Breadth versus Depth

Sen Geng Leonardo Pejsachowicz Michael Richter

Xiamen University, Princeton University, Royal Holloway - London

Preliminary Draft

March 14, 2017

Abstract

We consider a fundamental trade-off in search: when choosing between multiple unknown alternatives, is it better to learn a little about all of them (breadth) or a lot about a single one (depth)? In choice settings where a distribution is exogenous, we find that breadth is optimal for "small" problems and that depth is optimal for "large" ones. On the other hand, in IO settings, where firms endogenously choose distributions, we find breadth to be always optimal. We consider an application to a political setting where voters learn about candidates and find a rational justification for a heretofore unexplained fact, voters tend to learn only about their preferred candidate. Finally, we consider extensions to fat-tails and correlation, and find that in these extensions, breadth is superior.

Keywords: Incomplete Information, Breadth, Depth, Search

1 Introduction

In this paper, we consider the problem of an agent choosing between many objects each of which has different values along several attributes. The agent does not know these values and thus, she must search the alternatives in order to determine what to choose. But, she will always face a problem of incomplete information when choosing in that she will only be able to partially reveal the values of the alternatives that she faces. Specifically, she has two possible search strategies: i) "depth" where she searches one object intensely and learns all of its attribute values or ii) "breadth" where she searches one attribute intensely and learns each object's value in this attribute. After conducting her search, the agent makes her selection. Her choice is straightforward in that she selects the alternative with the highest expected value. But, her search problem on the other hand is highly non-trivial and features a fundamental informational trade-off: "Is it better to learn a lot about a single option or a single fact about all options?"

In several real-life examples, both of the aforementioned search methods are employed in varying contexts. For example, when one searches for an airline flight, the search provider will normally provide a list of flights ordered by price. That is, the searcher learns about a single attribute (price) for all objects. While there are certainly other attributes that are relevant for choice (departure/landing times, airline company, seat location, aircraft model, checked bag fees, changeability), the ranking focuses upon price and an agent perceives the alternatives through this attribute. Search-by-attribute shows up in several other settings as well such as phones (camera quality), restaurants (Yelp rating), college choice (USNews ranking) and stock selection (PE ratio). Turning to the search-by-object method, an investor may be referred to a stock (by either a broker, a news channel, or perhaps the stock of his employer) and investigate that stock in great detail in order to determine whether to make an investment. Investment banks also employ research analysts who focus on a single stock and investigate that stock in depth before making recommendations and price targets. Other settings where an agent may search-by-object are restaurants (food critics/friends' recommendation), colleges, technology products[1], or spouses.

In general, when an agent faces a complicated choice, it is not possible to make a perfectly informed decision, perhaps due to limited attention or time. Instead, the agent can only learn some of the underlying values prior to making her decision. Here, we consider the starkest possible trade off, pitting investigation of a singleattribute search against investigation of a single object. That different methods are used across a variety of settings suggests that empirically there is no clear ranking between the two. We find that theoretically this is the case as well, as the performance of each search strategy will depend upon the number of alternatives, attributes, and the underlying value distribution. But, in spite of this depdendence, in this paper, we will be able to carve out broad classes of problems where we can establish superiority of each search method.

The current research project is connected to a small but growing literature in search among multi-attribute products. **?** considers a yes/no adoption choice for a single good where the agent searches across different attributes with different distributions and costs. Both **?** and **?**, consider search environments with multiple attributes and focus on externalities generated by agents searching different attributes. The first take place in a strategic game setting against a monopolist, and the second is a matching model. Closer to the present work, as it is also focused on the depth/ breadth tradeoff in search, is **?**. Here agents, facing (not ex-ante) differentiated products, and simultaneously decide on search intensity - which determines precision of the signal - and on the number of searched objects. Unlike us they fix a specific distribution for the signal, whose convenient analytic properties lead to closed form (logit) formulae for the market shares of the sellers. This allows them to place the individual decision problem inside a strategic model of online intermediaries. Finally, Ke, Shen, and Villas-Boras (2017) consider a model with a finite number of objects and a

[1] For each product, such as a router, The Wire Cutter, `http://thewirecutter.com/` recommends "the best router" which it describes in depth. The Sweet Home, `http://thesweethome.com/` provides a similar service for home products.

continuum of attributes which evolve only while being searched according to Brownian motions and characterize the optimal switching and purchase decisions.

The rational inattention literature, which gives decision makers agency in the design of their information, has touched upon similar themes. In particular **?** look at optimal portfolio choice with an initial signal selection stage and show how under different classes of preferences and information technologies the agent might prefer to either learn mostly about a single asset, or to distribute his limited attention between many. Finally the phenomenon of saliency, individuals focusing on a single - salient - feature of the alternatives at hand in complex decision problems, has recently been the focus of much research in behavioral economic theory, in particular in **?** and **?**. Both papers posit a context dependent weighting function that the agent applies to the various attributes of a good when choosing from a set. We partly see our work as a rational explanation for this class of behavioral models.

