School of Economics and Finance

Competing for Attention: Is the Showiest Also the Best?

Paola Manzini and Marco Mariotti

Working Paper No. 743 May 2015 ISSN 1473-0278

Queen Mary
University of London

Competing for Attention: Is the Showiest also the Best?*

Paola Manzini[†] Marco Mariotti[‡]

This version: April 2015

Abstract

There are many situations in which alternatives ranked by quality wish to be chosen and compete for the imperfect attention of a chooser by selecting their own salience. The chooser may be “tricked" into choosing more salient but inferior alternatives. We investigate when competitive forces ensure instead that “the showiest is the best", that is, when the best alternative is maximally salient (and the one that gets picked most often) in equilibrium. We prove that the structure of externalities in the technology of salience is key. Broadly speaking, positive externalities favour correlation between quality and salience.

J.E.L. codes: D010.

Keywords: Consideration sets, bounded rationality, stochastic choice.

*We thank Rani Spiegler, Chris Tyson and Roberto Veneziani for helpful comments and discussions. Financial support from ESRC grant RES-000-22-3474 is gratefully acknowledged.

†School of Economics and Finance, University of St. Andrews, Castlecliffe, The Scores, St. Andrews KY16 9AL, Scotland, U.K. (e-mail: paola.manzini@st-andrews.ac.uk) and IZA, Bonn.

‡School of Economics and Finance, Queen Mary University of London, Mile End Road, London E1 4NS, U.K. (e-mail: m.mariotti@qmul.ac.uk).

1 Introduction

Choice requires attention to what options are available, and attention may be imperfect. We study the competition between alternatives that want to be chosen by an imperfectly attentive chooser and select their own *salience*, namely the factors that lead to being noticed. This structure broadly fits several economic, political and social situations, as exemplified further below.

If a chooser suffers from cognitive imperfections he will not always pick the best alternative. Even so, for some leading models of choice with errors the probability (or frequency) of choice of an alternative can still be used as an indicator of its quality. For instance, in the popular multinomial logit (or Luce), where the true value of an alternative is imperfectly perceived, the resulting probabilities of choice are proportional to utility values.[1] But if instead the source of errors is the imperfect attention paid to alternatives with a given salience, this proportionality might not hold.[2] The most chosen alternatives may simply be the most salient, not the best ones.

What happens when alternatives can set their own salience? Do strategic forces tend to correct choice errors caused by lack of attention? To answer this question, we present a model of strategic salience that abstracts as much as possible from the specific nature of the alternatives. They could be politicians, producers, sexual partners, and so on. In this way, while we lose the details that colour any specific application, we are able to tease out some general relationships and mechanisms. Broadly speaking, we find that the key to the equilibrium relationship between quality and salience is the type of *externalities* that own salience has on the visibility of the other alternatives: positive externalities (in a sense to be made precise later) push in favour of the best alternative being also the most salient and the most chosen.

We use the concept of a *consideration set* (Wright and Barbour [22], Eliaz and Spiegler [7], [8]; Masatlioglu, Nakajima and Ozbay [13]; Manzini and Mariotti [12]) as a tool to

[1] The logit model is equivalent to the maximisation of a utility perturbed with i.i.d. Gumbel distributed utility errors, a classical result due to Holman and Marley, as attributed in Luce and Suppes [11], McFadden [14] and Yellot [23].

[2] See e.g. Manzini and Mariotti [12] and Brady and Rhebeck [2].

model agents with an imperfect ability or willingness to consider all the objects of choice that are physically available.[3] In these models, an agent's preference is only maximised on the subset of alternatives in the menu that the agent actively considers, or pays attention to, not necessarily on the whole menu. While obviously the consideration set is a theoretical construct that is not directly observable, there is evidence of situations in which the consideration set is strictly smaller than the menu (e.g.Sovinsky Goeree [20] documents that purchasers of personal computers are typically not aware of all available models when making a choice).[4]

That alternatives can influence the consideration set of the chooser is natural in disparate contexts. A minor politician can make an outrageous statement to get noticed by the media and thus enter the voters' consideration set, but he will likely incur a cost in terms of credibility. A new restaurant can offer special deals in the early days to lure consumers in. A single person in search of a partner can increase expenditure on hairdressing to get noticed. He may or may not like hairdressing: importantly, both situations will be captured in our model.

These examples also serve to illustrate a focal special case. If the aim of e.g. a political or promotional campaign is merely to generate *awareness* of say a candidate, outlet or brand (rather than to influence tastes) then it is natural to assume that awareness of one alternative is largely unaffected by the advertising campaigns for, and the awareness of, rival alternatives. Notably, this is a feature of Butters' [4] classical 'mailbox' model of informative advertising, in which a consumer is informed of a's existence if a has dropped an ad in the agent's mailbox.[5] And the empirical study by van Nierop *et al.* [15] does suggest that the probability of being noticed is independent across alternatives.

