difference in this setting comes from the choice probability of the worst product after the addition, because the number of products offered increases by one. This implies that the addition of a new product depends whether the benefit of adding it, equal to the probability of choosing the worst product after addition, outweighs the cost of this product. The proposition follows since the probability of choosing the worst product when there are $n^*_{\delta_\theta}$ products is maximized when the attention parameter is equal to $\frac{1}{n^*_{\delta_\theta}}$, and the condition in the proposition implies that the probability of choosing the worst product is higher when the attention parameter is $\hat{\theta}$. Note that if $\theta \times n^*_{\delta_\theta} < 1$ for all $\theta \in (0,1)$, then the monotonicity w.r.t attention holds for all $\theta \in (0,1)$. In fact, since the firm does not offer any products when $\theta = 0$, we can say this holds for all $\theta \in [0,1)$. Because the firm offers the least costly $n^*_{\delta_\theta}$ products, this means a population with a higher attention parameter is offered all goods offered to a population with smaller attention.

The next corollary reports the comparative statics we discussed above. In addition, we provide a result stating the monotonicity of profits w.r.t attention level, which we will show more generally in the next section.

Corollary. *The profit of the firm increases as θ increases. If furthermore $\theta \times n^*_{\delta_\theta} < 1$ for all θ, then the consumer welfare improves as the consumer becomes more attentive except at $\theta = 1$.*

General Attention Distribution

We will first define the notion of *stochastic dominance*. We use two results from the literature. Our first result follows from the equivalence between any finite degree of stochastic dominance and integral of a utility function whose derivative has to satisfy a sign requirement (corresponding to the level of dominance).[4] The second result relies on a theorem proved by Fishburn [18], which relates the degree of stochastic dominance to the moments of the corresponding random variable. We first define what m^{th}-order stochastic dominance means. Because we only focus on probability distributions, we provide the definition only for this case, adapted from Fishburn [18] [5]:

Definition 1. *Let F be a probability distribution from the set of continuous probability distributions $\mathbf{F}$. We define F^m for all $m \in \{1,2,\dots\}$ recursively. Let $F^1 = F$ and then define F^m for all $m \geq 2$ as follows:*

$$F^m(x) = \int_0^x F^{m-1}(\theta)dF(\theta)$$

The m^{th}-order stochastic dominance, a binary relation $\geq_m$ on $\mathbf{F}$, can be defined as:

$$F \geq_m G \quad \text{if and only if} \quad F^m(\theta) \leq G^m(\theta) \quad \forall \theta \in [0,1]$$

$$F >_m G \quad \text{if and only if} \quad F \geq_m G \quad \text{and} \quad F \neq G$$

for all $x \geq 0$ where $F, G \in \mathbf{F}$.

Interpreting stochastic dominance becomes harder as the degree of dominance increases. Therefore, we can limit ourselves to the first two degrees, even though we state the result more generally. Consider two distributions F and G. In order to simplify interpretation, imagine that we have the same monopolist entering into two different countries. If $F \geq_1 G$, then $\mathbb{E}_F(\theta) \geq \mathbb{E}_G(\theta)$. Because θ is the attention parameter, this means that the population with distribution F has a higher expected attention level. If $F \geq_2 G$, then either $\mathbb{E}_F(\theta) \geq \mathbb{E}_G(\theta)$ or $\mathbb{E}_F(\theta) = \mathbb{E}_G(\theta)$ and $\mathbb{E}_F(\theta^2) \leq \mathbb{E}_G(\theta^2)$. Assume that the latter case is true, so G is a mean-preserving spread of F. This means that although both populations have the same level of

[4] We use the version stated in Eechkoudt et al. [14]. This is a well-known result in the literature, see references in Eechkoudt et al. [14]

[5] Fishburn [18] provides a more general definition only using right-continuity.

expected attention, G has a higher variation in the distribution of attention. In the following result, we compare the maximum profit level and the number of products offered to achieve it when we can compare distributions according to any finite degree of stochastic dominance. We denote the maximum value of monopolist's profit under distribution F as Π_F^*.

Theorem 1. *Assume that $F, G \in \mathbf{F}$ and $F \geq_m G$ for some $m \in \{1, 2, \dots\}$. Then, the following are true:*

- $\Pi_F^* \geq \Pi_G^*$.
- *If $\mathbb{E}_F(\theta^s) = \mathbb{E}_G(\theta^s)$ for all $s \leq m-1$, then $n_F^* \geq min\{n_G^*, m\}$.*

The proof of this result is in the Appendix (like other proofs not stated in the main body). Observe that if $m \geq n$, then $min\{n_G^*, m\} = n_G^*$ since by definition $n_G^* \leq n$. Therefore, if we know that $F \geq_m G$ for such an m with first $m-1$ moments being equal, then we can simply state that $n_F^* \geq n_G^*$. On the other hand, if we know that F first-order stochastic dominates G, then the second part of the result we stated becomes simpler because the equality of moments condition disappears. We state these two observations as a corollary.

Corollary.

- *If $F \geq_1 G$, then $\Pi_F^* \geq \Pi_G^*$ and $n_F^* \geq 1$.*
- *If $F \geq_m G$ for some $m \geq n$ and $\mathbb{E}_F(\theta^s) = \mathbb{E}_G(\theta^s)$ for all $s \leq m-1$, then $n_F^* \geq n_G^*$.*

As we emphasized previously, interpreting stochastic domination gets harder as the degree of stochastic dominance increases. In the following examples, we interpret this result for the cases of first-order and second-order stochastic dominance. We compare two simple distributions, and try to provide an intuition for why the result we stated holds.

Example 1. Assume that the firm has the least amount of information regarding the attention distribution, i.e. F is a uniform distribution over $[0, 1]$. The probability of choosing a product in S is equal to $\frac{1}{|S|+1}$. This implies that the monopolist adds k^{th} product if and only if:

$$\frac{k}{k+1} - c[k] > \frac{k-1}{k} - c[k-1]$$

where $c[k-1]$ denotes the cost of $(k-1)$th least costly product, and similarly for k. This condition is equivalent to:

$$\frac{1}{k(k+1)} > c[k] - c[k-1]$$

Clearly, the expression on the left decreases as k increases. Assume for simplicity the additional cost for offering the next best product is constant and equal to c. There will be a cutoff product k^* s.t. adding k^*th product is profitable, but adding the better ones are not.

Now take another distribution G, which is uniform over $[\alpha, 1]$ s.t. $\alpha > 0$. G first-order stochastically dominates F. The probability of choosing the default in a menu S is equal to $\frac{(1-\alpha)^{|S|}}{|S|+1}$,and so the payoff of the firm for a menu of size k is equal to:

$$1 - \frac{(1-\alpha)^k}{k+1} - c(L[k])$$

From this expression, it is easy to see that the firm gets a higher maximum profit from distribution G. The firm adds the k^{th}-product (assuming again a constant additional cost for the next best product) iff:

$$\frac{(1-\alpha)^{k-1}(1+\alpha k)}{k(k+1)} > c$$

So, whether the firm offer more products to a population with distribution G is not clear. If $(1-\alpha)^{k-1}(1+\alpha k) \geq 1$, then we can say that the condition for the distribution G is satisfied whenever it is satisfied for F. This holds for all α when $k=1$, if $\alpha \leq \frac{1}{2}$ when $k=2$, and if α is less than or equal to approximately 0.23 for $k=3$.

Example 2. Let $F = \delta_\theta$ and G be the uniform distribution on $[0,1]$. Note that for any $\theta > \frac{1}{2}$, $F >_1 G$, which can be seen directly from the result we used in the first part of the Proof of Theorem 1 since $\mathbb{E}_F(\theta) > \mathbb{E}_G(\theta)$. If $\theta = \frac{1}{2}$, then F is equivalent to G w.r.t first-order dominance, but $F >_2 G$ because G is a mean-preserving spread of F. Theorem 1 shows that in both cases the optimal profit under F is at least as much as the optimal profit under G. This in turn depends on the aggregate default probability, because an increase in this means that on aggregate a lower measure of consumers purchase a product offered by the monopolist, which lowers the profit of the firm. For any menu $S \in \mathbb{X}$ and θ, the default probability under F is equal to $[1-\theta]^{|S|}$, while this is equal to $\mathbb{E}_G([1-\theta]^{|S|})$ under G. Because $[1-\theta]^{|S|}$ is convex, by Jensen's inequality we have:

$$\mathbb{E}_G([1-\theta]^{|S|}) \geq (1-\mathbb{E}_G(\theta))^{|S|}$$

When $F >_1 G$, we need to have $\mathbb{E}_F(\theta) > \mathbb{E}_G(\theta)$. For F, by definition we have $\mathbb{E}_F(\theta) = \theta$. Therefore, $\theta > \mathbb{E}_G(\theta)$, and hence $[1-\theta]^{|S|} < (1-\mathbb{E}_G(\theta))^{|S|}$. So, we conclude that the probability of buying nothing is greater in G compared to F. In general, because the measure of consumers with at least θ level of attention is always higher under F, we expect a higher measure of consumers purchasing a product under F.

Now consider $\theta = \frac{1}{2}$, in which case we have $F >_2 G$. This implies that the default probability is equal to $(\frac{1}{2})^{|S|}$ under F, while it is greater than or equal to $1-(\frac{1}{2})^{|S|}$ under G. Because $2(\frac{1}{2})^{|S|} \leq 1$ for any $|S| \geq 1$, the default probability is always greater under G unless $S = \emptyset$. Intuitively, this is for the reason that in G, the variation in θ creates a higher probability of facing a consumer with a smaller attention parameter and hence higher default probability.

What happens if the attention parameter of the individual is not fixed across products? Let $X = \{x_1, \ldots, x_n\}$ and $x_{i+1} \succ x_i$ for all $i \in \{1, \ldots, n-1\}$. We denote the product-specific attention parameter (which is independent from other factors such as the menu it is presented in) as θ_i for x_i. Note that in this setting, F is a joint distribution of random variables $\{\theta_i\}_{\{i\in\{1,\ldots,n\}\}}$ on $[0,1]^n$. We ask whether we can reach the conclusion we reached in Section 4: The monopolist prefers a more attentive population to the less attentive one. The answer to this question is positive provided the joint distribution is independent. Let F_i denote the marginal distribution of F for r.v. θ_i. Take another independent joint distribution G. We say $F \geq_1 G$ (first-order stochastic domination) G if and only if $F_i \geq_1 G_i$ for each component i.

Proposition 4.2. *Assume that F and G are two independent joint distributions s.t. $F \geq_1 G$. Then, $\Pi^*_F \geq \Pi^*_G$.*

Proof. The aggregate default probability in this setting is equal to $\mathbb{E}_F(\Pi_{\{i\in\{1,\ldots,n\}}[1-\theta_i])$. Because of the independence assumption, we can write this as $\Pi_{\{i\in\{1,\ldots,n\}}(1-\mathbb{E}_{F_i}(\theta_i))$. This directly implies the conclusion using the profit function of the monopolist and the fact that first-order stochastic dominance implies the dominance in expected value. $\square$

Educating the Consumer

What will happen if the monopolist decides to educate the consumers? To be able to answer this question, we first need to understand what educating a consumer means in this setting. A consumer with attention parameter θ considers each alternative with θ probability. If another consumer considers each alternative with at least θ, then we can say this consumer is more attentive. So, educating a consumer in our model implies that the consumer becomes more attentive. Because different consumers can have different parameters, we need to define the notion of being more attentive w.r.t the distribution of attention in the population.

Definition 2. *F is said to be more attentive compared to another distribution G if $F \geq_1 G$.*

Imagine we sample a subset of the population and decide to educate them. In this case, what we expect as a result of education is an increase in the attention parameter. This means that if a consumer with attention parameter θ is sampled and educated, we expect her to be a type with attention parameter θ' s.t. $\theta' \geq \theta$. So, if we educate a population distributed F and as a result reach another distribution F', then the probability that F' has a type higher than θ should be at least the probability that F has a type higher then θ. This is what the definition of first-order stochastic dominance gives to us. Furthermore, we know first-order stochastic dominance implies that the expected level of attention is higher in the dominating distribution, so the definition we provide implies the definition which can be provided using the expected levels of attention.

We showed that if $F \geq_1 G$, then $\Pi^*_F \geq \Pi^*_G$. Thus, we can say for sure that educating the population is beneficial for the monopolist, while the consequences are unclear for the consumers. To see, we can check the second part of the same result, Theorem 1. The corollary of this result shows that when $F \geq_1 G$, we know that $n^*_F \geq min\{n^*_G, 1\} = 1$ because unless $G = \delta_0$, the monopolist offers at least a single product. If we would have more information regarding this, then we could provide a comparison related to the consumer welfare. To see, assume that indeed $n^*_F \geq n^*_G$. We know that the monopolist offers the least costly products, so the best product offered for a population with F is weakly better than the best product offered to G. Also, we know that the probability of choosing the best product offered is highest among all alternatives in the market. Finally, we know that the probability that the best product is chosen is equal to the expected level of attention in the population. Because $F \geq_1 G$ implies that $\mathbb{E}_F(\theta) \geq \mathbb{E}_G(\theta)$, we can conclude that a weakly better product is chosen with the highest probability in F, which is higher than the probability the best product offered is selected under G.