2 The Model

2.1 Framework

A single decision maker faces a choice between N objects, whose utility value is given by the sum of their values along N different attributes. While in general the number of attributes and objects need not coincide, we will generally focus on the case where they do as this is the fairest comparison between the two studied search methods.[2] The decision maker does not know any object's value for any attribute, but they are all drawn i.i.d. from the same symmetric, mean zero distribution F. For an agent to decide between the breader and depth search methods, this mean zero requirement is just a normalization. One may safely consider the decision that the agent faces

[2]If the number of objects differs from the number of attributes, then one search method will reveal more values than the other, giving it an informational advantage.

as being defined by a random matrix with objects denoted as rows and attributes as columns. For this formulation, the entry x_{ij} stands for the realized value of the random variable j-th attribute of the i-th object.

To make a perfectly informed decision, the agent would be best off if she could observe all attribute values for every object. She would then select the object with the maximal sum. But, as argued in the introduction, there are a variety of situations where this may be unrealistic. Instead, we consider two fundamentally different strategies.

According to the first, which we refer to as search of an object, the agent searches a single object and realizes each of its attribute values (that is, a single row in the matrix). If an agent searches by this method, then her decision rule is straightforward. Take the searched object if its payoff is positive and otherwise take an unsearched alternative, which yields expected payoff 0. Thus, her choice method has a flavor of satisficing (**?**). She selects the searched alternative if it has an above-average realization and she slected a different unsearched alternative if the searched alternative has a below-average realization.

The second search method is search of an attribute, where an agent searches a single attribute and realizes that attribute's value for every object. Here too, an agent's decision rule is simple. She takes the object with the highest realized value for that attribute. Due to the i.i.d. and zero mean assumptions, if an agent selects an object with realized value x for one attribute, then her expected payoff is x as well. Thus, there are no concerns of Bayesian updating, what she sees is what she gets. The two search methods are pictorally depicted below.

	A_1	A_2	A_3
O_1	x_{11}	x_{12}	x_{13}
O_2	F	F	F
O_3	F	F	F

	A_1	A_2	A_3
O_1	x_{11}	F	F
O_2	x_{21}	F	F
O_3	x_{31}	F	F

Figure 1: Search of an object (left panel) and search of an attribute (right panel).

As explained earlier, when an agent searches an object, say O_1, she decides to keep it if it has an above-average realization, that is $x_{11} + x_{12} + x_{13} \geq 0$. If she searches an attribute, then she chooses the object which has the highest realization for that attribute and she receives expected value 0 from all unsearched attributes. Therefore, an agent's expected payoffs are:

$$\text{Search of an Object} = \mathbb{E}\left[\max\left(\sum_{i=1}^{N} x_{1i}, 0\right)\right]$$

$$\text{Search of an Attribute} = \mathbb{E}\left[\max\left(x_{11}, \ldots, x_{N1}, 0\right)\right].$$

From the above formulas, one can see that the payoff of searching an object is the sum of N random variables, truncated at 0 while the payoff of searching an attribute is the max of N draws. Thus, the basic trade-off of breadth versus depth can be mathematically translated into the choice between the sum or the max of N random variables.

2.2 Small N

We first consider the case where $N = 2$. As a motivating example, suppose that F is a simple binomial distribution that takes each of the values 1, -1 with probability $1/2$. For this distribution, there are four possible row realizations and the corresponding object/attribute search payoffs are denoted in the following table.

Example: Binomial Distribution

Row	Depth Payoff	Column	Breadth Payoff
$[1,\ 1]$	2	$\begin{bmatrix}1\\1\end{bmatrix}$	1
$[1,\ -1]$	0	$\begin{bmatrix}1\\-1\end{bmatrix}$	1
$[-1,\ 1]$	0	$\begin{bmatrix}-1\\1\end{bmatrix}$	1
$[-1,\ -1]$	0	$\begin{bmatrix}-1\\-1\end{bmatrix}$	-1

In the above table, one can notice that there is not a realization-by-realization ranking between the breadth and depth search methods. If both the realizations come up positively, then it is better that they are concentrated on a single object because then the agent receives both of these positive realization. On the other hand, if the realizations are mixed, then the breadth method has a higher payoff because the positive realization can be obtained and the negative realization avoided. Nevertheless, perhaps suprisingly, in the simple case above, the payoff of search of an object is (1/4)*2=1/2 and the payoff of search of an attribute is (3/4)*1 + (1/4) * -1= 1/2. Thus, the two search methods have the same benefit in the case of a coin flip.

We now turn a further examples, and then to our first proposition.

Uniform Distribution: For the uniform distribution on $[-1, 1]$, $f(x) = 1/2$ for $-1 \leq x \leq 1$ and $F(x) = (x + 1)/2$. Let G denote the distribution of the maximum of two draws from F. To calculate the value of searching an attribute, notice that the $G(c) = Prob(max(x_1, x_2) \leq c) = Prob(x_1 \leq c)^2 = F(c)^2 = (c + 1)^2/4$ and so $g(x) = (x + 1)/2$ and

$$\text{Breadth Payoff} = \int_{-1}^{1} xg(x)dx = 1/3.$$

On the other hand, the distribution of the sum, H can be calculated as $H(c) = Pr(x_1 + x_2 \leq c) = \int_{-1}^{1} \int_{-1}^{\min(1, c-x_1)} f(x_1)f(x_2)dx_2dx_1$. In the case where $c \geq 0$, this is $H(c) = (-c^2 + 4c + 4)/8$ and $h(c) = (-c + 2)/4$. Taking into account that the sum of two uniforms ranges from -2 to 2 and that the object is only chosen when it has a non-negative realization, the benefit of searching an object is:

$$\text{Depth Payoff} = \int_0^2 xh(x)dx = 1/3.$$

Notice that while the expected benefit of the breadth and depth search methods for the uniform distribution are the same, the distributions are not, as depicted below.