Assume that while salience is endogenous, the *quality* of an alternative (its position

[3]See also Roberts and Lattin [17] and Shocker et al. [19].

[4]The failure to seriously consider all alternatives may in fact stem from several sources. It could, for example, be the outcome of a search process, or of ideological prejudice (e.g. Wilson [21]). While in the paper we focus for simplicity on the attention interpretation, the possibility of other interpretations should be borne in mind.

[5]Butters' model is discussed more fully in section 7.

in the preference ranking of the chooser) is fixed. In this setting, competition for attention between alternatives gives rise to an interesting strategic situation, which we call an *attention game*. Attention games can have a rich externality structure (unlike e.g. the superficially similar *contest* games; see section 7 for details), which we exploit in our results.

Proposition 1 shows that if there are *no externalities*, in the sense that alternatives can fully control the attention they get (though obviously not the probability with which they are chosen), then 'the showier is the better': once asymmetries in the technology of salience are netted out, it can never happen that an alternative is strictly more salient in equilibrium than a better alternative. And therefore, because whenever two alternatives are both noticed the inferior one cannot be chosen (recall that the chooser maximises within the consideration set), alternatives of better quality are chosen more often that all those of lower quality. The showier is also the more chosen.

In some respects this result holds under quite general conditions. For example, a remarkable fact is that virtually no restrictions are needed on the strategy sets and on the cost structure of salience. Indeed, it is not necessary to assume that increasing salience requires spending additional resources: the result goes through when generating salience is a costless or even rewarding activity (this is another fundamental differences between attention games and contest games), or when there is a mixture of these cases.

Part of this conclusion can be generalised to a larger class of games that have *positive externalities*, in the sense that a weak supermodularity condition on the technology of salience is satisfied, and that alternatives do not damage the visibility of other alternatives by increasing own salience (Proposition 2). A shouting competition to get attention in a market is an example of a situation with negative externalities, while Butters' [4] model is an example with no externalities, and an industry in which a firm attracting attention to itself also attracts attention to the whole industry is an example of positive externalities.[6]

[6] An interesting example of positive externalities (in our sense) is provided by the idea in de Clippel, Eliaz and Rozen [5]: a market leader's high salience might induce otherwise passive (inattentive) con-

When alternatives lack full control on the attention they get and there may be harmful spillovers by rivals, no clear link between equilibrium salience and quality seem to exist. In this case salience may confound the quality ranking (Claim 1), but we also show that with a standard 'Luce' technology a the-showiest-is-the-best result is reinstated (Claim 2).

A different route to inferior alternatives being chosen with the highest probability is when salience is not a directional phenomenon but rather is *contextual*, in the sense that visibility depends on the alternative's location in a space of characteristics in relation to the location of other alternatives - e.g. dressing in white when the others dress in black (Claim 3).

Attention games turn out to have a nice hierarchical structure that ensures the existence of Nash equilibria in *pure* strategies in any finite game in several subclasses, notably including games with no externalities (Propositions 3, 4, 5, 6).

After introducing attention games (section 2) we present the main results in section 3. Section 4 describes situations with negative externalities, and section 5 situations of contextual salience. We discuss issues of existence in section 6. Section 7 discusses some related literature. Section 8 summarises and contains additional discussion and reference to the literature.

2 Attention games

Alternatives in a finite set $A = \{a_1, ..., a_n\}$ wish to be chosen by a chooser with imperfect attention. The chooser evaluates alternatives by means of a strict preference ordering $\succ$ on A. The position of an alternative in the ranking is its *quality*, with $a_i \succ a_j$ iff $i < j$.

The chooser maximises $\succ$ on a *consideration set* $C(A) \subseteq A$ of alternatives (the set of alternatives the chooser actively considers), which is formed stochastically in the manner explained below. When $C(A)$ is empty, the chooser is assumed to pick a de-

sumers to give closer scrutiny to the market as a whole. (Note that in their model this is not such good news, for there firms *dislike* attention!).

fault alternative a^* (e.g. walking away from the shop, remaining without a partner, abstaining from voting).

The probability that an alternative a_i belongs to $C(A)$ depends on a set of variables $\sigma_j \in \mathcal{R}$, $j = 1, ..., n$, one for each alternative. These variables indicate the ability of each alternative to attract attention (and possibly to dampen or increase the attention paid to the other alternatives). The strategy set for each a_i is a subset $S \subset \mathcal{R}$. Unless otherwise specified, we need not assume any additional structure on S. We call $\sigma_i \in S$ the *salience* of a_i, and a list $\sigma = (\sigma_1, ..., \sigma_n) \in S^n$ a *salience profile*. As usual we write σ_{-i} to denote the profile σ with the i^{th} entry omitted, $(\sigma_1, ..., \sigma_{i-1}, \sigma_{i+1}, ..., \sigma_n)$, and (σ_i, σ_{-i}) to denote σ. Given two vectors $v, v' \in \mathcal{R}^k$, $k \leq n$,we write $v \geq v'$ to signify $v_i \geq v'_i$ for all $i = 1, ..., k$.