5 Incumbent Facing Competition

Imagine that the firm in the previous section already created the market by offering at least one product. We assume that this firm becomes the default provider for all consumers in the population, but it faces a competitor. We denote these two firms as **i** and **c** respectively for the incumbent and the competitor challenging this firm. A *market* is defined as a vector of menus offered by each firm, so it is an element of $\mathbb{X} \times \mathbb{X}$, and we denote a generic element as $\mathbf{M} = (S_{\mathbf{d}}, S_{\mathbf{c}})$. A product x is offered in $\mathbf{M}$ if and only if $x \in \bigcup \mathbf{M} = S_{\mathbf{d}} \cup S_{\mathbf{c}}$.[6] As in the case with the single firm, there is a default option in each market $\mathbf{M}$.

[6]We usually abuse the notation and say $x \in \mathbf{M}$ in this case.

We are going to look at the pure strategy equilibrium in two cases: Both firms simultaneously choose a menu to offer, and the incumbent has a first-mover advantage. We can summarize the main incentives of firms as follows: For the incumbent, selling only the default option is highly attractive (especially if the population is not very attentive), since producing it is costless. On the other hand, the competitor would prefer increasing the variety in the market because there is no gain if consumers choose their default, provided that the cost of increasing the variety is not too high.

First, assume that both firms choose menus simultaneously. In Section 3, we defined the notions of fully attentive and inattentive populations. We learned from our analysis of the monopoly case the following about these benchmark cases: If the monopolist faces a fully attentive population, then it offers $\succsim$-worst and hence the cheapest product. On the other hand, if the monopolist faces a fully inattentive population, then it chooses to offer nothing. Our first goal in this section is to see the effects of introducing competition for such populations. We are going to show that competition improves consumer welfare in both cases. In fact, the results we state are more general compared to the extreme cases.

The following result shows that if the population is sufficiently inattentive, then **i** can deter the entrance of the competitor meaning that **c** offers no products in equilibrium. This proposition is almost immediate: For any M offered in any equilibrium, the total share firm i gets in a market (M,E) for some E offered by the competitor is bounded above by $\mathbb{E}(\theta)$, and hence c_*. Therefore, offering a product is not profitable for the firm, and can deviate by dropping it. Hence, the competitor offers no products, and the incumbent only provides the default product keeping the full market share by deterring entrance.

Proposition 5.1. *If $\mathbb{E}(\theta) < c_*$, then the unique equilibrium of the game is $(\{d\},\{\emptyset\})$.*

Thus, the default provider gets a share of one with zero costs, meaning that its market power is equal to its market share. This contrasts with the incumbent's preferences when it faces no competition: A more attentive population is preferred when the firm is alone in the market, while this proposition shows that the firm which has the monopoly power achieves the maximum market power when the population is sufficiently inattentive.

What happens if the population is sufficiently attentive? In this case, being default provider does not work for the incumbent, so competition forces both firms to offer the best product.

Proposition 5.2. *If $\mathbb{E}(\theta) > 1 - c_*$ and $c^* - c_* < \frac{1}{2}$, then the unique equilibrium of the game is $(\{b(X)\},\{b(X)\})$.*

Next, we show that a completely different result can be achieved when the attention distribution changes. Before doing that, we are going to show two lemmas, which are themselves informative about the equilibrium play. Let $(S_{\mathbf{d}}, S_{\mathbf{c}})$ denote a generic pure strategy equilibrium profile.

Lemma 2. *If $c^* < \mathbb{E}(\theta[1-\theta]^{n-1})$ and $\mathbb{E}(\theta) < c^* + c_*$, then $S_{\mathbf{d}} \subset S_{\mathbf{c}}$ and $S_{\mathbf{d}} \cup S_{\mathbf{c}} = X$.*

This lemma shows that if there is a pure strategy equilibrium, then this cannot be symmetric: Entrant offers all products offered by the monopolist, and even more. If we assume that there is a tighter upper bound on the expected attention level of the population, then firms offer no products in common.

Lemma 3. *If $\mathbb{E}(\theta) < 2c_*$, then $S_{\mathbf{c}} \cap S_{\mathbf{d}} = \emptyset$.*

The condition we use for the next result has two parts: The first one says that the expected level of attention in the population is sufficiently low ($\mathbb{E}(\theta) < 2c_*$) so that there is enough consumers in the population who do not choose the $\succsim$-best product offered in the market. On the other hand, the other part ($c^* < \mathbb{E}(\theta[1-\theta]^{n-1})$) says that the attention distribution F is such that the measure of the consumers with positive level of attention, $\theta \in (0,1]$, is sufficiently high so that the share who purchase a product and not the default offered is above a certain threshold.

Theorem 2. *If $c^* < \mathbb{E}(\theta[1-\theta]^{n-1})$ and $\mathbb{E}(\theta) < 2c_*$, then $(\{d\}, X)$ is the unique pure strategy equilibrium.*

This result shows that if the distribution of attention satisfies some conditions guaranteeing that the the population is not on the extremes, then in the unique pure strategy equilibrium the incumbent only offers the default product.[7] The competitor, knowing that all offered products are sold with a sufficiently high strictly positive probability, offers the full set of products X.

Next result shows that under a further weak condition, the incumbent keeps a higher market power and market share.

Corollary. *If $\mathbb{E}(\theta) < 1 - \sqrt[n]{\frac{1-c(X)}{2}}$, then $\mathbf{i}$ has a higher market power. Furthermore, if $\mathbb{E}(\theta) < 1 - \sqrt[n]{\frac{1}{2}}$, then the final market share of $\mathbf{i}$ is higher.*

What happens if some consumers decide to try the new firm in the market? Specifically, imagine that a certain share of consumers who does not consider any alternatives in the market does not go to the incumbent for sure, but chooses a product offered by the entrant with $\gamma \leq \frac{1}{2}$ probability. In other words, we have a duopoly of firms s.t. the incumbent has a higher market share (equal to $1-\gamma$). It is not hard to show that if γ is sufficiently low, then the incumbent does not extend the product line also in this case, and the competitor offers full line of products. Thus, an incumbent believing that it will keep a sufficiently high share of loyal consumers will behave similar.

Uniqueness

An important question is whether there is any mixed strategy equilibrium of this game. Later, we show a case in which there is no pure strategy equilibrium, and provide the unique equilibrium if the game. However, we might have the converse: The unique pure strategy equilibrium above is the unique equilibrium of the game including mixed strategies. A mixed strategy σ is a probability distribution over the set of all menus, so $\sigma \in \triangle(\mathbb{X})$. We denote the support of σ as $S(\sigma)$. To differentiate between the mixed strategies used by different firms, we use σ_i and $S(\sigma_i)$ for firm $i \in \{\mathbf{d}, \mathbf{c}\}$. The following object defines the aggregate probability that consumers go to their default provided the competitor plays a mixed strategy σ and the incumbent does not extend the product line:

$$\alpha_d(\sigma) = \sum_{M \in S(\sigma)} \sigma(M)\mathbb{E}([1-\theta]^{|M|}) \tag{4}$$

Similarly, we can define the aggregate probability of default when the incumbent extends the product line by offering M as $\alpha_d(M,\sigma)$. When it is clear from the context, we will drop the σ

[7] If we know furthermore that under distribution F we do not face any consumer with an attention level below a certain threshold, then we could state the theorem by only imposing a condition on the expected level of attention. In particular, the following conditions are sufficient to state this result: $F(2/n) = 0$ and $\mathbb{E}(\theta) \in (\frac{c^*}{[1-2c_*]^{n-1}}, 2c_*)$.

from the notation and just use $\alpha_d, \alpha_d(M)$. By Jensen's inequality, we have:

$$\alpha_d(\sigma) \geq 1 - \sum_M \sigma(M)[1 - \mathbb{E}(\theta)]^{|M|} \tag{5}$$

Let us define the following object for simplifying the notation:

$$\mathscr{B}(m) = \sum_{i=0}^{m} (-1)^i \binom{m}{i} \mathbb{E}(\theta^i) \tag{6}$$

for $m \in \{0, \ldots, n\}$. Note that $[1 - \mathbb{E}(\theta)]^{|M|} = \mathscr{B}(|M|)$ for any M. Finally, let us define:

$$\mathscr{B}(\sigma) = \sum_{i=0}^{m} \sigma(M) \mathscr{B}(m) \tag{7}$$

Therefore, we can write simply $\alpha_d(\sigma) \geq 1 - \mathscr{B}(\sigma)$. For the next proposition, assume that the conditions in Theorem 2 holds.

Proposition 5.3. *If $\mathscr{B}(\sigma) < \frac{\mathscr{B}(|M|) - \mathscr{B}(|M \cup E|) + c(M)}{1 - \mathscr{B}(|M \cup E|)}$ for all M and σ, then $(\{d\}, X)$ is the unique equilibrium of the game.*

Comparative Statics

Imagine that the consumers are educated. According to our definition, this implies that the attention distribution in the population moves to a distribution which first-order stochastic dominates the previous one, say F^e. Next result shows that under a further condition on the previous distribution,

Corollary. *If $F^e \geq_1 F$, F satisfies the conditions in Theorem 2, and also $F(1/n) = 0$, then $(\{d\}, X)$ is the unique pure strategy equilibrium under distribution F^e.*

Proof. Because $F^e \geq_1 F$, $\mathbb{E}_{F^e}(\theta) \geq \mathbb{E}_F(\theta)$ and therefore $\mathbb{E}_F(\theta) < 2c_*$. Let $h(\theta) = -\theta[1 - \theta]^{n-1}$. To conclude the proof, we are going to show that $h(\theta)$ satisfies the sign requirement. Note that $\frac{\partial h}{\partial \theta} = [1 - \theta]^{n-2}(n\theta - 1)$. This satisfies the sign requirement because $F^e \geq_1 F$ and $F(1/n) = 0$ implies by the definition of first-order dominance that $F^e(1/n) = 0$. Therefore, $\mathbb{E}_{F^e}(h(\theta)) \geq \mathbb{E}_F(h(\theta))$, implying that $c^* < \mathbb{E}_{F^e}(\theta[1 - \theta]^{n-1}) \leq \mathbb{E}_F(\theta[1 - \theta]^{n-1})$. □

We can show a similar result for the general case of any finite level of stochastic dominance, provided that we know $\mathbb{E}_G(\theta) < 2c_*$. If $F \geq_m G$ and $G(m/n) = 0$, then $(\{d\}, X)$ is the unique pure strategy equilibrium under distribution G. The proof of this follows simply applying the derivative repeatedly and verifying the sign requirement is satisfied for the corresponding level. These results can be interpreted by using the definition of more attentive we provided in the previous section. Recall that to establish Theorem 2, we used the condition:

$$c^* < \mathbb{E}_F(\theta[1 - \theta]^{n-1}) < \mathbb{E}_F(\theta) < 2c_*$$

So, the population is sufficiently inattentive not to purchase only the best product offered in the market, while the distribution is s.t. even the worst product is sold with a sufficiently high probability. Note that the latter implies that instead of going with the default, people choose the worst product offered in the market. This means that the lowest types in the population with an attention parameter sufficiently close to 0 are very small in the population. The first corollary says that the population with F is more attentive than G, and furthermore $G(1/n) = 0$. So, we know that the population with G would buy the best product with a smaller probability, while they also do not have many people choosing the default. As the level of stochastic dominance increases (meaning it gets weaker), the requirement on G gets stronger.

Mixed Strategies

A natural question to ask is whether the results hold more generally or not, since a pure strategy equilibrium does not exist necessarily.

Theorem 3.

- *If $c^* < \mathbb{E}(\theta[1-\theta]^{n-1})$ and $\mathbb{E}(\theta) < 2c_* + c^*$, then we have:*

$$\bigcup S(\sigma_{\mathbf{d}}) \subseteq \bigcup S(\sigma_{\mathbf{c}}) = X$$

- *If $\mathbb{E}(\theta) \leq 1 - \sqrt[n]{1-c(X)}$, then $\{d\} \in S(\sigma_{\mathbf{d}})$ and $\mathbf{m}$ receives a payoff of $\alpha_d(\sigma_{\mathbf{c}})$ in equilibrium.*

The first part provides a condition in which there is full variety in equilibrium, and all products offered by the monopolist are also offered by the entrant. Then, we show that if the expected attention level is below a certain threshold depending on the cardinality and the total cost of X, the monopolist offers $\{d\}$, i.e. it does not offer any new products with strictly positive probability in equilibrium. An increase in the cardinality of X (holding the total cost fixed) implies a decrease in the upper threshold of the expected level of attention because $1-c(X) < 1$, while an increase in the total cost of X (holding n fixed) implies an increase in the threshold.