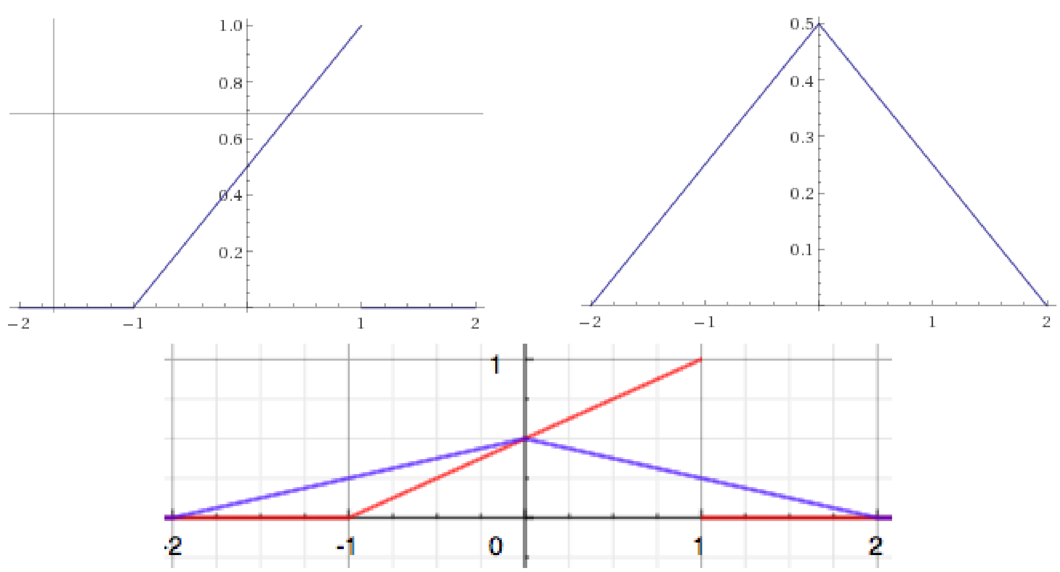

Figure 2: Densities of the max of two uniforms (left panel) and the sum of two uniforms (right panel).

Thus, we have established that for the Bernoulli and Uniform distributions, while the distributions of the maxes and sums are different, they have the same expectation and therefore the breadth and depth search methods have the same payoffs for these distributions. The following proposition shows that these examples are instances of a more general result which holds when $N = 2$ and F is symmetric.

Proposition 2.1 *When $N = 2$ and F is symmetric, breadth=depth. That is, the payoffs of searching an object or searching an attribute are the same.*

Proof. We proceed with a coupling argument. When an agent searches, she realizes two values,x_1, x_2. Because the distribution is symmetric, each of the four possible vectors $(\pm x_1, \pm x_2)$ are realized with equal probability, specifically $f(x_1) \cdot f(x_2)$. Without loss of generality, assume that $x_1 \geq x_2 \geq 0$.

Search of an object: The agent will keep the object whenever the vector (x_1, x_2) or $(x_1, -x_2)$ is realized and therefore the payoff of searching an object is $1/4(x_1 + x_2) + 1/4(x_1 - x_2) = 1/2 \cdot x_1$.

Search of an attribute: For the vectors (x_1, x_2) or $(x_1, -x_2)$, the agent will select the object with the positive x_1 realization and receive a payoff of x_1. For the vector $(-x_1, x_2)$, she willl select the object with the x_2 realization and receive a payoff of x_2. For the vector $(-x_1, -x_2)$, she will select the object with the $-x_2$ realization, because while both realizations are negative, this is the less negative realization. Therefore, her payoff of searching an attribute is $1/2 \cdot x_1 + 1/4 \cdot x_2 + 1/4 \cdot (-x_2) = 1/2 \cdot x_1$.

Thus, her payoff of either search method is the same conditional on one of the vectors $(\pm x_1, \pm x_2)$ being realized. As looking at these groupings provides a partition of all possible realizations, and the search methods offer the same payoff for each cell of this partition, it is the case that both search methods offer the same expected payoff.

■

The above proof proceeds by partitioning the space of all possible realizations based upon their absolute values. We then demonstrate that in each cell of this partition, the two search methods offer the same average payoff (although they generally do not offer the same payoff for any two vectors within the cell).

In each cell, there is a vector where all realizations are negative. For the agent who searches an object, she can avoid these realizations by choosing the other unsearched object. But, the agent who searches an attribute must choose one of the two objects. In such a situation, she chooses the object with the less negative realization but necessarily obtains negative utility. If there was an outside option so that she did not have to choose any object when they both have negative realizations, then the payoff of searching an attribute would strictly increase. On the other hand, adding an outside option to the case where she inspects an object does not increase her payoff because there is already an uninspected object which serves as an outside option. Therefore, the presence of an outside option breaks the tie between the search methods and leads to the following corollary.

Corollary 2.1 *If $N = 2$, F is symmetric, and the agent has an outside option, then searching an attribute is strictly better than searching an object.*

We now turn our attention to other N. As it turns out, the technique of partitioning the space of realizations and analyzing the benefits of each search method for each cell can still be applied for some N. Specifically, for N in the range $[3, 6]$, it can be discovered that searching an attribute generally leads to a strictly higher payoff than searching an object for each cell in the partition.