The technology of salience is described by functions $p_i : S^n \to (0, 1)$, $i = 1, .., n$. Each p_i associates a salience profile with the probability of membership of $C(A)$ for a_i: that is, $p_i(\sigma)$ is the probability that a_i is noticed when the salience profile is σ. We assume that these probabilities are interior, namely there is always an (arbitrarily small) positive probability of being noticed or of not being noticed, independently of salience. The only further assumption we make for the moment on the functions p_i is the following:

Own Monotonicity: For all i, for all $\sigma \in S^n$: $\sigma_i > \sigma'_i \Rightarrow p_i(\sigma_i, \sigma_{-i}) > p_i(\sigma'_i, \sigma_{-i})$.

Own Monotonicity stipulates that increasing one's own salience strictly increases the probability of being noticed, whatever the salience of the other alternatives. An example of this type of function, which we will consider later in sections 4 and 7, takes the *Luce form*

$$p_i(\sigma) = \frac{\sigma_i}{\sigma_1 + ... + \sigma_n} \tag{1}$$

with the σ_is chosen on a strictly positive domain.[7] In this example, an increase in salience of the other alternatives is harmful to an alternative. While this is natural in some contexts, the opposite effect may occur. For instance, an alternative a_i which is

[7] We call this the Luce form in view of the Luce [10] stochastic choice rule, popular in econometrics in the multinomial logit version.

similar to, or dominated by, another alternative a_j may make the latter more prominent, so that an increase in the salience of a_i also increases a_j's visibility (the probability of it getting noticed). These possibilities are well known and documented in marketing science and psychology as the 'similarity' and 'attraction' effects.

The chooser picks the preferred alternative among those he considers. Therefore, the probability $\pi_i(\sigma)$ that alternative a_i is chosen at a salience profile σ is the probability of the compound event that it is considered and that none of the better alternatives is considered, that is

$$\pi_i(\sigma) = p_i(\sigma) \prod_{k<i} (1 - p_k(\sigma))$$

The payoff to each alternative is the probability of being chosen minus a (possibly negative) cost associated with the salience level that has been selected. An interpretation of this payoff function is that alternatives vie for one single chooser who chooses one alternative, yielding a unit payoff. Another interpretation is that alternatives face a continuum of identical choosers each of whom picks one alternative (the latter interpretation is adopted in Eliaz and Spiegler [7] as well as in Bordalo, Gennaioli and Shleifer [3]).

The payoff to alternative a_i for a pure strategy profile σ is

$$z_i(\sigma) = \pi_i(\sigma) - e_i(\sigma_i)$$

where $e_i : S \rightarrow \mathcal{R}$. We make no further assumption on the functions e_i. In particular, e_i may be convex or concave or neither, and may not be monotonic increasing. So e_i can represent both costly *effort* (for example an advertising budget) and *elation*, when increasing salience is pleasurable at least on some range (for example, hairdressing to become salient in competition for sexual partners).

An *attention game* is denoted (A, S, z), where $z = (z_1, z_2, ..., z_n)$ with the z_i defined above and satisfying Own Monotonicity.

An attention game has *absolute salience* if it satisfies the following condition:

Absolute Salience: for all $\sigma, \sigma'$and for all i: $\sigma_i = \sigma'_i \Rightarrow p_i(\sigma) = p_i(\sigma')$

That is, an attention game with absolute salience is one in which an alternative can decide its own probability of being *noticed* independently of the salience choices by the other alternatives. Of course, even in this case the strategic situation is not trivial: the *payoff* of an alternative typically still depends on the salience profile, since it depends on the probability of being chosen which in turns is determined (for all alternatives except the best) by the salience profile. As noted before, the situation captured by absolute salience fits for example the case of when ads for an alternative merely have the function of making the chooser actively aware of the alternative ('did you know that people who read book A also read book B?'; 'have you considered cycling to work'?). When salience is not absolute it is *relative*.

We study the pure strategy Nash equilibria of attention games.

3 Does salience reveal quality?

Obviously, an alternative that has access to a superior salience technology might produce more salience in equilibrium. The interesting issue is what the intrinsic strategic incentives to become salient are in an attention game once we strip out such technological asymmetries. We exhibit environments in which, when alternatives are ex-ante symmetric except for the difference in quality (they have access to the same technology of salience), the best alternative is guaranteed to have a highest salience level in equilibrium. In some of these, the stronger property holds that the equilibrium salience order (weakly) correlates with the quality order.