The relevance of the previous result depends on the non-existence of pure strategy equilibrium. We will demonstrate that this is possible using a simple distribution. In particular, we consider a discrete distribution with two types consisting of the fully attentive and fully inattentive. We assume that the former type has a measure λ, while low type has $1-\lambda$. We denote such a discrete distribution as F_λ. This corresponds to the case of all-or-one consideration pattern as mentioned in Armstrong and Vickers [4].[8] We first provide the non-existence of a pure strategy equilibrium provided the population distribution satisfies certain conditions.

Proposition 5.4. *If the population distribution is F_λ and $\lambda \in (\max\{2(c^*-c_*), 2c_*\}, 2c^*)$, then there is no pure strategy equilibrium.*

For a population with a distribution F_λ, the second part of Theorem 3 implies the following corollary immediately:

Corollary. *If $\lambda \leq 1 - \sqrt[n]{1-c(X)}$, then $\mathbf{m}$ gets $1-\lambda$ in equilibrium.*

Next, we show the equilibrium play of the game when the population is distributed F_λ and the assumption of Proposition 5.4 is satisfied.

Proposition 5.5. *Assume the distribution of attention in the population is equal to F_λ and $\lambda \in (\max\{2(c^*-c_*), 2c_*\}, 2c^*)$. The game has a unique equilibrium in which firms play a mixed strategy s.t.:*

[8] Varian is such a model in which firms compete in prices. Baye et al.[**?**] solves for the equilibrium in the general case of any finite number of firms.

- $\sigma_{\mathbf{d}}(\{b(X)\}) = \frac{2(c^*-c_*)}{\lambda}$ & $\sigma_{\mathbf{d}}(\{d\}) = 1 - \frac{2(c^*-c_*)}{\lambda}$.
- $\sigma_{\mathbf{c}}(\{b(X)\}) = 2(1-\frac{c^*}{\lambda})$ & $\sigma_{\mathbf{c}}(\{w(X)\}) = \frac{2c^*}{\lambda} - 1$

with equilibrium payoffs $1-\lambda$ *and* $\lambda-(2c^*-c_*)$ *respectively for* **m** *and* **e**.

In equilibrium, three products are offered: $b(X)$, $w(X)$ and d. The unique common product offered by both firms is $b(X)$, and the competitor best responds to the incumbent offering the default by offering $w(X)$. As the share of high type increases, the probability that the entrant offers $b(X)$ increases, meaning that the entrant behaves more aggressively to capture the market knowing that the population she faces is sufficiently attentive. On the other hand, an increase in the cost of $b(X)$, c^*, decreases the probability that it offers $b(X)$. The situation is the converse for the incumbent: The probability that the incumbent offers $b(X)$ is positively related to the maximum cost difference between products, while it is inversely related to the share of the high type who considers every alternative with probability one. Because $\lambda \leq 1$ and $c^* \geq c_*$ by definition, the next corollary is immediate.

Corollary. *The market power of the default provider is always higher.*

If we look at the case with two types, the comparative statics get much easier. In this context, educating the consumer is equivalent to an increase in λ. The industry profit (total profits of two firms) is equal to $1-(2c^*-c_*)$, so it is independent from the share of fully attentive consumers in the population. The latter has an impact on the distribution of the industry profits to the firms: An increase in education, which corresponds to an increase in λ, decreases the share the default seller gets from the industry profits, while it has the opposite effect for the competitor.

We can also compare the equilibrium profits of these two firms with the benchmark cases of a sufficiently attentive and inattentive populations. In the former, the unique equilibrium (provided $c^*-c_* < \frac{1}{2}$) is both firms offering $b(X)$. Thus, each firm gets a payoff of $\frac{1}{2}-c^*$, implying that the total profit is equal to $1-2c^*$. When the population is fully attentive, the incumbent offers the default product and the competitor does not enter the market, implying that the profit of the incumbent is equal to 1 and the competitor is 0. This makes the total profit in the fully inattentive benchmark equal to 1. If $\lambda > \frac{1}{2}+(c^*-c_*)$, meaning the share of fully attentive consumers is not sufficiently high, then the competitor receives a higher equilibrium payoff in the case of two-types compared to fully attentive population. On the other hand, if $c^* > \lambda - \frac{1}{2}$, i.e. the share of fully attentive is sufficiently low, then the equilibrium profit the default provider gets higher under two-types.

First-Mover Advantage

Assume now that the incumbent has the first-mover advantage, and selects $S_{\mathbf{i}}$. Given this choice, the competitor chooses $S_{\mathbf{c}}$ s.t.:

$$S_{\mathbf{c}} \in \underset{S}{\operatorname{argmax}}\, p(S \setminus S_{\mathbf{i}}, S_{\mathbf{i}} \cup S) + \frac{1}{2} p(S_{\mathbf{i}} \cap S, S_{\mathbf{i}} \cup S) - c(S)$$

Note that the lemmas showed in the previous case also hold here, but now the incumbent has the first-mover advantage. In particular, if the incumbent offered x and $\mathbb{E}(\theta) < 2c_*$, then the competitor never offers a product offered by the incumbent. Also, $\mathbb{E}(\theta[1-\theta]^n) > c^*$ implies that the competitor always adds a product not offered by the incumbent. So, if these twp assumptions hold, then any pure strategy equilibrium $(S_{\mathbf{i}}, S_{\mathbf{c}})$ should be a partition of X. This implies that the problem of the incumbent is to choose $S_{\mathbf{i}}$ s.t.:

$$S_{\mathbf{i}} \in \operatorname*{argmax}_{S} p(S_{\mathbf{i}}, X) + p(d, X) - c(S_{\mathbf{i}})$$

From this, it easily follows that the incumbent extends the product line (product proliferation strategy) if and only if $p(S_{\mathbf{i}}, X) > c(S_{\mathbf{i}})$, and otherwise only offers the default product. Because $\mathbb{E}(\theta[1-\theta]^n) > c^*$, it is easy to see that the incumbent extends the product line. Hence, under the same assumptions as in Theorem 5, the resulting equilibrium is dependent on whether the incumbent has first-mover advantage or not. Let us denote a pure strategy equilibrium (assuming it exists) as $(S^*_{\mathbf{i}}, S^*_{\mathbf{c}})$. It is straightforward to show that both firms have no profitable deviation to a mixed strategy. We reach the following conclusion provided that the conditions of Theorem 5 are satisfied:

Proposition 5.6. *The incumbent follows a product proliferation strategy if it has the first-mover advantage, and otherwise keeps its market share by not extending the product-line.*

6 Further Discussion

6.1 Preference-Dependent Attention

In the original Manzini and Mariotti model, there is an attention parameter θ_x for each alternative x. So, a cognitive type θ needs to be defined as a vector of attention parameters over the set of products X. Below, we will present a simplified version of this which relates the attention parameter to the satisficing consumer defined in Simon [35].

Imagine that each consumer has a certain satisficing level, represented by an alternative in X. If the consumer's type is θ, then we can denote this alternative with x_θ. We assume that the consumer does not consider any product that does not satisfice her, i.e.:

$$\theta_x = \begin{cases} 1 & x \succsim x_\theta \\ 0 & o.w. \end{cases}$$

Assume that there are finitely many types. In particular, for each product x_i, there is a unique corresponding type θ_i with λ_i probability.[9] Let us first discuss the monopoly case. The problem of the monopolist becomes:

$$\max_{M \in \mathbb{X}} \sum_{b(M) \succsim x_i} \lambda_i - c(M)$$

It is easy to see that offering more than 1 product only increases costs, so the monopolist will offer a single product. If there is at least one product x_i s.t. $\sum_{x_i \succsim x_j} \lambda_j > c_{x_i}$, then the monopolist would offer a product. Assume that the measure of consumers that have the satisficing level of the worst product, λ_n, is sufficiently high so that $\lambda_n > c_n = c_{w(X)}$. So, offering $\{w(X)\}$ is profitable for the firm. Consider the next-best product, x_{n-1}. If $\lambda_{n-1} > c_{n-1} - c_n$, then the monopolist would offer $\{x_{n-1}\}$ instead of $\{x_n\}$. Continuing in this fashion, it is straightforward to show the following:

Proposition 6.1. *The monopolist offers* $\{x_k\}$ *s.t.:*

[9] One can think of this as a partitioning of the set of all types, which is assumed to be $[0,1]$. For example, the interval $[a,b] \subset [0,1]$ of types can correspond to a type θ_i with satisficing level x_i.

- $\sum_{i=k+1}^{l} \lambda_i \leq c_l - c_k$ *for all* $l \geq k+1$.
- $\sum_{i=j}^{k} \lambda_i > c_k - c_j$ *for all* $j \leq k$.

This shows that under suitable conditions, the monopolist will exclude high-types above a certain threshold from its supply, meaning that it offers a product which is not satisficing from their point of view. Next, we provide two simple examples directly following from the previous proposition demonstrating two different cases:

Example 3. Assume that $c_i = c$ and $\lambda_i > 0$ for all i. Then, the monopolist would offer $\{b(X)\}$. In this case, no type is excluded from the monopolist's supply.

Example 4. Assume that λ_i is a strictly increasing function of i s.t. $\sum_{i \in \{1,\ldots,n\}} \lambda_i = 1$. Next, we define the cost of product x_i, c_i. Let $\varepsilon_i > 0$ for $i \geq k$.

$$c_i = \begin{cases} \sum_{j \geq i}(\lambda_j - \varepsilon_j) & i \geq k \\ \sum_{j \geq i} \lambda_j - \sum_{j \geq k} \varepsilon_j & o.w. \end{cases}$$

With this arrangement, it is easy to show that the monopolist offers $\{x_k\}$, excluding any type θ_i s.t. $i \leq k-1$. These are the types with a strictly higher satisficing level compared to x_k.

We defined and discussed a way to model preference-dependent attention parameter in the case of a monopoly. Now, we will discuss how things change when we introduce competition. When competition is introduced, there is unique pure strategy equilibrium in which both firms offer $b(X)$.

Proposition 6.2. *Assume that $\lambda_i > c_i - c_{i+1}$ for all $i \in \{1,\ldots,n\}$ s.t. $c_{n+1} = 0$. There is a unique pure strategy equilibrium $(\{b(X)\},\{b(X)\})$.*

Proof. It is straightforward to see that both firms offer at most a single product (with monopolist being able to offer only the default), because provided that a consumer is attentive, she chooses the best product offered in the menu. Let $x_{\mathbf{m}}$ denote what is offered by the monopolist in equilibrium, and similarly $x_{\mathbf{e}}$. In equilibrium, $x_{\mathbf{e}} \succsim x_{\mathbf{m}}$, since otherwise the entrant receives a negative payoff and deviating to any $\{y\}$ s.t. $y \succsim x_{\mathbf{m}}$ is profitable. Furthermore, because $\lambda_i > c_i - c_{i+1}$ for all $i \in \{1,\ldots,n\}$, deviation to the next best product is always profitable, implying that $x_{\mathbf{e}} = b(X)$. Note that in this case the monopolist receives a negative payoff if $x_{\mathbf{m}} \neq b(X)$, and deviating to $\{b(X)\}$ is profitable because $\frac{1}{2} > c_{b(X)}$. Finally, we need to show that $(\{b(X)\},\{b(X)\})$ is an equilibrium, which follows easily because any deviation to a menu including $b(X)$ only increases the costs, while if $b(X)$ is not offered the deviator gets a negative payoff. □

6.2 Menu-Dependent Attention

The second simplification is that the attention parameter is menu-independent. Manzini and Mariotti [28] provide a result saying that we can rationalize any probabilistic choice data if the attention parameter is menu-dependent. Demirkan and Kimya [11] analyzes the case when the attention parameter is menu-dependent using the hazard rate function. Brady and Rehbeck [6] provides another model in which they attach a weight to each menu and define logit attention, making the choice probabilities menu-dependent. Both model generalizes the initial model presented by Manzini and Mariotti. Another way to model this is defining the attention parameter as follows:

$$\theta_m(M) = \begin{cases} 1 & |M| \leq m \\ 0 & o.w. \end{cases}$$

Note that the attention parameter is still identical for different alternatives, but it only differs according to the cardinality of the menu. Indeed, consumers are overwhelmed by high number of available alternatives in the market, as reported in Iyengar and Lepper [26].