Theorem 2.1 *If $3 \leq N \leq 6$, then it is weakly better to search an attribute than to search an object. Moreover, except for the case where F is Bernoulli and $N = 3$ or 5, it is strictly better to search an attribute than to search an object.*

Proof. Let $a_1 \geq \ldots \geq a_N \geq 0$. Let the event $A_N = \{|X_i| = a_i\}$. Notice that both search methods reveal N random variables and each row is equally likely as a column. Therefore, it suffices to show that

$$2^N \mathbb{E}[\max(X_1, \ldots, X_N, 0)|A_N] \geq 2^N \mathbb{E}\left[\max(\sum_{i=1}^{N} X_i, 0)|A_N\right].$$

Given that the underlying distribution is symmetric, the left-hand side is straightforward to calculate, it is $\sum_{i=1}^{N-1} a_i 2^{N-i}$. The reason is that conditional on A_N, there is a $1/2$ probability that $X_1 = a_1$ contributing a_1 to the max payoff. Furthermore, there is a $1/4$ probability that $X_1 = -a_1$ and $X_2 = a_2$ contributing a_2 to the max payoff. In general, there is a $1/2^i$ probability that $X_j = -a_j$ for all $j < i$ and $X_i = a_i$. The above formula features 2^{N-i} rather than probabilities because the expectation was multiplied by 2^N. The single subtelty is that there is a $1/2^N$ chance that $X_j = -a_j, \ \forall j$ and in this case, the agent would choose object N because it has the least negative realization, contributing $-a_N$ to the payoff. This exactly cancels out with the case where $X_j = -a_j \ \forall j < N, \ X_N = a_N$ and that is why the sum runs from $i = 1$ until $N - 1$.

The right-hand side of the expectation is significantly more complicated to calculate. But, notice that the right hand side can also be expressed as $\sum_{i=1}^{N} a_i b_i$. I will show that the left-hand expectation is greater than the right-hand side expectation by demonstrating a first-order stochastic dominance, that is, $\sum_{i=1}^{j} 2^{N-i} \geq \sum_{i=1}^{j} b_i$. Without loss of generality, one may restrict attention to the case that $j = N$ because smaller j are dealt with according to smaller N. The previous proposition demonstrates the base case when $N = 2$.

The reason that the b_i are difficult to calculate is that the term $\max\left(\sum_{i=1}^{N} X_i, 0\right)$ depends upon $a_1, \ldots, a_N$. Nevertheless, to calculate an upper bound of $\sum_{i=1}^{N} b_i$ is straightforward because it requires counting vectors where the majority of coefficients are positive. More specifically, there is one vector with N positive terms, N vectors with $N - 1$ positive terms and one negative term and generally $\binom{N}{k}$ vectors with $N - k$ positive terms and k negative terms. Only terms where

$N - k > k$ need to be considered and the relevant calculations are conducted below.

- $N = 3$: $2^N - 2 = 6 \geq 1 \cdot 3 + 3 \cdot 1 = 6$
- $N = 4$: $2^N - 2 = 14 > 1 \cdot 4 + 4 \cdot 2 = 12$
- $N = 5$: $2^N - 2 = 30 \geq 1 \cdot 5 + 5 \cdot 3 + \binom{5}{2} \cdot 1 = 30$
- $N = 6$: $2^N - 2 = 62 > 1 \cdot 6 + 6 \cdot 4 + \binom{6}{2} \cdot 2 = 60$
- $N = 7$: $2^N - 2 = 126 \not\geq 1 \cdot 7 + 7 \cdot 5 + \binom{7}{2} \cdot 3 + \binom{7}{3} \cdot 1 = 140$

According to the above, we have strict inequalities when $N = 4, 6$ and equalities when $N = 3, 5$. The importance of the equality cases is that the coefficient sums then only give a weak dominance as weight is being shifted from the lower terms a_3 to the higher terms a_1. So, it is still the case that breadth offers a higher payoff in this scenario, unless $a_1 = a_3$ in every cell. That is, for the Bernoulli distribution and $N = 3,\ 5$, the two search methods offer the same payoff and otherwise, inspecting an attribute offers a strictly higher payoff.

■

The above theorem states that searching an attribute generally outperforms searching an object for all symetric distributions when N is small. Moreover, the Bernoulli case demonstrates that there can be parity issues and that generally, the superiority of one search method to another may be non-monotonic.

For larger N, which search method is better depends upon the underlying distribution F. In fact, it can be shown that for every $N \geq 7$, there are two symmetric distributions F_N and G_N so that for F_N it is strictly better to search an attribute and for G_N it is strictly better to search an object. In fact, the Bernoulli distribution is an example of G_N because for that distribution, searching an attribute is better when $N \leq 6$ and searching an object is better when $N \geq 7$.