An attention game is *symmetric* when the following holds:

Symmetry:
(i) For all i, j and all $x, y \in S$:

$$p_i\left(\sigma_1, ..., \sigma_{i-1}, x, \sigma_{i+1}, ..., \sigma_{j-1}, y, \sigma_{j+1}, ..., \sigma_n\right) = p_j\left(\sigma_1, ..., \sigma_{i-1}, y, \sigma_{i+1}, ..., \sigma_{j-1}, x, \sigma_{j+1}, ..., \sigma_n\right)$$

(ii) For all i: $e_i = e$ for some $e : S \to \mathcal{R}$.

The first part of Symmetry says that the effectiveness of salience for getting noticed for a given configuration of the other alternatives' salience is the same for each alternative. More precisely, holding the salience of all alternatives fixed except for a_i and a_j, the attention attracted by a_i with a level of salience x when a_j has salience y is the same as the attention attracted by a_j with salience x when a_i has salience y. The second part of the condition simply says that achieving any level of salience has the same payoff implications for any two alternatives.

Remarkably, the main characterisation results below require no restriction on the structure of the strategy sets S, nor on that of the cost function e.

Proposition 1 (The showier is the better) *Let G be a symmetric attention game with absolute salience. Let σ be a pure strategy equilibrium of G. Then, $\sigma_i > \sigma_j \Rightarrow j > i$.*

Proof: We use a revealed preference argument. By contradiction, suppose that for some i, j with $i < j$ we had $\sigma_j > \sigma_i$. Since σ_j is optimal for j at σ, it must be in particular:

$$p_j(\sigma)\prod_{k<j}(1-p_k(\sigma)) - e(\sigma_j) \geq p_j(\sigma_i, \sigma_{-j})\prod_{k<j}(1-p_k(\sigma_1, ..., \sigma_{j-1}, \sigma_i, \sigma_{j+1}, ..., \sigma_n)) - e(\sigma_i)$$

(where we have used Symmetry (ii)). Because the game has absolute salience we have

$$p_k(\sigma_1, ..., \sigma_{j-1}, \sigma_i, \sigma_{j+1}, ..., \sigma_n) = p_k(\sigma)$$

and so the above inequality implies:

$$\left(p_j(\sigma) - p_j(\sigma_i, \sigma_{-j})\right)\prod_{k<j}(1-p_k(\sigma)) \geq e(\sigma_j) - e(\sigma_i).$$

Similarly, from the optimality of σ_i we obtain:

$$\left(p_i(\sigma_j, \sigma_{-i}) - p_i(\sigma)\right)\prod_{k<i}(1-p_k(\sigma)) \leq e(\sigma_j) - e(\sigma_i).$$

By absolute salience,

$$p_j(\sigma_i, \sigma_{-j}) = p_j(\sigma_1, ..., \sigma_{i-1}, \sigma_j, \sigma_{i+1}, ..., \sigma_{j-1}, \sigma_i, \sigma_{j+1}, ..., \sigma_n)$$

and then by Symmetry (i) $p_j(\sigma_i, \sigma_{-j}) = p_i(\sigma)$. Similarly, we have $p_j(\sigma) = p_i(\sigma_j, \sigma_{-i})$. Moreover, by Own Monotonicity and $\sigma_j > \sigma_i$, $p_i(\sigma_j, \sigma_{-i}) - p_i(\sigma) > 0$. It follows that the above inequalities imply

$$\prod_{k<i}(1-p_k(\sigma)) \leq \prod_{k<j}(1-p_k(\sigma))$$

which is impossible given the range assumption on the p_k. ∎

So, for games of absolute salience, there is a (weak) correlation between equilibrium salience and quality ranking: not only is the best alternative one of those with the highest salience, but also a higher quality implies a higher salience at any level of the quality hierarchy.

The the-showi*est*-is-the-*best* feature of this result (but not the general correlation between quality and salience) is preserved if we weaken absolute salience to "positive externalities, as expressed by the following conditions:

Weak Supermodularity: For all i, all $\sigma_{-i}, \sigma'_{-i} \in S^{n-1}$ with $\sigma'_{-i} \geq \sigma_{-i}$, and all $x, y \in S$ with $x > y$:

$$p_i\left(x, \sigma'_{-i}\right) - p_i\left(y, \sigma'_{-i}\right) \geq p_i\left(x, \sigma_{-i}\right) - p_i\left(y, \sigma_{-i}\right)$$

Cross Monotonicity: For all i, all $\sigma_{-i}, \sigma'_{-i} \in S^{n-1}$ with $\sigma'_{-i} \geq \sigma_{-i}$, $p_i\left(\sigma_i, \sigma'_{-i}\right) \geq p_i\left(\sigma_i, \sigma_{-i}\right)$.