Let us assume that each consumer has a type θ_m corresponding to the attention limit m. The number of types are finite, so there is a discrete distribution where each type m has probability λ_m. We assume that $\lambda_i > 0$ for all $i \in \{1, \ldots, m\}$ s.t. $\sum_{i \in \{1,\ldots,m\}} \lambda_i = 1$. Consider the case of the monopolist: If the monopolist offers a menu M, then types θ_m with $m < |M|$ cannot consider any alternative and therefore do not buy any alternative.[10] If $m \geq |M|$, then these types consider every alternative and choose $b(M)$. Thus, the profit of a monopolist from offering M is equal to:

$$\sum_{m \geq |M|} \lambda_m - c(M)$$

Note that by decreasing the cardinality of the menu, the monopolist can capture more consumers. This implies that the monopolist would offer a single product, and to maximize its profit, it would offer the cheapest one. Therefore, the monopolist will offer $\{w(X)\}$ no matter what the distribution of types is.

Now we discuss the menu-dependent attention parameter when there is competition. We are going to show that under a plausible assumption on the distribution of types, there is a unique pure strategy equilibrium of the game that is in stark contrast with the result in the monopolistic setting. Let us define the cumulative mass function Λ as follows:

$$\Lambda(m) = \sum_{j \leq m} \lambda_j$$

and denote the least costly m products in X as L_m, which is equal to the m $\succsim$-worst products. Let $X = \{x_1, \ldots, x_n\}$ with $x_{i+1} \succsim x_i$ for $i \in \{1, \ldots, n-1\}$, and $c_{x_i} = c_i$.

Proposition 6.3. *Assume that $\lambda_1 > c^* - c_*$, $\lambda_n > 2(c^* - c_*)$ and there is m^* s.t. $\lambda_k \leq 2c_{k+1}$ for all $n-1 \geq k \geq m^*$. Then, $(L_{m^*-1} \cup \{b(X)\}, \{b(X)\})$ is the unique equilibrium of the game.*

This shows that the behavior of the firms depend on how the attention parameter is specified. When attention parameter is identical across products, the default probability decreases as the variety increases in the market. On the other hand, an attention parameter that includes the complexity effect by inversely depending on the menu size changes this by affecting the default probability positively as variety increases. This is clearly seen from the result: The incumbent keeping its market share and gaining a higher market power only by offering the default product previously now offers m^* products in equilibrium.

Appendix A Proofs

Proof of Theorem 1

[10] Of course, this is a very simplified model, what we would expect normally is to consider a subset of alternatives with a cardinality less than m.

- We use an equivalence result from Eeckhoudt et al [14]. It implies in our context that:

$$F \geq_m G \iff \mathbb{E}_F(h(\theta)) \geq \mathbb{E}_G(h(\theta))$$

for all functions h s.t. the sign of $h^{(s)}$, where $h^{(s)}(\theta) = \frac{\partial^s u(\theta)}{\partial \theta^s}$, is equal to $(-1)^{s+1}$ for all $s \in \{1,...,m\}$. We will first show that this requirement can be weakened to the following:

Lemma 4. *If $F \geq_m G$, then $\mathbb{E}_F(h(\theta)) \geq \mathbb{E}_G(h(\theta))$ for all h s.t.*

$$sign(h^{(s)}(\theta)) = \begin{cases} (-1)^{s+1} & s \leq s^* \\ 0 & s > s^* \end{cases} \tag{8}$$

Proof. Assume $F \geq_m G$. We can write $\mathbb{E}_F(h(\theta)) - \mathbb{E}_G(h(\theta))$ in the following way:

$$\sum_{k=1}^{m-1} (-1)^k h^{(k)}(1)[F^{(k)}(1) - G^{(k)}(1)] + (-1)^m \int_0^1 h^{(m)}(x)[F^{(m-1)}(x) - G^{(m-1)}(x)]dx \tag{9}$$

following the steps in Eeckhoudt et al [14]. By definition of stochastic dominance, $F \geq_m G$ if and only if $F^{(m-1)}(x) \leq G^{(m-1)}(x)$ for all $x \in [0,1]$ and $F^{(i)}(1) \leq G^{(i)}(1)$ for all $i \in \{1,\ldots,m-2\}$. Assuming that h satisfies the requirement as in Equation 8, Equation 9 reduces to:

$$\sum_{k=1}^{s^*} (-1)^k h^{(k)}(1)[F^{(k)}(1) - G^{(k)}(1)]$$

and this is still nonnegative because the sign of h^k is $(-1)^{k+1}$ and $F^{(k)}(1) \leq G^{(k)}(1)$. This concludes the proof.

□

Let $h_k(\theta) = -[1-\theta]^k$. This function satisfies the sign requirement 8. Therefore, $F \geq_m G$ implies that $\mathbb{E}_F(h_k(\theta)) \geq \mathbb{E}_G(h_k(\theta))$ for any k. The monopolist maximizes the objective function $1 - p(d,S) - c(L[|S|])$ by choosing $|S|$. Note that $p(d,S) = \mathbb{E}_F([1-\theta]^{|S|}) = \mathbb{E}_F(-h_{|S|}(\theta))$. So, the problem of the monopolist is equivalent to:

$$1 + \mathbb{E}_F(h_{|S|}(\theta)) - c(L[|S|])$$

which implies the optimal profit for any size $|S|$ is higher under F. Hence, $\Pi_F^* \geq \Pi_G^*$.

- Theorem 2 in Fishburn [18] shows that if $F \geq_m G$ for some m and $\mathbb{E}_F(\theta^s) = \mathbb{E}_G(\theta^s)$ for all $s \leq m-1$, then $(-1)^m \mathbb{E}_F(\theta^{m+1}) > (-1)^m \mathbb{E}_G(\theta^{m+1})$. To complete the proof, we are going to derive an equation according to which the monopolist decides to add an additional product. This decision will rely on a comparison between the marginal benefit of adding the product versus marginal cost of it. Because $p(d,S) = \mathbb{E}_F([1-\theta]^{|S|})$, we can use the binomial expansion to get the following expression:

$$p(d,S) = \sum_{i=0}^{|S|} (-1)^i \binom{|S|}{i} \mathbb{E}_F(\theta^i)$$

Observe that which products are offered by the monopolist is immaterial, so the monopolist will choose the least costly $|S|$ products to offer for any menu S. This implies that the monopolist's problem is equivalent to the following:

$$\underset{S\in\mathbb{X}}{\operatorname{argmax}} \sum_{i=1}^{|S|}(-1)^{i+1}\binom{|S|}{i}\mathbb{E}_F(\theta^i) - c(L[|S|])$$

Thus, the monopolist's problem is to choose how many products to offer given the population is distributed according to F. In general, we can derive the following condition, according to which the monopolist decides to add $(k+1)^{th}$ product. We use Pascal's Theorem[11] to get this expression:

$$\sum_{i=1}^{k+1}(-1)^{i+1}\binom{k}{i-1}\mathbb{E}_F(\theta^i) \geq c[k+1] \tag{10}$$

Because $\mathbb{E}_F(\theta^s) = \mathbb{E}_G(\theta^s)$ for all $s \leq m-1$, the left hand-side of equation 10 is equal under F and G when $k \leq m-2$. So, provided that $n_G^* \leq m-1$, $n_F^* \geq n_G^*$. By what is shown in Fishburn [18], $(-1)^{m-1}\mathbb{E}_F(\theta^m) > (-1)^{m-1}\mathbb{E}_G(\theta^m)$, so $(-1)^{m+1}\mathbb{E}_F(\theta^m) > (-1)^{m+1}\mathbb{E}_G(\theta^m)$. This shows that left hand-side of equation 10 is strictly greater under F compared to G when $k = m-1$. Therefore, if $n_G^* \leq m$, $n_F^* \geq n_G^*$. If $n_G^* > m$, then this means equation 10 is satisfied for distribution G for any $k \leq m-1$, which implies it should be satisfied for F. This completes the proof.

Proof of Proposition 4.1

- A fully inattentive population is a distribution F with a probability mass of 1 at $\theta = 0$. The probability that such a population chooses the default option is equal to 1. Because the default option in this case corresponds to buying nothing, the firm gets a negative profit when it offers a nonempty menu. Therefore, it will choose to offer nothing and get a payoff of 0.

  On the other hand, in a fully attentive population, $p(b(S),S) = \int_0^1 \theta[1-\theta]^0 dF = \int_0^1 \theta dF = 1$ for any $S \in \mathbb{X}$.[12] Thus, the probability that such a population buys the $\succsim$-best offered product is equal to 1. Therefore, the monopolist maximizes its profit when it offers the cheapest product, which is also the $\succsim$-worst product.

- Assume that the distribution F has a probability mass of 1 at some $\theta \in (0,1)$. The aggregate probability of choosing the default is equal to $p(d,S) = [1-\theta]^{|S|}$. Therefore, the problem of the monopolist is:

  $$\underset{M\in\mathbb{X}}{\operatorname{argmax}}[1-[1-\theta]^{|S|} - c(L[|S|])]$$

  Let us denote the cost of j^{th}-cheapest product in X as $c[j]$. We deduce from the problem of the monopolist that it adds $(k+1)^{th}$ product if and only if $1-[1-\theta]^{k+1} - c(L[k+1]) > 1-[1-\theta]^k - c(L[k])$, which is equivalent to the following:

  $$\theta[1-\theta]^k > c[k+1]$$

[11] $\binom{s}{t} = \binom{s+1}{t} - \binom{s}{t-1}$ for $s \geq t$.
[12] Because of the convention that $0^0 = 1$.

For simplicity, let $f_k(\theta) = \theta[1-\theta]^k$. One can show that this function is linear if $k=0$, and strictly concave for any $k \geq 1$. The maximum of this function is achieved when θ is equal to $\frac{1}{k+1}$. Denoting this value at which the maximum is achieved as $\theta^*(k)$. We know that unless $\theta = 0$, the monopolist offers at least one product. So, consider $k \geq 1$. All $f_k(\theta)$ takes the same value at $\theta = 0$ and $\theta = 1$ when $k \geq 1$. Thus, all these functions have the same shape (strictly concave) with the maximum value point shifting left. Let $I_k(\theta) = \{\theta' \in (0,1) : f_k(\theta') \geq f_k(\theta)\}$. It is easy to show that this interval gets smaller as k increases, with the leftmost point being equal by definition (which is θ itself). Therefore, any $\hat{\theta}$ s.t. $|\hat{\theta} - \frac{1}{n^*_{\delta_\theta}}| \leq |\theta - \frac{1}{n^*_{\delta_\theta}}|$ is contained in all $I_k(\theta)$ for $k \leq n^*_\theta$. This concludes the proof.

Proof of Lemma 5.4

Assume otherwise and let $\mathbf{M} = (S,S)$ for some $S \in \mathbb{X}$. Consider the following deviation by $\mathbf{e}$ when $S \neq X$: It adds the $\succsim$-worst product not offered in the market, $w(X \setminus S)$. Denote the final market after the deviation as $\mathbf{M}'$. After this deviation, the choice probability of no product offered by $\mathbf{e}$ changes, while the choice probability of the default decreases. The change in $\mathbf{e}$'s payoff can be written as:

$$p(w(X \setminus S), \mathbf{M}') - c_{w(X\setminus S)}$$

Note that $p(w(X\setminus S),\mathbf{M}') \geq p(w(X),X) = \mathbb{E}_F(\theta[1-\theta]^{n-1})$, and by assumption the final quantity is strictly greater than c^*. This implies that the deviation is profitable because $c^* \geq c_{w(X\setminus S)}$ for any S. Since S is arbitrary, this shows that there is a symmetric equilibrium only if $S = X$. We are going to reach a contradiction by showing that $b(X) \notin S_{\mathbf{d}}$ in equilibrium, so S cannot be equal to X.