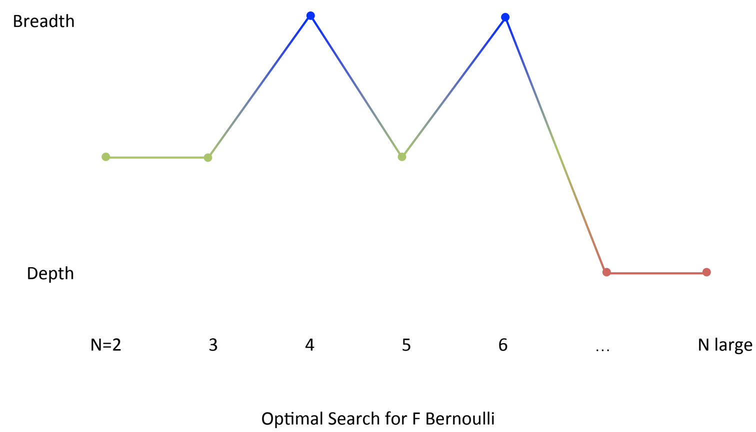

Figure 3: Breadth vs. Depth for the Bernoulli Distribution

2.3 Large N

In the previous section, we have seen that even when problems are relatively simple, there are both situations in which attribute search might prevail and ones in which object search takes over. Here we turn to asymptotics in order to understand whether for N large enough -for a sufficently complex problem - we might obtain a more clear-cut answer. It turns out that even here that is not the case as we will provide, first for object and then for attribute search, a large class of distributions under which each is optimal in the limit.

As one might expect, the key tools in the study of the value of $V_X^N(obj)$ ($V_X^N(att)$) as N grows are the limit theorems for sums (maxima) of i.i.d. random variables, which identify conditions under which such sums (maxima) converge to a given non-degenerate distribution. Because many distributions can have the same limit, here are able to provide some results for non symmetric distributions of X, although we will always require the limit to be symmetric.

Throughout this section, we will assume that we are given a sequence $\{X_n\}_{n=1}^{\infty}$ of i.i.d. random variables distributed as X with $\mathbb{E}(X) = 0$.

2.3.1 Finite variance

Here we turn our attention to the case in which $\mathbb{E}(X^2) = \sigma^2 < \infty$. In this the limit value of object search is easily pinned down. By the central limit theorem,

$$\frac{V_X^N(obj)}{\sqrt{N}} = \frac{\mathbb{E}\left(\max\left\{\sum_{n\leq N} X_n, 0\right\}\right)}{\sqrt{N}} \approx \sigma\sqrt{\frac{1}{2\pi}} \tag{1}$$

for large enough n, where we used the inverse Mills ratio formula to determine the expectation of the normal above the mean. Hence all we need to do is to study $V_X^N(att)$ as $N \to \infty$. In the theorem below we are able to show, thanks to results by Gumbel (1954) establishing bounds for the expected maxima of an i.i.d. sequence, that it will eventually lay below the value of object search:

Theorem 2.2 *If X has variance $\sigma^2 < \infty$ then $V_X^N(obj) > V_X^N(att)$ for N large enough.*

Proof. By Gumbel (1954), for any r.v. X with mean μ and finite variance σ^2 it must be that

$$\mathbb{E}(\max_{n\leq N}\{X_n\}) \leq \mu + \sigma\frac{N-1}{\sqrt{2N-1}}$$

Now suppose we truncate the variable X at some non-negative value c, let this r.v. be called $X^{|c}$. If we calculate the Gumbel bound for this r.v., we get

$$\mathbb{E}(\max\{X_1^{|c}, X_2^{|c}, ..., X_N^{|c}, 0\}) = \mathbb{E}(\max_{n\leq N}\{X_N^{|c}\}) \leq$$

$$\leq \mu_c + \left[(c-\mu_c)^2 F(c) + \int_c^\infty (x-\mu_c)^2 \, dF(x)\right]\frac{N-1}{\sqrt{2N-1}}$$

where $\mu_c = cF(c) + \displaystyle\int_c^\infty x \; dF(x)$. Because the first and second moments exists, the terms $\displaystyle\int_c^\infty x \; dF(x)$ and $\displaystyle\int_c^\infty x^2 \; dF(x)$ (and thus also $\displaystyle\int_c^\infty (x-\mu_c)^2 \; dF(x)$) converge to zero as c

grows, and $cF(x) + \int_c^\infty x \; dF(x)$ is asymptotically c. So we can take c large enough to make the coefficient of $\frac{N-1}{\sqrt{2N-1}}$ as small as we want. It follows that for any ϵ and δ small we can find a c such that

$$\mathbb{E}(\max_{i \leq n}\{X_i^{|c}\}) \leq (c+\epsilon) + \delta \frac{N-1}{\sqrt{2N-1}}$$

Since $X^{|c}$ dominates X, it must be that for such c (and for any N)

$$V_X^N(att) \leq c + \delta \frac{N-1}{\sqrt{2N-1}} + \epsilon \leq c + \delta \sqrt{\frac{N}{2}} + \epsilon$$

and thus

$$\frac{V_X^N(att)}{\sqrt{N}} \leq \frac{c}{\sqrt{N}} + \frac{\delta}{\sqrt{2}} + \frac{\epsilon}{\sqrt{N}}$$

While

$$\frac{V_X^N(obj)}{\sqrt{N}} = \sqrt{\frac{1}{2\pi}} + o(1)$$

So for n large enough $V_X^N(obj) > V_X^N(att)$. Q.E.D.

The intuition here is the following: as long as the variance of X is finite, the expected contribution of any attribute j to the value of an object is small w.r.t. the value of the sum $\sum_n x_n$. Hence even if observing an attribute from a growing number of objects increases the likelihood of finding a very high realization, this is not enough to offsett the low expected informativeness about the value of each alternative, and thus the value of attribute search eventually goes to zero. On the other hand searching the object always gives full information about the alternative observed, hence the value of object search converges to the expectation of $\max\{X, 0\}$ for X normally distributed, which is always positive.