Both conditions impose positive spillovers of own salience on the rivals, with Weak Supermodularity acting on the first difference of the p_i and Cross Monotonicity acting on the absolute levels.

Cross Monotonicity says that an alternative cannot harm the visibility of the rivals by raising own salience. This is verified in games of absolute salience. Another situation captured by the property is when advertising by an individual (e.g. a firm selling a new type of product) draws attention to similar individuals (the firms selling similar products) and hence to the whole group (industry).

Weak Supermodularity says that the effectiveness for getting noticed of an increase in an alternative's own salience increases with the salience of the other alternatives. Note that this is a condition on the function p_i only: *the whole payoff function need not be supermodular even when Weak Supermodularity is satisfied*. If salience is absolute then Weak Supermodularity is trivially satisfied. The condition might also be verified in the similarity within an industry example given above: the salience effort of an individual firm in a little known industry is likely to be less effective than when other firms selling similar products are well-known.

Proposition 2 (The showiest is the best). *Let σ be a pure strategy equilibrium of a symmetric attention game which satisfies Weak Supermodularity and Cross Monotonicity. Then $\sigma_1 \geq \sigma_j$ for all $j > 1$.*

Proof: Suppose by contradiction that $\sigma_j > \sigma_1$ for some $j > 1$. The optimality condition for a_j at σ implies in particular:

$$p_j(\sigma)\prod_{k<j}(1-p_k(\sigma)) - e_j(\sigma_j) \geq p_j(\sigma_1,\sigma_{-j})\prod_{k<j}(1-p_k(\sigma_1,...,\sigma_{j-1},\sigma_1,\sigma_{j+1},...,\sigma_n)) - e_j(\sigma_1)$$

By Cross Monotonicity and $\sigma_1 < \sigma_j$ we have

$$\begin{aligned} p_k(\sigma) &\geq p_k(\sigma_1,...,\sigma_{j-1},\sigma_1,\sigma_{j+1},...,\sigma_n) \text{ for all } k \neq j \\ &\Rightarrow \prod_{k<j}(1-p_k(\sigma_1,...,\sigma_{j-1},\sigma_1,\sigma_{j+1},...,\sigma_n)) \geq \prod_{k<j}(1-p_k(\sigma)) \end{aligned}$$

and therefore

$$p_j(\sigma)\prod_{k<j}(1-p_k(\sigma)) - e_j(\sigma_j) \geq p_j(\sigma_1,\sigma_{-j})\prod_{k<j}(1-p_k(\sigma)) - e_j(\sigma_1)$$

or

$$(p_j(\sigma) - p_j(\sigma_1,\sigma_{-j}))\prod_{k<j}(1-p_k(\sigma)) \geq e_j(\sigma_j) - e_j(\sigma_1) \quad (2)$$

To avoid cumbersome notation, denote, for $x, y \in S$,

$$(y, x, \sigma_{-1j}) = (y, \sigma_2, ..., \sigma_{j-1}, x, \sigma_{j+1}, .., \sigma_n)$$

the profile in which a_1 plays y, a_j plays x and all other alternatives play as in σ. By Weak Supermodularity (2) implies

$$(p_j(\sigma_j,\sigma_j,\sigma_{-1j}) - p_j(\sigma_j,\sigma_1,\sigma_{-1j}))\prod_{k<j}(1-p_k(\sigma)) \geq e_j(\sigma_j) - e_j(\sigma_1).$$

Then, by Symmetry

$$(p_1(\sigma_j,\sigma_j,\sigma_{-1j}) - p_1(\sigma_1,\sigma_j,\sigma_{-1j}))\prod_{k<j}(1-p_k(\sigma)) \geq e(\sigma_j) - e(\sigma_1)$$

or

$$\left(p_1\left(\sigma_j,\sigma_{-1}\right)-p_1\left(\sigma\right)\right)\prod_{k<j}\left(1-p_k\left(\sigma\right)\right)\geq e\left(\sigma_j\right)-e\left(\sigma_1\right)$$

Since $\sigma_1<\sigma_j$, by Own Monotonicity we have $p_1\left(\sigma_j,\sigma_j,\sigma_{-1j}\right)-p_1\left(\sigma_1,\sigma_j,\sigma_{-1j}\right)>0$, and by the range assumption on the p_k we have $1>\prod_{k<j}\left(1-p_k\left(\sigma\right)\right)>0$. We conclude that

$$\left(p_1\left(\sigma_j,\sigma_{-1}\right)-p_1\left(\sigma\right)\right)>e\left(\sigma_j\right)-e\left(\sigma_1\right).$$

But this means that a_1 would improve by deviating from σ_1 to σ_j at profile σ, a contradiction. ■

The the-showiest-is-the-best result is unrelated to any kind of signalling argument. There is no hidden quality to signal to the chooser via costly investment. The reason why lower quality alternatives never produce more salience in equilibrium does not derive either from lower levels of resources or lower unit costs of salience production, both types of asymmetry having been ruled out: every alternative can choose from exactly the same set at exactly the same cost or benefit. The result is purely a function of the cognitive process postulated for the chooser.