Assume to the contrary $b(X) \in S_{\mathbf{d}}$. Because $S_{\mathbf{c}} = S = X$, the resulting change in payoff if $\mathbf{m}$ deviates to $X \setminus \{b(X)\}$ is equal to:

$$c_{b(X)} - \frac{1}{2} p(b(X),X)$$

Since $p(b(X),X) = \mathbb{E}_F(\theta)$ and $\mathbb{E}_F(\theta) < c^* + c_*$, $2c_{b(X)} = 2c^* > \mathbb{E}_F(\theta)$, and hence the deviation is profitable. So, $b(X) \notin S_{\mathbf{d}}$, a contradiction. So, if there is any pure strategy equilibrium profile $(S_{\mathbf{d}}, S_{\mathbf{c}})$, then $S_{\mathbf{d}} \neq S_{\mathbf{c}}$. Assume to the contrary $S_{\mathbf{d}} \setminus S_{\mathbf{c}} \neq \emptyset$. Consider a deviation implemented by $\mathbf{m}$ from $S_{\mathbf{d}}$ to $S_{\mathbf{d}} \setminus \{y\}$ s.t. $y \in S_{\mathbf{d}} \setminus S_{\mathbf{c}}$ and let $\mathbf{M}'$ denote the resulting market. The probability of choosing default increases because the variety in the market decreases. Note that the choice probability of a product z changes only if $y \succ z$. Hence, the change in payoff of $\mathbf{m}$ resulting from the deviation is equal to:

$$[p(d,\mathbf{M}') - p(d,\mathbf{M})] + \sum_{\{z \in S_{\mathbf{d}} : y \succ z \;\&\; z \neq d\}} [p(z,\mathbf{M}') - p(z,\mathbf{M})] + c_y - p(y,\mathbf{M}) \tag{11}$$

By definition, $p(d,\mathbf{M}') - p(d,\mathbf{M}) = \mathbb{E}_F(\theta[1-\theta]^{n(\mathbf{M}')})$ where $n(\mathbf{M}') = n(\mathbf{M}) - 1$ (which is equal to the probability that $\succsim$-worst product is chosen in $\mathbf{M}$). Furthermore, $p(y,\mathbf{M}) = \mathbb{E}_F(\theta[1-\theta]^{|B(y,\mathbf{M})|}) < \mathbb{E}_F(\theta)$ and $c_y \geq c_*$ for any y. Therefore, Equation 11 is greater than or equal to:

$$\mathbb{E}_F(\theta[1-\theta]^{n-1}) + c_* - \mathbb{E}_F(\theta)$$

since $n(\mathbf{M}) \leq n$. Because we assume that $\mathbb{E}_F(\theta[1-\theta]^{n-1}) > c^*$, this expression is strictly greater than $c^* + c_* - \mathbb{E}_F(\theta)$, which is greater than or equal to 0 by assumption. This implies that the deviation is profitable. Because **M** is an arbitrary market and y is any product in $S_{\mathbf{d}} \setminus S_{\mathbf{c}}$, for **m** it is profitable to deviate from any $S_{\mathbf{d}}$ s.t. $S_{\mathbf{d}} \setminus S_{\mathbf{c}} \neq \emptyset$.

Proof of Lemma 3
It is enough to show that **e** finds profitable to deviate any product offered in common. Note that dropping such products do not change the default probability in the market because the variety remains the same. Assume to the contrary $S_{\mathbf{c}} \cap S_{\mathbf{d}} \neq \emptyset$ and consider a deviation by **e** from $S_{\mathbf{c}}$ to $S_{\mathbf{c}} \setminus S_{\mathbf{d}}$. The change in payoff is equal to:

$$c(S_{\mathbf{c}} \cap S_{\mathbf{d}}) - \frac{1}{2}p(S_{\mathbf{c}} \cap S_{\mathbf{d}}, X)$$

Note that the market is equal to X by the previous step in the proof. By definition, $c(S_{\mathbf{c}} \cap S_{\mathbf{d}}) \geq |c(S_{\mathbf{c}} \cap S_{\mathbf{d}})|c_*$ and $p(S_{\mathbf{c}} \cap S_{\mathbf{d}}, X) < |S_{\mathbf{c}} \cap S_{\mathbf{d}}|p(b(X), X) = |S_{\mathbf{c}} \cap S_{\mathbf{d}}|\mathbb{E}_F(\theta)$. Because by assumption $\mathbb{E}_F(\theta) \leq 2c_*$, the deviation is profitable.

Proof of Proposition 5.2
Let us first show that no firms offer a menu with multiple products. Assume that $M \in S(\sigma_i)$ s.t. $|M| \geq 2$. For any E offered by the competitor, the total share it gets from offering $M \setminus \{b(M)\}$ is bounded above by $1 - \mathbb{E}(\theta)$, because $\mathbb{E}(\theta)$ is equal to the choice probability of the best product offered in a market (M, E). Since $\mathbb{E}(\theta) > 1 - c_*$, $1 - \mathbb{E}(\theta) < c_*$, and therefore a deviation in which M drops any product but $b(M)$ is profitable. Hence, no firm offers a menu with multiple products in equilibrium and we have $S(\sigma_i) \subseteq \{\{x\} : x \in X\}$. Observe that all firms offer the same $\succsim$-worst product in equilibrium, because otherwise the firm offering the worse product would get a negative payoff by the assumption $\mathbb{E}(\theta) > 1 - c_*$. Let us denote this product with w. The payoff firm i receives from $\{w\}$ is equal to $\frac{1}{2}\mathbb{E}(\theta)\sigma_{\mathbf{d}}(\{w\}) - c_w$. If **c** deviates to $\{b(X)\}$ from $\{w\}$, then the resulting payoff can be written as:

$$[1 - \sigma_{\mathbf{d}}(\{b(X)\})]\mathbb{E}(\theta) + \frac{1}{2}\sigma_{\mathbf{d}}(\{b(X)\})\mathbb{E}(\theta) - c_{b(X)}$$

which is equal to $1 - \frac{\mathbb{E}(\theta)}{2}\sigma_{\mathbf{d}}(\{b(X)\}) - c_{b(X)}$. Therefore, the change in payoff is equal to:

$$1 - \frac{\mathbb{E}(\theta)}{2}[\sigma_{\mathbf{d}}(\{b(X)\}) + \sigma_{\mathbf{d}}(\{w\})] - [c_{b(X)} - c_w]$$

Because $\sigma_{\mathbf{d}}(\{b(X)\}) + \sigma_{\mathbf{d}}(\{w\}) \leq 1$ and $c^* - c_* < \frac{1}{2}$, this change is strictly positive and therefore the deviation is profitable. This shows that in equilibrium, the competitor offers $\{b(X)\}$. The best-response of **d** to this is offering $\{b(X)\}$, because if **d** offers another product, it will get a negative payoff and gets a zero payoff if it offers none.

Proof of Theorem 2
The proof is demonstrated in several steps.

Step 1: There is no symmetric equilibrium in pure strategies.

Follows by Lemma 2 because $2c_* < c^* + c_*$.

Step 2: $S_{\mathbf{d}} \subset S_{\mathbf{c}}$.

This follows by Lemma 2.

Step 3: $S_{\mathbf{c}} \supseteq X \setminus S_{\mathbf{d}}$.

Consider a deviation by firm $\mathbf{e}$ to $S_{\mathbf{c}} \cup \{y\}$ s.t. $y \notin \bigcup \mathbf{M}$, and denote the resulting market as $\mathbf{M}_y$. This deviation is variety-increasing, and hence $p(d,\mathbf{M}) > p(d,\mathbf{M}_y)$. The change in the payoff of firm $\mathbf{e}$ is equal to:

$$\frac{1}{2}[p(S_{\mathbf{d}},\mathbf{M}_y) - p(S_{\mathbf{d}},\mathbf{M})] + [p(S_{\mathbf{c}} \setminus S_{\mathbf{d}},\mathbf{M}_y) - p(S_{\mathbf{c}} \setminus S_{\mathbf{d}},\mathbf{M})] + [p(y,\mathbf{M}_y) - c_y] \tag{12}$$

Note that Equation 12 is greater than equal to $\frac{1}{2}[p(S_{\mathbf{c}},\mathbf{M}_y) - p(S_{\mathbf{c}},\mathbf{M})] + [p(y,\mathbf{M}_y) - c_y]$. Furthermore, $p(S_{\mathbf{c}},\mathbf{M}_y) = 1 - p(d,\mathbf{M}_y) - p(y,\mathbf{M}_y)$ and $p(S_{\mathbf{c}},\mathbf{M}) = 1 - p(d,\mathbf{M})$. This implies that Equation 12 is greater than or equal to:

$$\frac{1}{2}[p(d,\mathbf{M}) - p(d,\mathbf{M}_y)] + \frac{1}{2}p(y,\mathbf{M}_y) - c_y \tag{13}$$

By definition, $p(d,\mathbf{M}) - p(d,\mathbf{M}_y) = \mathbb{E}_F(\theta[1-\theta]^{n(\mathbf{M})})$. Furthermore, because $n(\mathbf{M}) \leq n-1$, $\mathbb{E}_F(\theta[1-\theta]^{n(\mathbf{M})}) \geq \mathbb{E}_F(\theta[1-\theta]^{n-1})$. Since $p(y,\mathbf{M}_y) \geq p(w(\mathbf{M}_y),\mathbf{M}_y) = \mathbb{E}_F(\theta[1-\theta]^{n(\mathbf{M})})$, Equation 13 is greater than or equal to $\mathbb{E}_F(\theta[1-\theta]^{n(\mathbf{M})}) - c_y$. By the assumption that $\mathbb{E}_F(\theta[1-\theta]^{n-1}) > c^*$, the deviation is profitable. This implies the conclusion using the previous Lemma.

In the proof of the first Lemma, we showed that $b(X) \notin S_{\mathbf{d}}$. The same proof is valid also here because $S_{\mathbf{c}} = X$, so $b(X) \notin S_{\mathbf{d}}$.

Step 3: $S_{\mathbf{c}} \cap S_{\mathbf{d}} = \emptyset$.

It is enough to show that $\mathbf{e}$ finds profitable to deviate any product offered in common. Note that dropping such products do not change the default probability in the market because the variety remains the same. Assume to the contrary $S_{\mathbf{c}} \cap S_{\mathbf{d}} \neq \emptyset$ and consider a deviation by $\mathbf{e}$ from $S_{\mathbf{c}}$ to $S_{\mathbf{c}} \setminus S_{\mathbf{d}}$. The change in payoff is equal to:

$$c(S_{\mathbf{c}} \cap S_{\mathbf{d}}) - \frac{1}{2}p(S_{\mathbf{c}} \cap S_{\mathbf{d}},X)$$

Note that the market is equal to X by the previous step in the proof. By definition, $c(S_{\mathbf{c}} \cap S_{\mathbf{d}}) \geq |c(S_{\mathbf{c}} \cap S_{\mathbf{d}})|c_*$ and $p(S_{\mathbf{c}} \cap S_{\mathbf{d}},X) < |S_{\mathbf{c}} \cap S_{\mathbf{d}}|p(b(X),X) = |S_{\mathbf{c}} \cap S_{\mathbf{d}}|\mathbb{E}_F(\theta)$. Because by assumption $\mathbb{E}_F(\theta) \leq 2c_*$, the deviation is profitable.

Thus, we have shown that if there is a pure strategy equilibrium $(S_{\mathbf{d}},S_{\mathbf{c}})$, then $(S_{\mathbf{d}},S_{\mathbf{c}})$ is a partition of X. In the next step, we show that under the assumptions of this result, the only partition that can be an equilibrium is $(\{d\},X)$ (Recall that $\{d\}$ is always offered by the monopolist, so the monopolist offers no 'new' products from X.)

Step 4: $S_{\mathbf{d}} = \{d\}$.

Consider a deviation from $S_{\mathbf{d}}$ to $S_{\mathbf{d}} \setminus \{x\}$. The change in the payoff is equal to:

$$c_x - p(x,X) + [p(d,X \setminus \{x\}) - p(d,X)]$$

Note that $p(d,X \setminus \{x\}) - p(d,X) = \mathbb{E}_F(\theta[1-\theta]^{n-1})$. By definition, $p(x,X) = \mathbb{E}_F(\theta[1-\theta]^{|B(x,X)|})$. Therefore, the change in payoff is equal to:

$$c_x - \mathbb{E}_F(\theta[1-\theta]^{|B(x,X)|}) + \mathbb{E}_F(\theta[1-\theta]^{n-1})$$

The deviation is not profitable if and only if this expression is less than or equal to 0. Because $\mathbb{E}_F(\theta[1-\theta]^{|B(x,X)|}) \leq \mathbb{E}_F(\theta)$, if the deviation is not profitable we need to have:

$$c_x \leq \mathbb{E}_F(\theta) - \mathbb{E}_F(\theta[1-\theta]^{n-1})$$

By assumption, $\mathbb{E}_F(\theta) < 2c_*$ and $\mathbb{E}_F(\theta[1-\theta]^{n-1}) > c^*$, which implies further the following:

$$c_x < 2c_* - c^*$$

This is easiest to satisfy for the cheapest product, $w(X)$. But this does not hold even for $w(X)$, because it would imply $c^* < c_*$, a contradiction.

Step 5: $(\{d\},X)$ is an equilibrium.

Finally, we demonstrate that this is indeed an equilibrium by showing that there is no profitable deviation. To see, we first show that $\mathbf{m}$ does not have any profitable deviation to some $\sigma_{\mathbf{d}}$. Assume to the contrary the deviation is profitable to some mixed strategy $\sigma_{\mathbf{d}}$. For any $M \in S(\sigma_{\mathbf{d}})$, the payoff to this menu is equal to $p(d,X) + \frac{1}{2}p(M,X) - c(M)$, which is strictly greater than $p(d,X)$. Thus, $p(M,X) > 2c(M)$. The first term is less than or equal to $|M|\mathbb{E}_F(\theta)$, while the second term is greater than or equal to $2|M|c_*$. So, we have $\mathbb{E}_F(\theta) > 2c_*$, a contradiction.