2.4 Infinite Variance

The above arguments regarding the case where N is large dealt purely with distributions which were "thin-tailed", that is, they had finite variance. However, in many applications, for example, stock prices, the case of unbounded variance may be of interest. As it turns out, the tail of a distribution controls the relative optimality of the two search methods when N is large.

Theorem 2.3 *Let X be a zero mean r.v. with distribution F satisfying $1 - F(x) \to \frac{c}{x^{-\alpha}}$ and $\frac{1-F(x)}{F(-x)} \to 1$ as x goes to infinity, for some positive constant c and some $1 < \alpha < 2$. Then there is an $1 < \alpha^* < 2$ such that for all $\alpha < \alpha^*$ search-by-attribute beats search by object, and for all $\alpha > \alpha^*$ the opposite is true.*

Proof. By Mikosch (1999), X satisfies conditions such that

i) the sum $\sum_{n \leq N} X_n$ converges to an α-stable distribution $S_\alpha(1,0,0)$ at rate $\left(\frac{Nc}{C_\alpha}\right)^{\frac{1}{\alpha}}$, where $C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x dx\right)^{-1}$.

ii) the maxima $\max_{n \leq N} X_n$ converges to the standard Frechet distribution $F_\alpha(x) = e^{-x^{-\alpha}}$ at a rate $(Nc)^{\frac{1}{\alpha}}$.

Hence under these conditions both the sum and the maxima have nondegenerate limits, and we can thus calculate the integral above zero for each of the two and confront them. For the max things are easy :since the standard Frechet has non-negative support the value coincides with the mean, which is $\Gamma(1 - \frac{1}{\alpha})$, where Γ is the classical Gamma function. For the sum things are trickier, as most α-stable laws do not have a closed form. Nevertheless for a symmetric stable law we can use the calculations in **?** which provide the expectation of the absolute value, and divide by two.

Doing so, we obtain the value $\Gamma(1-\frac{1}{\alpha})\frac{1}{\pi}$. Hence we have that as N grows

$$\frac{V_N^X(att)}{V_N^X(obj)} \to C_\alpha^{\frac{1}{\alpha}} \pi \frac{(Nc)^{\frac{1}{\alpha}}\Gamma(1-\frac{1}{\alpha})}{(Nc)^{\frac{1}{\alpha}}\Gamma(1-\frac{1}{\alpha})} = C_\alpha^{\frac{1}{\alpha}}\pi.$$

If the last term is bigger than 1, eventually attribute search is more valuable than object search, and vice versa if it is lower than one. Since $\int_0^\infty x^{-\alpha}\sin x dx$ is $\frac{\pi}{2}$ for $\alpha = 1$, for every $\alpha > 1$ but close enough, $C_\alpha^{\frac{1}{\alpha}}\pi \approx 2$,hence for low values of α attribute search prevails. On the other side $C_\alpha \to 0$ as α goes toward 2, hence eventually object search is more valuable. ■

The above theorem demonstrates that the relative payoffs of the two search methods, in the limit, depends upon the fatness of the tails of the underlying distribution. Specifically, for very fat-tailed distributions, not only is the agent choosing in a riskier world, but there is a second heretofore unmentioned endogenous form of risk, the agent searches in a riskier fashion.

3 Correlation through Hidden States

We now generalize the model to allow for correlation. As before, there are N objects each with N attributes. Each object i has a hidden value v_i which is i.i.d. drawn from a distribution $N(0,\sigma)$. The payoff for an agent of choosing an object is it's hidden valuation. Object i's j^{th} attribute is denoted $a_{i,j}$. Attributes are now signals in the sense that they help the agent to infer the hidden value which is the source of utility rather than being the source of utility themselves. Formally, there is a probability p so that $a_{i,j} = v_i$ with probability p and $a_{i,j}$ is an i.i.d. draw from $N(0,\sigma)$ with probability $1-p$.

Theorem 3.1 *With hidden value payoffs and unbounded support, as $N \to \infty$, attribute search is preferable to object search.*

Proof: As before, we must calculate approximations for the value of attribute and object search. Notice that when one searches an object, with probability $Np(1-p)^{N-1} + (1-p)^N$ the set of attributes will contain the hidden value at most once. With the remaining probability, denoted P_N, the attributes will contain the hidden value at least twice. As F is a continuous distribution, the hidden value will then be deducible in these situations. The agent will then select the object if it has a positive deduced value and reject it if does not. This yields a payoff of $P_N \mathbb{E}[X|X > 0] = \frac{P_N \sigma}{\sqrt{2\pi}}$ where $X \sim N(0, \sigma)$. With probability $1 - P_N$, the agent cannot perfect identify the value of the object, but an upper bound for the agent's payoff is to assume that she still can. In this, her payoff in that scenario is $\frac{(1 - P_N)\sigma}{\sqrt{2\pi}}$ and therefore an upper bound to the agent's value from object search is $\frac{\sigma}{\sqrt{2\pi}}$. This is a natural bound as it represents her payoff under perfect information about one object.