Why do the conditions on p_i turn out to be important for the result? Let's focus on the incentives to raise own salience. On the one hand, increasing own salience attracts attention to oneself. But - depending on the p_is - there may also be a second effect: that of *detracting* attention from other alternatives. This second part of the incentive is *always stronger* for lower quality alternatives. Good alternatives "do not care" whether worse alternatives are noticed or not - their payoff only depends on the probability that even better alternatives are noticed. In the extreme case, the best alternative only cares about its own probability of getting noticed, so that the only incentive it has is of the first type. Thus if the p_is are decreasing in the salience of other alternatives, there is a tendency for a salience rise to be the more profitable the *lower* the quality of an alternative. Cross Monotonicity removes this tendency. Assume now that a worse alternative w finds it profitable to raise salience from l to $h>l$ while a better alternative b chooses l. Can this be an equilibrium? Suppose that b is the only other alternative beside w. Note that b's gain in visibility when moving from l to h would be scaled up compared to that of w, by a factor strictly greater than one, given by the (reciprocal of the) probability that b is not noticed: only if b is not noticed, in fact, will w be picked if noticed. So the only reason why b might not want to follow w in raising salience to h is that raising own salience becomes less effective for becoming more noticeable when the rivals raise *their* salience, which w has done. Weak supermodularity eliminates precisely this effect. So if it were profitable for the worse alternative w to raise salience to h, it would be a fortiori profitable for the better alternative b, and therefore the initial configuration could not be an equilibrium.

However, the above reasoning might not hold if there was a third alternative t that is even better than b, and b's possible increase in salience was beneficial to t while w's increase in salience was not. In this case, b might prefer to keep its salience low in order not to draw attention to t, while w would not face this negative incentive. It is for this reason that, once we move beyond games of absolute salience and allow for positive externalities, *only the top alternative* is sure to produce more salience in equilibrium: the peculiarity of the top alternative is that it is the only one that does not need to worry about the effect of its own salience on the attention paid to the other alternatives. In this sense there is a difference between the way in which the top alternative is better than the second best and that in which the second best is better than the third best. The next example illustrates this fact with a three-player attention game satisfying all the conditions of proposition 2 in which the second best alternative is less salient in equilibrium than the third best one.

Example 1 Let $\sigma_i \in \{l, h\}$ with $i \in \{1, 2, 3\}$ and $h > l$. There is a common effort function e with $e(l) = 0$ and $e(h) = \frac{3}{32}$. Let $p_i(\sigma)$ be given by the following table:

	hhh	hhl	hlh	hll	lhh	lhl	llh	lll
$p_1(\sigma)$	$\frac{3}{4}$	$\frac{5}{8}$	$\frac{5}{8}$	$\frac{5}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$p_2(\sigma)$	$\frac{3}{4}$	$\frac{5}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$p_3(\sigma)$	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{1}{8}$	$\frac{5}{8}$	$\frac{1}{8}$

The profile $\sigma = (h, l, h)$ is an equilibrium, since we have:

$$\pi_1(h,l,h) = \frac{17}{32} \geq \frac{1}{8} = \pi_1(l,l,h)$$
$$\pi_2(h,l,h) = \frac{3}{32} \geq \frac{3}{32} = \pi_2(h,h,h)$$
$$\pi_3(h,l,h) = \frac{21}{256} \geq \frac{21}{512} = \pi_3(h,l,l)$$

It is easy to check that Weak Supermodularity and Cross-Monotonicity hold (details available from the authors).

Note in Example 1 that

$$p_1(h,l,h) - p_1(h,l,l) = \frac{5}{8} - \frac{5}{8} < \frac{3}{4} - \frac{5}{8} = p_1(h,h,h) - p_1(h,l,h)$$

This means that when the worst alternative a_3 increases its own salience from l to h there is no spillover on the visibility of the best alternative a_1; but the same increase at (h,l,h) on the part of the middling alternative a_2 would have a positive externality on a_1. This deters a_2 from raising its own salience up to the level of a_1.

Observe in passing how effects such as those just discussed make clear the stark difference between *All Pay Contests* and attention games, to which they might at first sight seem to bear some similarity: in an auction, an increase in one's own bid can only be weakly detrimental for a rival. Attention games have a richer externality structure. We expand on this point in section 7.