Now, consider a deviation by $\mathbf{e}$ to some $\sigma_{\mathbf{c}}$. Assume to the contrary it is profitable. The payoff to any $E \in S(\sigma_{\mathbf{c}})$ is equal to $p(E,E) - c(E)$. A deviation from E to $E \cup \{x\}$ is profitable. To see, note that the change in payoff if $\mathbf{e}$ implements this deviation is equal to:

$$[p(E,(\{d\},E\cup\{x\})) - p(E,(\{d\},E))] + [p(x,(\{d\},E\cup\{x\})) - c_x] \tag{14}$$

The first term in the above equation is equal to $p(d,(\{d\},E)) - p(d,(\{d\},E\cup\{x\}))$. This is equal to $\mathbb{E}_F(\theta[1-\theta]^{|E|})$, which is greater than or equal to $\mathbb{E}_F(\theta[1-\theta]^{n-1})$, and hence strictly greater compared to c^*. This implies the deviation is profitable because $p(x,(\{d\},E\cup\{x\})) > 0$ and $c^* \geq c_x$ for all $x \in X$. Because E is arbitrary, this implies that we can find a sequence of profitable deviations until $\mathbf{e}$ offers X. But when X is offered in a mixed strategy, $\mathbf{e}$ gets the same payoff as it gets with the pure strategy of offering X, a contradiction.

Proof of Theorem 5.3

Assume to the contrary firms play a mixed strategy profile $\sigma = (\sigma_{\mathbf{i}}, \sigma_{\mathbf{c}})$. Consider the incumbent. If it does not extend the product line (offering $\{d\}$ in the support), then the payoff of doing this should be equal to the payoff of any menu $M \in S(\sigma_{\mathbf{i}})$. If it extends the product line in any case, then the payoff to $\{d\}$ should be less than or equal to the payoff to such an M. The firm gets a payoff of $\alpha_d(\sigma_{\mathbf{c}})$ from not extending the line, while its payoff for extending the line with M is equal to:

$$\alpha_d(M,\sigma_c) + \sum_{S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E)[p(M\setminus E, M\cup E) + \frac{1}{2}p(M\cap E, M\cup E) - c(M) \tag{15}$$

First, let us show that $\alpha_d(\sigma_{\mathbf{c}}) - \alpha_d(M,\sigma_{\mathbf{c}}) \geq \alpha_d(\sigma_{\mathbf{c}}) \times (1 - \mathscr{B}(|M|))$. We can write $\alpha_d(\sigma_{\mathbf{c}}) - \alpha_d(M,\sigma_{\mathbf{c}})$ as $\mathbb{E}([1-\theta]^{|E|}(1-[1-\theta]^{|M\cup E|-|E|}))$. It is straightforward to show that the term inside the expectations is convex, but it requires a bit of work related to algebraic manipulations, so we omit that part. Given the convexity, Jensen's inequality implies that:

$$\begin{aligned}\alpha_d(\sigma_{\mathbf{c}}) - \alpha_d(M,\sigma_{\mathbf{c}}) &\geq \sum_{S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E)[1-\mathbb{E}(\theta)]^{|E|}(1-[1-\mathbb{E}(\theta)]^{|M\cup E|-|E|})\\ &= \alpha_d(\sigma_{\mathbf{c}}) \times (1 - \mathscr{B}(|M\cup E|))\end{aligned}$$

Observe that $\mathscr{B}(|M\cup E|) \geq \max\{\mathscr{B}(|M|), \mathscr{B}(|E|)\}$. Furthermore, in the main body we demonstrated that:

$$\alpha_d(\sigma_{\mathbf{c}}) \geq 1 - \mathscr{B}(\sigma)$$

The rest of Expression 15 is less than or equal to:

$$\sum_{S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E)p(M, M \cup E) - c(M)$$

which is maximized if the products in M are the best $|M|$ products in $M \cup E$, in which case they would be equal to:

$$\sum_{i=0}^{|M|-1} \mathbb{E}(\theta[1-\theta]^i) - c(M)$$

Note that $\mathbb{E}(\theta[1-\theta]^i) = 1 - \mathscr{B}(|M|)$. Since the payoff to $\{d\}$ should be less than or equal to the payoff to M, we need to have:

$$1 - \mathscr{B}(|M|) - c(M) \geq (1 - \mathscr{B}(\sigma))(1 - \mathscr{B}(|M \cup E|))$$

which implies that:

$$\mathscr{B}(\sigma) \geq \frac{\mathscr{B}(|M|) - \mathscr{B}(|M \cup E|) + c(M)}{1 - \mathscr{B}(|M \cup E|)}$$

a contradiction to our assumption.

Proof of Theorem 3

- Let $\sigma = (\sigma_{\mathbf{d}}, \sigma_{\mathbf{c}})$ denote a mixed strategy profile. For a menu M offered in $S(\sigma_{\mathbf{d}})$, the payoff $\mathbf{m}$ receives is equal to:

$$\sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E)p(M \setminus E, (M, E)) + \frac{1}{2} \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E)p(M \cap E, (M, E)) + \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_e(E)p(d, (M, E)) - c(M) \tag{16}$$

  while the payoff $\mathbf{e}$ receives from offering E is given by:

$$\sum_{M \in S(\sigma_{\mathbf{d}})} \sigma_{\mathbf{d}}(M)p(E \setminus M, (M, E)) + \frac{1}{2} \sum_{M \in S(\sigma_{\mathbf{d}})} \sigma_{\mathbf{d}}(M)p(E \cap M, (M, E)) - c(E) \tag{17}$$

  We are going to demonstrate the proof using two lemmas.

  Lemma 5. *If $\mathbb{E}_F(\theta[1-\theta]^{n-1}) > c^*$, then $X \setminus \bigcup S(\sigma_{\mathbf{d}}) \subseteq \bigcap S(\sigma_{\mathbf{c}})$.*

  Proof. Assume to the contrary $x \notin E$ for some $x \in X \setminus \bigcup S(\sigma_{\mathbf{d}})$. Consider a deviation by $\mathbf{e}$ to $S_{\mathbf{c}} \cup \{x\}$. The change in payoff is equal to:

$$\sum_{\{M \in S(\sigma_{\mathbf{d}}): x \notin M\}} \sigma_{\mathbf{d}}(M)[p(E \setminus M, (M, E \cup \{x\})) - p(E \setminus M, (M, E))] + \frac{1}{2} \sum_{\{M \in S(\sigma_{\mathbf{d}}): x \in M\}} \sigma_{\mathbf{d}}(M)[p(E \cap M, (M, E \cup \{x\})) - p(E \cap M, (M, E))] + [\sum_{M \in S(\sigma_{\mathbf{d}})} \sigma_{\mathbf{d}}(M)p(x, (M, E \cup \{x\})) - c_x]$$

This equation is greater than or equal to:

$$\frac{1}{2}\sum_{M\in S(\sigma_{\mathbf{d}})}\sigma_{\mathbf{d}}(M)[p(E,(M,E\cup\{x\}))-p(E,(M,E))]$$
$$+[\sum_{M\in S(\sigma_{\mathbf{d}})}\sigma_{\mathbf{d}}(M)p(x,(M,E\cup\{x\}))-c_x]$$

Because $p(E,(M,E\cup\{x\}))=1-p(d,(M,E\cup\{x\}))-p(x,(M,E\cup\{x\}))$ and $p(E,(M,E))=1-p(d,(M,E))$, the above equation is equal to the following:

$$\begin{aligned}&\frac{1}{2}\sum_{M\in S(\sigma_{\mathbf{d}})}\sigma_{\mathbf{d}}(M)[p(d,(M,E))-p(d,(M,E\cup\{x\}))]\\&+\frac{1}{2}\sum_{M\in S(\sigma_{\mathbf{d}})}\sigma_{\mathbf{d}}(M)p(x,(M,E\cup\{x\}))-c_x\end{aligned}\tag{18}$$

By definition, $p(d,(M,E))-p(d,(M,E\cup\{x\}))=\mathbb{E}_F(\theta[1-\theta]^{n((M,E))})$ and $p(x,(M,E\cup\{x\}))=\mathbb{E}_F(\theta[1-\theta]^{|B(x,(M,E\cup\{x\}))|})$. Since $|B(x,(M,E\cup\{x\}))|\leq n((M,E))\leq n-1$, both of these expressions are greater than or equal to $\mathbb{E}_F(\theta[1-\theta]^{n-1})$, which implies that Equation 18 is greater than or equal to:

$$\mathbb{E}_F(\theta[1-\theta]^{n-1})-c_x$$

By the assumption that $\mathbb{E}_F(\theta[1-\theta]^{n-1})>c^*$, this is strictly greater than 0, so the deviation is profitable, a contradiction. $\square$

Lemma 6. *If $\mathbb{E}_F(\theta\{1-[1-\theta]^{n-1}\})<2c_*$, then $\bigcup S(\sigma_{\mathbf{d}})\subseteq\bigcup S(\sigma_{\mathbf{c}})$.*

Proof. Since $\mathbb{E}_F(\theta[1-\theta]^{n-1})>c^*$ and $\mathbb{E}_F(\theta)<2c_*+c^*$, we have $\mathbb{E}_F(\theta\{1-[1-\theta]^{n-1}\})<2c_*$. Assume to the contrary $\bigcup S(\sigma_{\mathbf{d}})\setminus\bigcup S(\sigma_{\mathbf{c}})\neq\emptyset$, and let x be a product in this set. Consider a deviation by $\mathbf{m}$ in which it drops x from a menu $M\in S(\sigma_{\mathbf{d}})$ s.t. $x\in M$. First, note that the default probability in all markets increases because this deviation decreases the variety in all markets. The change in default probability is equal to:

$$\sum_{E\in S(\sigma_{\mathbf{c}})}\sigma_{\mathbf{c}}(E)[p(d,(M\setminus\{x\},E)-p(d,(M,E))]\tag{19}$$

On the other hand, because $\mathbf{m}$ drops x, it loses any share due to offering this product:

$$-\sum_{E\in S(\sigma_{\mathbf{c}})}\sigma_{\mathbf{c}}(E)p(x,(M,E))=-\mathbb{E}_F(\theta)\tag{20}$$

In addition to these, the choice probability of products that are $\succsim$-worse than x change. These products either offered by $\mathbf{m}$ alone, or by both firms. Therefore, the change in payoff can be written as follows:

$$\sum_{E\in S(\sigma_{\mathbf{c}})}\sigma_{\mathbf{c}}(E)[p(M\setminus(E\cup\{x\}),(M\setminus\{x\},E))-p(M\setminus(E\cup\{x\}),(M,E))]$$
$$+\frac{1}{2}\sum_{E\in S(\sigma_{\mathbf{c}})}\sigma_{\mathbf{c}}(E)[p((M\setminus\{x\})\cap E,(M\setminus\{x\},E)-p(M\setminus\{x\})\cap E,(M,E))]$$

Note that $p(M \setminus (E \cup \{x\}), (M \setminus \{x\}, E)) = 1 - p(d, (M \setminus \{x\}, E))$ and $p(M \setminus (E \cup \{x\}), (M, E)) = 1 - p(d, (M, E)) - p(x, (M, E))$. Therefore, the above equation is greater than or equal to:

$$\frac{1}{2} \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E)[p(d,(M,E)) - p(d,(M \setminus \{x\},E))] + \frac{1}{2} \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E) p(x,(M,E)) \quad (21)$$

Combining Equations 19,20, and 21, using the definitions, and because $c_x \geq c_*$, we conclude that the change in payoff is greater than equal to the following:

$$\frac{1}{2} \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E)[\mathbb{E}_F(\theta[1-\theta]^{n((M \setminus \{x\},E))}) - \mathbb{E}_F(\theta[1-\theta]^{|B(x,(M,E))|})] + c_*$$

Note that $\mathbb{E}_F(\theta\{1-[1-\theta]^{n-1}\}) = \mathbb{E}_F(\theta) - \mathbb{E}_F(\theta[1-\theta]^{n-1}\})$ is greater than or equal to $\mathbb{E}_F(\theta[1-\theta]^{n((M \setminus \{x\},E))}) - \mathbb{E}_F(\theta[1-\theta]^{|B(x,(M,E))|})$ for any (M,E). Thus, the assumption that $\mathbb{E}_F(\theta\{1-[1-\theta]^{n-1}\}) < 2c_*$ implies that this final equation is strictly positive, and therefore the deviation is profitable. This concludes the proof. □

To conclude the proof of the proposition, note that by Lemma 5 we have $X \setminus \bigcup S(\sigma_{\mathbf{d}}) \subseteq \bigcap S(\sigma_{\mathbf{c}})$. Furthermore, we showed in Lemma 6 that $\bigcup S(\sigma_{\mathbf{d}}) \subseteq \bigcup S(\sigma_{\mathbf{c}})$. These two imply that $\bigcup S(\sigma_{\mathbf{c}}) \supseteq X \setminus \bigcup S(\sigma_{\mathbf{c}})$, which is possible only if $\bigcup S(\sigma_{\mathbf{c}}) = X$.