If the agent searches by attribute, we can assume without loss of generality that he searches attribute 1. The agent will then select the object j with the highest attribute value, a_{1j}. In this case, with probability p he receives a_{1j} and with probability $1 - p$ he receives an iid draw from $N(0, \sigma)$ which yields 0 in expectation. Therefore, his expected payoff is $p\mathbb{E}[\max(X_1, \ldots, X_N, 0)]$ where $X_1, \ldots, X_N \sim N(0, \sigma)$. Clearly $\mathbb{E}[\max(X_1, \ldots, X_N, 0)] \to \infty$. ■

The above proof can also be used for more general distributions F with mean 0. In this case, we see that attribute search dominates object search if $\mathbb{E}[X|X > 0] < p\mathbb{E}[\max(X_1, \ldots, X_N, 0)]$. This is true for any distribution with infinite support as the RHS will go to infinity and in fact is true for many distributions with finite support. However, when $p < 1/2$, there is one prominent case that goes the other way, the standard heads/tails distribution where $P(X = 1) = P(X = -1) = 1/2$.

Theorem 3.2 *With sum payoffs and unbounded support, as $N \to \infty$, attribute search is preferable to object search.*

Proof. If the hidden value of an object is v, then on average, p realizations of the attribute space will be v. Moreover, with extremely high probability less than or equal to $2p$ realizations are v. [Need to bound the remaining probability which should be fine.] If the rest of the observations are drawn from F, then an upper bound for the payoff from object search is $2pN \max(v, 0)$ + $c_2\sqrt{N}$. Taking expectations over v then yields $2pNC + c_2\sqrt{N} = c_2N + c_2\sqrt{N}$.

On the other hand, when searching by attribute, let m denote the maximum value realized. There is a probability p that the realized value is the hidden value. In this case, the agent can expect a payoff of $m + (N-1)m$ from the remaining hidden values being realized. With the remaining probability, the agent gets a payoff of m. Therefore, the agent's payoff is $m + p^2(N-1)m$. Taking expectations as before, $\max \mathbb{E}[m] \to \infty$ and therefore attribute search is greater than object search.

■

4 IO

In this section, we consider the case of N firms and a single consumer. The number of consumers is not an essential feature of the following model. Each firm produces one good composed of N attributes. Each firm i chooses a distribution F_i over $\mathbb{R}_+$ with $\mathbb{E}[F_i] = 1$. Notice that we are restricting attention to strategies where a firm's distributional choice F_i is used for every attribute. A firm's goal is to maximize market share.

4.1 Attribute Search Equilibria

In this subsection, we characterize all symmetric equilibria where agents search by attribute. Recall that the firm's objective is to maximize the probability of purchase, that is

$$\max_{F_i} \int_x \Pi_{j \neq i} F_j(x) dF_i(x) \quad \text{s.t.} \quad \mathbb{E}[F_i] = 1$$

Out first lemma demonstrates that no symmetric equilibrium may have a mass point.

Lemma 4.1 *In any symmetric equilibrium $(F,\dots,F)$, the distribution F cannot have any mass points.*

Proof. If F has mass p at $x > 0$, then any firm can profitably deviate by shifting most of this mass to $x+\epsilon$ and the rest to 0. This increases his probability of winning by at least $p^N(N-1)/N$ (probability that every other firm draws this mass point times probability that i was not selected in that event) and increases his probability of losing by at most ϵ (probability that x was previously drawn and now 0 is). Such a deviation is beneficial for any small enough ϵ. A similar argument show that there can be no mass points at 0 because a firm can ϵ increase the value of this mass point and zero out a smaller than ϵ share of his distribution (so as to preserve the condition $\mathbb{E}[F_i] = 1$). ■

Our next lemma demonstrates that in any symmetric equilibrium, the distribution F must be gapless.

Lemma 4.2 *In any symmetric equilibrium $(F,\dots,F)$, the distribution F can not have any gaps.*

Proof. Suppose that there is a gap from x to y. That is, $F(x) = F(y) < 1$ and $z(\epsilon) := F(y+\epsilon) - F(y) > 0$ for all $\epsilon > 0$. Then, a firm can strictly increase its chances of being chosen ■

Next, notice there can be no gaps in the distributions that the firms employ. If there were, then again, a firm could decrease all weight placed a little bit above the gap and increase the values of his distribution elsewhere as a profitable deviation.

Thus, the firm's problem can be approached through the calculus of variations with $\mathcal{L}(x, F_i, f_i) = \Pi_{j\neq i}F_j(x)f_i(x) - \lambda F_i(x)$. Assuming that all other agents use the same symmetric strategy G, the Lagrangian becomes $\mathcal{L} = G^{N-1}(x)f(x) + \lambda F(x)$.The Euler-Lagrange first order condition is $\frac{\partial \mathcal{L}}{\partial F} - \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial f} = 0$. That is,

$$-\lambda - (N-1)G^{N-2}(x)g(x) = 0$$

Rearranging gives $(N - 1)G^{N-2}(x)g(x) = \lambda$. Integrating both sides $G^{N-1}(x) = C + \lambda x$. Therefore, $G(x) = (C + \lambda x)^{1/(N-1)}$. Now, we make some notes on boundary conditions. First, $G(0) = 0$. Notice that if $G(a) = 0$ for some $a > 0$, then the agent would prefer to take all weight on low-values and move it to 0. Therefore, $0 = C^{1/(N-1)} \Rightarrow C = 0$. Second, as discussed above, there are no atoms. Furthermore, there is some common upper bound b. Thus, $1 = G(b) = (\lambda b)^{1/(N-1)}$. Thus, $\lambda = 1/b$ and $G(x) = (x/b)^{1/(N-1)}$. Finally, the condition that $\int_x xg(x)dx = 1$ can be used to determine the upper bound b which is in fact $b = N$. Thus, the unique symmetric Nash Equilibrium among the firms is $F(x) = (x/N)^{1/(N-1)}$ on $[0, N]$.