Example 1 also shows that the statements of the results cannot be strengthened to obtain a *strict* correlation between salience and quality. The the-showiest-is-the-best property ensures that if there is an alternative that is *uniquely* maximally salient in equilibrium, then that alternative must also be the best one; but it does not exclude that alternatives of lower quality tie with the best for salience. However, even if we have to settle for a weak correlation, an immediate but important implication of Proposition 2 is:

Corollary 1 (The best gets picked most often) *Let σ be a pure strategy equilibrium of a symmetric attention game satisfying Weak Supermodularity and Cross Monotonicity. Then alternatives of higher quality are chosen with strictly greater probability. That is, $i < j \Rightarrow \pi_i(\sigma) > \pi_j(\sigma)$.*

Corollary 1 can be read from a revealed preference perspective. Suppose that neither the salience of alternatives nor the chooser's preferences are observable to an outside party, but that this party knows the structure of the game. Then, Corollary 1 implies that the observer could still infer - under the assumption of no technological disadvantage for the best alternative - the preference of the chooser for the top alternative from choice data, simply by checking the choice frequencies. However the frequencies of other alternatives might not faithfully reveal their quality ranking: quality is revealed by choice only for the top alternative.

A second perspective from which Corollary 1 can be read is as showing circumstances in which competitive forces counterbalance the cognitive limitations of the chooser. Competition pushes the best alternatives to raise their salience sufficiently to overcome the distorsive effects of imperfect attention on the relative popularity of the alternatives.

4 Negative externalities: examples with the Luce form

When salience is relative, and Weak Supermodularity or Cross Monotonicity fail, the neat equilibrium properties obtained in the previous section may break down. Perverse equilibria become possible, in which the worst alternative is selected with the strictly highest probability ('perverse' is meant in contrast to the situation that holds for standard models of choice errors, such as the logit).

Claim 1 *There are attention games of relative salience with equilibria in which the worst alternative is chosen with the strictly highest probability.*

This claim is shown with a two-alternative example in which $S = \{l, h\}$ with $h > l$ and the probability of being noticed has the Luce form (1):

$$p_i(\sigma) = \frac{\sigma_i}{\sigma_i + \sigma_j}$$

for $i \neq j$. We impose the following restrictions on the admissible values of h and l:

$$\begin{aligned} h &> \frac{9}{32} > l > 0 \\ h+l &> \frac{1}{2} \\ h &\in \left[\frac{3}{8} - l - \frac{1}{8}\sqrt{9-32l}, \frac{3}{8} - l + \frac{1}{8}\sqrt{9-32l}\right] \end{aligned}$$

Let $e_i(x) = x$ for $i = 1,2$ and $x \in \{h, l\}$, so we drop the subscript. Then the profile $\sigma = (l, h)$ in which the showiest is the strictly worst is a strict Nash equilibrium for all admissible values of l and h. Not all admissible profiles of this type also have the property that alternative a_2 is picked with higher probability. Such profiles do exist, however (the simple calculation is available from the authors upon request).

However, at least limited forms of the the-showiest-is-the-best property also hold in some "natural" attention games despite them failing both Weak Supermodularity and Cross Monotonicity. We illustrate the point with another Luce form example in which alternatives can choose any positive level of salience. Even in the two player case, whether or not the supermodularity condition $p_i\left(x, \sigma'_j\right) - p_i\left(y, \sigma'_j\right) \geq p_i\left(x, \sigma_j\right) - p_i\left(y, \sigma_j\right)$ holds depends on the sign of $\sigma'_j\sigma_j - xy$. Moreover, any alternative's increase in own salience is detrimental to the rivals' chances of being noticed. Nevertheless, we show that the the-showiest-is-the-best result fully applies in the two-alternative case, or to the two best alternatives in any game with $n > 2$.

Claim 2 *Let $G = (A, S, z)$ be a symmetric attention game with $S = \mathcal{R}_{++}$; $p_i(\sigma) = \frac{\sigma_i}{\sigma_1 + \ldots + \sigma_n}$ for all i, all σ; and with e twice differentiable and weakly convex. Then at any equilibrium σ of G the salience chosen by the best alternative is never lower than that chosen by the second best alternative; that is, $\sigma_1 \geq \sigma_2$.*

Proof: Denote $k = \sigma_3 + \ldots + \sigma_n \geq 0$. At any interior equilibrium the FOCs for alternatives 1 and 2 must be satisfied

$$\begin{aligned} \frac{\partial\left(\frac{\sigma_1}{\sigma_1+\sigma_2+k}\right)}{\partial\sigma_1} &= e'(\sigma_1) \\ \frac{\partial\left(\frac{\sigma_2}{\sigma_1+\sigma_2+k}\frac{\sigma_2+k}{\sigma_1+\sigma_2+k}\right)}{\partial\sigma_2} &= e'(\sigma_2) \end{aligned}$$

where e' denotes the first derivative of e. Dividing side by side the two equations and simplifying we get

$$(k+\sigma_2)\frac{k+\sigma_1+\sigma_2}{k\sigma_1+k\sigma_2+2\sigma_1\sigma_2+k^2}=\frac{e'(\sigma_1)}{e'(\sigma_2)} \tag{3}$$