- Let us denote a generic mixed strategy equilibrium as $\sigma = (\sigma_{\mathbf{d}}, \sigma_{\mathbf{c}})$, and define the following object which we will use in the next result:

$$\alpha_d(M, \sigma_{\mathbf{c}}) = \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E) \mathbb{E}([1-\theta]^{|M \cup E|}) \quad (22)$$

We are going to show that $\{d\} \in S(\sigma_{\mathbf{d}})$. Assume it does not offer $\{d\}$, which is true only if $\mathbf{m}$ does not find any deviation from $M \in S(\sigma_{\mathbf{d}})$ to $\{d\}$ profitable. The payoff to offering $\{d\}$ is equal to $\alpha_d(\sigma_{\mathbf{c}})$. The payoff of $\mathbf{m}$ from any $M \neq \{d\}$ offered in the support is given by:

$$\begin{aligned} \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E) p(M \setminus E, (M,E)) + \frac{1}{2} \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E) p(M \cap E, (M,E)) \\ + \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_e(E) p(d,(M,E)) - c(M) \end{aligned} \quad (23)$$

Thus, for the deviation not to be profitable the following condition needs to be satisfied:

$$\alpha_d(\{d\}, \sigma_{\mathbf{c}}) - \alpha_d(M, \sigma_{\mathbf{c}}) \leq \frac{1}{2} \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E) p(M \cap E, (M,E)) + \sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E) p(M \setminus E, (M,E)) - c(M)$$

The right-hand side of this inequality is strictly less than $\sum_{E \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(E) p(M,(M,E)) - c(M)$. Note that $p(M,(M,E)) \leq 1 - p(d,(M,E))$, so we get the following inequality:

$$\alpha_d(\{d\}, \sigma_{\mathbf{c}}) - \alpha_d(M, \sigma_{\mathbf{c}}) < 1 - \alpha_d(M, \sigma_{\mathbf{c}}) - c(M)$$

and hence $\alpha_d(\{d\}, \sigma_{\mathbf{c}}) < 1 - c(M)$ for any $M \in S(\sigma_{\mathbf{d}})$. By definition, $\alpha_d(\{d\}, \sigma_{\mathbf{c}}) = \mathbb{E}_F(\mathbb{E}_{\sigma_{\mathbf{c}}}([1-\theta]^{|E|}))$, so $\alpha_d(\{d\}, \sigma_{\mathbf{c}}) \geq \mathbb{E}_F([1-\theta]^n)$. By Jensen's inequality and the observation that $[1-\theta]^n$ is a convex function (since $n \geq 2$), $\mathbb{E}_F([1-\theta]^n) \geq (1 - \mathbb{E}_F(\theta))^n$. So, we need to have:

$$(1 - \mathbb{E}_F(\theta))^n < 1 - c(X)$$

which contradicts with the assumption in the proposition, so $\{d\} \in S(\sigma_{\mathbf{d}})$.

Proof of Proposition 5.4
We first show that any pure strategy equilibrium (if it exists) should be either symmetric with each firm offering a single product or the entrant offers the cheapest product and the monopolist does not offer a new product. First, note that because high type only buys the $\succsim$-best product offered in the market and low type only goes for the default, offering more than a single product only increases the cost of the firm. Therefore, both firms offer at most a single product. Furthermore, note that if firms offer different singletons, then only one of them is sold to the high type in the market. Therefore, the only possibility is that they offer the same product and share the high-type consumers equally. So, if there is a pure strategy equilibrium, then the strategy profile is either $(\{x\},\{x\})$ or $(\{d\},\{x\})$ for some $x \in X$.

First, let us consider the former possibility. Note that unless $x = b(X)$, **e** can profitably deviate to $\{b(X)\}$. To wit, note that the change in payoff resulting from this deviation is equal to $[p(b(X),(\{x\},\{b(X)\})) - \frac{1}{2}p(x,(\{x\},\{x\}))] - [c^* - c_x]$. This is equal to:

$$\frac{1}{2}\lambda - [c^* - c_x]$$

which is strictly positive because $\lambda > \max\{2(c^* - c_*), 2c_*\}$. Thus, the only possible pure strategy equilibrium profile in this form is $(\{b(X)\},\{b(X)\})$. But from this market, **m** can profitably deviate to offering no new products, i.e. to $\{d\}$, since the change in payoff is equal to:

$$c^* - \frac{1}{2}\lambda$$

which is also profitable because $\lambda < 2c^*$. This shows that $(\{x\},\{x\})$ for any $x \in X$ cannot constitute a pure strategy equilibrium.

Now let us look at the second possibility in which the entrant offers $\{x\}$ for some $x \in X$ and the monopolist offers $\{d\}$. Clearly, **e** has always a profitable deviation to $\{w(X)\}$ from any $x \neq w(X)$. But when the market is $(\{d\},\{w(X)\})$, **m** can profitably deviate from $\{d\}$ to $\{w(X)\}$ because $\lambda > \max\{2(c^* - c_*), 2c_*\}$. This concludes the proof.

Proof of Proposition 5.5
We are going to demonstrate the proof in several steps.

Step 1: $|M|,|E| = 1\ \forall M \in S(\sigma_{\mathbf{c}})\ \&\ E \in S(\sigma_{\mathbf{d}})$.

When the population distribution is F_λ, the high-type consumers only purchase the $\succsim$-best product offered, while low-types only buy the default. This implies that in equilibrium a firm does not offer more than a single product.

Step 2: $w(S_{\mathbf{d}}) = w(S_{\mathbf{c}})$.

Assume this is not the case. If $w(S_{\mathbf{d}}) \succ w(S_{\mathbf{c}})$, then the payoff of **e** is equal to $-c_{w(S_{\mathbf{c}})}$, so this cannot be the case because **e** would deviate. If $w(S_{\mathbf{c}}) \succ w(S_{\mathbf{d}})$, then **m** gets $1 - \lambda - c_{w(S_{\mathbf{d}})}$, which is strictly less than $1 - \lambda$ and therefore cannot be the case. So, $w(S_{\mathbf{d}}) \sim w(S_{\mathbf{c}})$. Because any menu offered in the equilibrium is a singleton, $w(S_{\mathbf{d}}) = w(S_{\mathbf{c}})$.

Step 2: $b(S_{\mathbf{c}}) \succsim b(S_{\mathbf{d}})$.

Assume to the contrary $b(S_{\mathbf{d}}) \succ b(S_{\mathbf{c}})$. In this case, **m** receives a payoff of $1 - c_{b(S_{\mathbf{d}})}$, which is strictly greater than $1 - \lambda$ provided that $\lambda > c^*$. This shows that if there is a mixed strategy equilibrium and $\{d\} \in S(\sigma_{\mathbf{d}})$, then $b(S_{\mathbf{c}}) \succsim b(S_{\mathbf{d}})$.

Furthermore, note that it is straightforward to show that **e** has a profitable deviation to the $\succsim$-next best product, which implies that if $b(S_{\mathbf{c}}) \succ b(S_{\mathbf{d}})$, then the products offered by **e** $\succsim$-better than $b(S_{\mathbf{d}})$ are $b(S_{\mathbf{d}})$ and $\succsim$-next best product to $b(S_{\mathbf{d}})$. The other possibility is having $b(S_{\mathbf{c}}) = b(S_{\mathbf{d}})$. We next show that this is the case when $\lambda > 2(c^* - c_*)$.

Step 3: $\lambda > 2(c^* - c_*) \implies b(S_{\mathbf{c}}) = b(S_{\mathbf{d}})$.

Assume to the contrary $b(S_{\mathbf{c}}) \succ b(S_{\mathbf{d}})$. **e** receives a payoff of $\lambda - c_{b(S_{\mathbf{c}})}$ from $\{b(S_{\mathbf{c}})\}$. On the other hand, **e** receives $\frac{\lambda}{2}\sigma_{\mathbf{d}}(\{w\}) - c_w$ from $\{w\}$. The equality of these imply that:

$$\sigma_{\mathbf{d}}(\{w\}) = 2 - \frac{2}{\lambda}(c^* - c_w)$$

which is greater than or equal to 1 because $\lambda > 2(c^* - c_*)$. This is a contradiction to the fact that $\sigma_{\mathbf{d}}(\{d\}) > 0$.

From now on, we simply denote $b(S_{\mathbf{d}}) = b(S_{\mathbf{c}})$ as b.

Step 4: $S(\sigma_m) = \{\{b\}, \{d\}\}$**.**

To show this, it is enough to show that **m** never offers $\{w\}$. Note that the payoff to $\{b\}$ is equal to $1 - \frac{\lambda}{2}\sigma_{\mathbf{c}}(\{b\}) - c_b$, while the payoff to $\{w\}$ is $\frac{\lambda}{2}\sigma_{\mathbf{c}}(\{w\}) + (1-\lambda) - c_w$. These are equal to each other if and only if:

$$\lambda - (c_b - c_w) = \frac{\lambda}{2}(\sigma_{\mathbf{c}}(\{b\}) + \sigma_{\mathbf{c}}(\{w\}))$$

Because $\sigma_{\mathbf{c}}(\{b\}) + \sigma_{\mathbf{c}}(\{w\}) \leq 1$, the term on the right is less than or equal to $\frac{\lambda}{2}$. On the other hand, by assumption $\lambda > 2(c^* - c_*)$, which implies the left-hand side is strictly more than $\frac{\lambda}{2}$, a contradiction. Therefore, **m** only offers a single product $\{x\}$ and no new products ($\{d\}$).

Step 5: $S(\sigma_e) = \{\{b\}, \{w(X)\}\}$**.**

First, let us note that given $S(\sigma_m) = \{\{b\}, \{d\}\}$, the payoff to any $\{x\}$ offered by **e** is equal to:

$$\begin{aligned} \frac{\lambda}{2}(1 + \sigma_{\mathbf{d}}(\{d\})) - c_x x &= b \\ \lambda - c_x x &\succ b \\ \lambda\sigma_{\mathbf{d}}(\{d\}) - c_x b &\succ x \end{aligned}$$

Note that if **e** offers any x s.t. $b \succ x$, then this will be equal to $w(X)$. We also know that the best products offered by both firms are equal, so in equilibrium **e** offers no x s.t. $x \succ b$.

Now assume to the contrary **e** offers a single product in equilibrium, and this needs to be equal to b. If this is the case, $\sigma_{\mathbf{d}}$ violates the indifference condition, because the payoff of **m** from $\{b\}$ is strictly more than its payoff from $\{d\}$. Thus, for $(\sigma_{\mathbf{d}}, \sigma_{\mathbf{c}})$ to constitute an equilibrium with the properties we showed previously, $|S(\sigma_{\mathbf{c}})| > 1$. By what we remarked previously, **e** does not offer any product strictly better than b, and it offers $w(X)$ from any x it offers s.t. $b \succ x$. Thus, $S(\sigma_{\mathbf{c}}) = \{\{b\}, \{w(X)\}\}$.

.

Step 6: Derivation of the Mixed Strategy

Because **m** gets in equilibrium $1 - \lambda$ by Corollary to Proposition **??** and it gets $1 - \frac{\lambda}{2}\sigma_{\mathbf{c}}(\{b\}) - c_b$ from offering $\{b\}$, the equality of these imply that:

$$\sigma_{\mathbf{c}}(\{b\}) = \frac{2(\lambda - c_b)}{\lambda}$$

which implies that:

$$\sigma_{\mathbf{c}}(\{w(X)\}) = \frac{2c_b}{\lambda} - 1$$

Observe that since $\lambda > 2c^*$, this probability is strictly less than 1 We can easily derive the probabilities **m** attaching to the menus in the support. Note that **e** receives from offering $\{b\}$ the following payoff:

$$\lambda - \frac{\lambda}{2}\sigma_{\mathbf{d}}(\{b\}) - c_b$$

while her payoff from $\{w(X)\}$ is equal to:

$$\lambda \sigma_{\mathbf{d}}(\{d\}) - c_*$$

which implies using the fact that $\sigma_{\mathbf{d}}(\{d\}) = 1 - \sigma_{\mathbf{d}}(\{b\})$:

$$\sigma_{\mathbf{d}}(\{d\}) = \frac{\lambda - 2(c_b - c_*)}{\lambda}$$

and thus:

$$\sigma_{\mathbf{d}}(\{b\}) = \frac{2(c_b - c_w)}{\lambda}$$

The payoff of **m** from $\{b\}$ is equal to:

$$1 - \frac{\lambda}{2}\sigma_{\mathbf{c}}(\{b\}) - c_b$$

Observe that this is strictly greater than 0 because $\lambda > c^*$. Also, it is strictly less than 1 because $\lambda < 2c^*$.