In the case where $N = 2$, it should be noted that $F(x) = x/2$, that is F is a uniform distribution on $[0, 2]$ and the payoff from attribute and object search is the same for such a distribution. Therefore Attribute search and F is an equilibrium. For $N > 2$, instead of receiving 1 (the average value of an attribute), when the agent searches by attribute, she instead receives $\int_0^N x\left(\frac{d}{dx}F^N(x)\right)dx = xF^N(x)\big|_0^N - \int_0^N F^N(x)dx = N - \int_0^N (x/N)^{N/(N-1)}dx = N - \int_0^1 z^{N/(N-1)}Ndz = N(1-(N-1)/(2N-1)) = N * \frac{N}{2N-1} \approx N/2$. Notice that if the firms chose the ideal (from the agent's point of view) distribution on $[0, N]$, it would be a binomial distribution which takes the value 0 with probability $(N-1)/N$ and takes the value N with probability $1/N$. Searching such an attribute would yield the agent a payoff of $N*(1-((N-1)/N)^N) \to N*(e-1)/e$. Therefore, the equilibrium payoffs of the agents are $\frac{1/2}{(e-1)/e} \approx 0.791$, that is about 79.1% of the social optimum. Raw calculations for $N = 2, 3, 4$ and bounding for $N \geq 5$ demonstrate that the attribute search is an equilibrium.

Agent's payoffs calculations are:

N	Object Payoff	Attribute Payoff
2	$\int_0^2 \int_0^{2-x} (2-y-x) f(x) f(y) dy dx = 1/3$	$N * \frac{N}{2N-1} - 1 = 4/3 - 1 = 1/3$
3	$\int_0^3 \int_0^{3-x} \int_0^{3-y-x} (3-z-y-x) f(z) dz f(y) dy f(x) dx = \frac{\pi}{5} \approx 0.63$	$N * \frac{N}{2N-1} - 1 = 9/5 - 1 = 0.8$
4	$\frac{16\sqrt{3}\pi\Gamma(1/3)^2}{7 * 3^4 * \Gamma(8/3)} \approx 0.73$	$9/7 \approx 1.29$
N	An upper bound is $\left(\frac{N-1}{N}\right)^N N \to \frac{N}{e}$	$N * \frac{N}{2N-1} - 1 \to \frac{N}{2}$

4.2 Object Search - observable F

There is a single consumer who searches by object. The firm's objective, as before, is to maximize the probability of purchase. An agent observes $F_1, \ldots, F_K$ and then chooses which firm to search. As the agent wishes to maximize his expected payoff, it must be the case that firms choose F_i in order to maximize agents' payoffs. But, there is a continuity problem.

Suppose that a firm i chooses F_i which is equal to z with probability $1/z$ and 0 otherwise where z is large. Then the agent's marginal payoff from searching i is $\left(\frac{1}{z^2}\right)(2z-2) + \left(\frac{2}{z} - \frac{1}{z^2}\right)(z-2)$ $= \frac{2(z-1)}{z^2} + \frac{(2z-1)(z-2)}{z^2} = \frac{2z^2-3z}{z^2} = \frac{2z-3}{z}$.

Notice that as z grows large, the above heads to 2 which is an upper bound on an agent's payoff. To see this, suppose that the agent first commits to choose i. This offers him an average payoff of 2. But, he will then switch away from that alternative if it's realized value is less than 2. Therefore, his marginal gain from searching is the value that he gets from switching from bad realizations and in particular is capped by $2 \cdot \Pr(Switching) < 2$. Therefore, the maximal payoff from searching an object is 2. But, there is no distribution which achieves such a payoff. Thus, an object search equilibriums fail to exist in this situation.

One way to address this is to put a cap on either the variance or the range of F. If we put a cap $z >> 1$, then it is unsurprising that the optimal F_i is to realize z with probability $1/z$ and 0 otherwise. We called such F_i earlier, “lottery distributions”. Notice that for such an F_i it is optimal for an agent to search by attribute if there is an outside option and it is optimal for an agent to search by object if there is not.

4.3 Object Search - unobservable F

If F is unobservable, then a firm wants to maximize the probability that they exceed 2, the agent’s ex-ante payoff. Take F_i to be a distribution which takes values $1 + \epsilon$ with probability $1/(1 + \epsilon)$ and 0 otherwise. The probability that the sum of two iid draws from F_i exceeds 2 is $1/(1 + \epsilon)^2 \approx 1$. So, the firm can choose a distribution which will lead it to be chosen with almost probability 1 conditional upon being originally selected. But, then by continuity, it must be able to select a distribution which guarantees it to be chosen with probability 1, and this is the degenerate distribution that puts all weight on 1 along with tie-breaking in favor of the selected firm. Interestingly, this is a distribution for which there is no value of searching and if the agent’s bore any cost for searching, then they would not.