Step 7: $b = b(X)$

If $b \neq b(X)$, **m** can deviate from $\{b\}$ to $\{b(X)\}$ and increase its payoff from $1 - \lambda$ to $1 - c^*$ because $\lambda > c^*$. Therefore, in equilibrium $b = b(X)$.

Step 8: $(\sigma_{\mathbf{d}}, \sigma_{\mathbf{c}})$ is the unique equilibrium.

We conclude the proof by showing that the described profile is the unique equilibrium of the game. Let us start with the monopolist. We know that it always offers $\{d\}$, and therefore checking a deviation from $\{b(X)\}$ is sufficient. For any pure strategy $\{x\}$ s.t. $x \notin \{d, b(X)\}$, the payoff **m** gets:

$$(1 - \lambda) + \frac{\lambda}{2}\sigma_{\mathbf{c}}(\{w(X)\}) - c_* x = w(X)$$
$$(1 - \lambda) + \lambda \sigma_{\mathbf{c}}(\{w(X)\}) - c_x x \neq w(X)$$

We know that $\sigma_{\mathbf{c}}(\{w(X)\}) = \frac{2c^*}{\lambda} - 1$. If $x = w(X)$, this implies that the change in the payoff resulting from the deviation is equal to:

$$\frac{\lambda}{2} \times (\frac{2c^*}{\lambda} - 1) - c_*$$

which is strictly less than 0 because $\lambda > 2(c^* - c_*)$. Similarly, one can show that the deviation is not profitable for the other case. Showing the mixed strategy case is similar, so there is no profitable deviation for **m**.

For $\mathbf{e}$, a deviation to the $\{\emptyset\}$ is clearly not profitable. Similarly, a deviation to $\{x\}$ s.t. $x \notin \{w(X), b(X)\}$ would result in a payoff of $\lambda \sigma_{\mathbf{d}}(\{d\}) - c_x$. Since $\sigma_{\mathbf{d}}(\{d\}) = \frac{\lambda - 2(c^* - c_*)}{\lambda}$, this is equal to $\lambda - 2(c^* - c_*) - c_x$, which is strictly less than the payoff $\mathbf{e}$ receives under $(\sigma_{\mathbf{d}}, \sigma_{\mathbf{c}})$. The mixed strategy case is similar to show. So, this profile is indeed an equilibrium.

For uniqueness, recall that we showed previously that under the condition of this theorem there is no pure strategy equilibrium. Throughout the proof, we demonstrated that in any mixed strategy equilibrium the structure of the supports of the firms are the same: Both firms offer one common product which is the best product offered. We showed $\mathbf{m}$ always offers $\{d\}$, and there is a profitable deviation unless the best product offered by firms is equal to $b(X)$. Finally, we also demonstrated that $\mathbf{e}$ can only offer $w(X)$ in addition to $b(X)$, showing the uniqueness.

Proof of Proposition 6.3

We first show that in any equilibrium, the monopolist offers a menu M in the support of its mixed strategy iff $M = L_{|M|-1} \cup \{b(M)\}$. Furthermore, $b(M) = b(X)$. Showing the former claim is straightforward, so let us show the latter. Assume it is not true, and consider replacing $b(M)$ with $b(X)$. Observe that $\mathbf{e}$ has an incentive to deviate to a smaller menu, because this decreases the share of people who go to the default firm and increase the payoff of $\mathbf{e}$. Thus, in equilibrium, the entrant only offers singleton menus. If $\{b(X)\} \notin S(\sigma_{\mathbf{c}})$, then the change in payoff after deviation is given by:

$$1 - \sum_{\{x\} \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(\{x\})(1 - \Lambda(|M \cup \{x\}|))S_{\mathbf{c}}(M, \{x\}) - (c^* - c_{b(M)})$$

where $S_{\mathbf{c}}(M, \{x\})$ denotes the share of the entrant provided the monopolist offers M and it offers $\{x\}$. This is greater than or equal to:

$$\lambda(|M|) - (c^* - c_*)$$

which is strictly positive by the assumption that $\lambda_1 > c^* - c_*$. Assume now $\{b(X)\} \in S(\sigma_{\mathbf{c}})$. The change in payoff resulting from the same deviation is at least:

$$\frac{1}{2} \sum_{\{x\} \in S(\sigma_{\mathbf{c}})} \sigma_{\mathbf{c}}(\{x\})(1 - \Lambda(|M \cup \{x\}|)) - (c^* - c_*)$$

which is greater than or equal to:

$$\frac{1}{2}(1 - \Lambda(n-1)) - (c^* - c_*)$$

and this is strictly positive because $\lambda_n > 2(c^* - c_*)$. Thus, $b(X) \in M$ for any $M \in S(\sigma_{\mathbf{d}})$. Since this holds in any equilibrium, $\mathbf{e}$ gets a positive payoff iff it offers $\{b(X)\}$, which means $S(\sigma_{\mathbf{c}}) = \{\{b(X)\}\}$. But the unique best response to this by $\mathbf{m}$ is offering $(L_{m^*-1} \cup \{b(X)\}, \{b(X)\})$. To see, note that this menu maximizes the payoff of the monopolist among all such menus. By the assumption of $\lambda_k \leq 2c_k$ for all $n-1 \geq k \geq m^*$, any M of the form $L_{k-1} \cup \{b(X)\}$ is not offered, because this assumption implies that $\Lambda(k) - \Lambda(m^*) \leq 2c(\{x_{m^*+1}, \ldots, x_k\})$ and hence a deviation to $L_{m^*-1} \cup \{b(X)\}$ is profitable. Note that there is always such an m^*, that is $m^* \geq 1$, since $\lambda_1 > c^* - c_*$ implies that the monopolist gets a strictly positive payoff by only offering $\{b(X)\}$, equal to $\frac{1}{2}(1 + \lambda_1) - c^* > \frac{1}{2}(1 - (c^* + c_*)) + \frac{1}{2}c^* > 0$. Also, from any smaller menu, the monopolist can profitably deviate by adding all products up to x_{m^*-1}. This shows that $L_{m^*-1} \cup \{b(X)\}$ is the unique best response, and hence $(L_{m^*-1} \cup \{b(X)\}, \{b(X)\})$ is the unique equilibrium.

References

[1] Abaluck, J. and Abi Adams-Prassl (2021). What do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses. The Quarterly Journal of Economics, Volume 136, Issue 3, August 2021, Pages 1611–1663

[2] Aguiar, Victor and Boccardi, Maria Jose and Kashaev, Nail and Kim, Jeongbin. Random Utility and Limited Consideration (June 21, 2021). Available at SSRN: https://ssrn.com/abstract=3871256 or http://dx.doi.org/10.2139/ssrn.3871256

[3] Armstrong, M., and Chen, Y. (2009). Inattentive Consumers and Product Quality. Journal of the European Economic Association, 7(2/3), 411–422.

[4] Armstrong, M. and Vickers, J. (2020). Patterns of competitive interaction. Econometrica 90 (1): 153-191

[5] Basov, S. and Danilkina, S. (2015). Bertrand Oligopoly with Boundedly Rational Consumers. The B.E. Journal of Theoretical Economics, 15(1), 107-123.

[6] Brady, R. L., and Rehbeck, J. (2016). Menu-Dependent Stochastic Feasibility. Econometrica, 84(3), 1203–1223.

[7] Cattaneo, M.D., Ma, X., Masatlioglu, Y., and Suleymanov, E. (2020) A random attention model. Journal of Political Economy, 128, 2796–836.

[8] Carsten Dahremöller, and Markus Fels (2015). Product lines, product design, and limited attention, Journal of Economic Behavior Organization, Volume 119, p. 437-456

[9] Dardanoni, V., P. Manzini, M. Mariotti, and C. J. Tyson. 2020. Inferring Cognitive Heterogeneity from Aggregate Choices. Econometrica 88(3): 1269–96.

[10] DellaVigna, Stefano and Matthew Gentzkow. (2019) Uniform Pricing in U.S. Retail Chains. The Quarterly Journal of Economics. Volume 134, Issue 4, p. 2011–2084

[11] Demirkan, Yusufcan and Mert Kimya, Hazard rate, stochastic choice and consideration sets. Journal of Mathematical Economics, Volume 87, 2020, Pages 142-150,

[12] de Clippel, G., Eliaz, K., and Rozen, K. (2014). Competing for Consumer Inattention. Journal of Political Economy 122(6), 1203–1234. https://doi.org/10.1086/677253

[13] Anderson, S. P. and de Palma, A. (1992). The Logit as a Model of Product Differentiation. Oxford Economic Papers, 44(1), 51–67.

[14] Eeckhoudt, L., Schlesinger, H., and Tsetlin, I. (2010). Simplifying and solving decision problems by stochastic dominance relations

[15] Eliaz, Kfir, and Ran Spiegler. 2011a. Consideration sets and competitive marketing. The Review of Economic Studies. 78(1): 235–262

[16] Eliaz, Kfir, and Ran Spiegler. 2011b. On the Strategic Use of Attention Grabbers. Theoretical Economics. 6: 127-155

[17] Ellison, Glenn. Bounded rationality in industrial organization. In Richard Blundell, Whitney Newey, and Torsten Persson, editors, Advances in Economics and Econometrics: Ninth World Congress (Econometric Society Monographs), chapter 5, pages 142–174. Cambridge University Press, Cambridge, 2006

[18] Fishburn, P. C. (1980). Stochastic Dominance and Moments of Distributions. Mathematics of Operations Research, 5(1), 94–100. http://www.jstor.org/stable/3689397

[19] Hauser, John R. and Wernerfelt, B. (1990), An Evaluation Cost Model of Consideration Sets, Journal of Consumer Research, Volume 16, Issue 4, p. 393–408

[20] Gallego, Guillermo and Li, Anran. Attention, Consideration then Selection Choice Model (September 24, 2017). Available at SSRN: https://ssrn.com/abstract=2926942 or http://dx.doi.org/10.2139/ssrn.2926942

[21] Goeree, M. S. (2008). Limited Information and Advertising in the U.S. Personal Computer Industry. Econometrica, 76(5), 1017–1074.

[22] Hefti, Andreas (2018) Limited Attention, Competition and Welfare. Journal of Economic Theory, Vol. 178, pp. 318–359.

[23] Heidhues, Paul and Botond Koszegi. (2018). "Behavioral Industrial Oganization" in Handbook of Behavioral Economics, Vol. 1, ed. D Bernheim, S DellaVigna, and D Laibson, Elsevier, 517-612.

[24] Honka, E., A. Hortacsu, and M. Wildenbest (2019): "Empirical Search and Consideration Sets," in Handbook of the Economics of Marketing, ed. by J.-P. Dube and P. Rossi. Amsterdam: Elsevier, 193–257.

[25] Hortaçsu, A., Madanizadeh, S. A., Puller, S. L. (2017). Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market. American Economic Journal: Economic Policy, 9(4), 192–226.

[26] Iyengar, Sheena and Mark Lepper. 2000. *When Choice is Demotivating: Can One Desire Too Much of a Good Thing?* Journal of Personality and Social Psychology, 79(6): 995-1006.

[27] Manzini, P., and Mariotti, M. (2018). Competing for attention: Is the showiest also the best?. The Economic Journal, 128(609), 827-844.

[28] Manzini, P. and Mariotti, M. (2014). Stochastic Choice and Consideration Sets. Econometrica, 82: 1153-1176.

[29] Martinovici, Ana and Pieters, F. G. M. (Rik) and Erdem, Tulin, Attention Trajectories Predict Brand Choice (February 14, 2021). Available at SSRN: https://ssrn.com/abstract=3393021 or http://dx.doi.org/10.2139/ssrn.3393021

[30] Matejka, F. (2015). Rigid pricing and rationally inattentive consumer Journal of Economic Theory, 158, 656–678.

[31] Matějka, Filip, and Alisdair McKay. 2012. Simple Market Equilibria with Rationally Inattentive Consumers. American Economic Review, 102 (3): 24-29.

[32] McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. P. Zarembka, ed. Frontiers in Econometrics. Academic Press, New York, 105–142

[33] Reutskaja, E., Nagel, R., Camerer, C. F., Rangel, A. (2011). Search Dynamics in Consumer Choice under Time Pressure: An Eye- Tracking Study. The American Economic Review, 101(2), 900–926.

[34] Shocker, A. D., Ben-Akiva, M., Boccara, B., Nedungadi, P. (1991). Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models, and Suggestions. Marketing Letters, 2(3), 181–197.

[35] Simon, H. A. (1956). Rational choice and the structure of the environment. Psychological Review, 63(2), 129–138.

[36] Spiegler, Ran. (2011). Bounded Rationality and Industrial Organization. New York: Oxford University Press.

[37] Spiegler, Ran. 2015. Choice complexity and market competition. Annual Review of Economics

[38] Varian, H. R. (1980). A Model of Sales. The American Economic Review, 70(4), 651–659.