Suppose by contradiction that $\sigma_2 > \sigma_1$. Then, by the weak convexity of e, $\frac{e'(\sigma_1)}{e'(\sigma_2)} \leq 1$. It follows from (3) that

$$\begin{aligned}(k+\sigma_2)\frac{k+\sigma_1+\sigma_2}{k\sigma_1+k\sigma_2+2\sigma_1\sigma_2+k^2} &\leq 1 \Leftrightarrow \\ \sigma_1 &\geq \sigma_2+k\end{aligned}$$

a contradiction in view of $k \geq 0$. ■

The examples in this section suggests that, at least with the Luce form, the possibility of perverse equilibria is related to restrictions in the strategy spaces of the players. When these spaces are sufficiently rich, as shown in Claim 2 the better alternative will always find it profitable to deviate from a profile at which the inferior alternative is more salient. This sensitivity to the domain somehow dampens hopes of finding a nice general condition that is necessary as well as sufficient for a the-showiest-is-the-best results to hold.

5 Contextual salience

There is a different way in which the the-showiest-is-the-best property may collapse. So far we have assumed (through Own Monotonicity) that salience, whether absolute or relative, is a 'directional' attribute for which 'the more is always the better': the more commercials you produce, the louder you shout, the glitzier your clothes, the more likely - ceteris paribus - you are to get noticed. In some scenarios, however, alternatives can only control variables whose values are not *intrinsically* positive or negative for the aim of attracting attention; whether they are depends, for each alternative, on what the other alternatives do. If everybody else dresses in green you will be salient by dressing in yellow, and viceversa. If all other candidates converge on a given political message, you will stand out by deviating from that message. We call this scenario one of *contextual* salience.

We study a simple stylised class of models of contextual salience that generates the perverse result. Suppose that $\sigma_i \in [0,1]$ is now a 'position' selected by alternative a_i in the unit interval. Whether or not the probability that a_i is noticed is increasing in σ_i depends on the entire profile σ: that is, Own Monotonicity may fail. In particular, we assume that an alternative's probability of being noticed is conferred by its distance from the 'average alternative' (excluding itself):

$$p_i(\sigma) = \alpha_i \left(v_i - \frac{\sum_{j \neq i} v_j}{(n-1)} \right)^2 \in [0,1]$$

where $\alpha_i \in (0,1)$ for all i can be seen for instance as a psychological parameter indicating how naturally inclined the chooser is to notice any given alternative. Finally assume for simplicity a null effort-elation function so that we can take

$$z_i(\sigma) = p_i(\sigma) \prod_{k<i} (1 - p_k(\sigma))$$

Claim 3 *There exist (for some values n and $\alpha_1, ..., \alpha_n$) pure strategy Nash equilibria of the game above in which the worst alternative is chosen with the highest probability.*[8]

The case of contextual salience we have studied has some superficial similarity with location games à la Hotelling. In that case, too, alternatives can gain by moving away from the nearest neighbour. However, the payoff structure is in fact very different. This results, unlike a location game, in an existence result with three players, as detailed in the proof of Claim 3.

The intuition for the above result is as follows. Consider a situation with three alternatives, in which the two better alternatives a_1 and a_2 are conformist, while a_3 is non-conformist: in particular, a_1 and a_2 bunch at one extreme position, whereas a_3 occupies a solitary position at the other extreme. Then, for players a_1 and a_2, the average position of the other players is $\frac{1}{2}$. Alternative a_1 cares only about its own probability of being noticed, and because of the convexity of the p_i functions this is maximal at maximum distance from the average position $\frac{1}{2}$, that is at the extremes: so, a_1 cannot

[8]The proof is relegated to an Appendix.

gain by deviating. Alternative a_2's probability of being noticed responds exactly like that of a_1 to changes in own position, but its probability of being *chosen* also depends (negatively) on a_1's probability of being noticed. By moving away from position 0, a_2 makes a_1 more noticeable, and therefore a_2 does not gain by deviating either. As for a_3, by locating as far away as possible from the rest, it maximises its own probability of being noticed. Of course this also makes the opponents more noticeable, but it is not difficult to find parameter configurations for which the benefits outweigh the costs. This shows that the configuration of positions is a Nash equilibrium, which, for appropriate choices of parameters, has the feature that the worst alternative is chosen with the highest probability.

6 Existence

While standard existence results apply for many reasonable specifications of the model, in this section we highlight some peculiar features of attention games, which may be useful in applications. Attention games with absolute salience are - thanks to their 'hierarchical' structure - very well behaved in terms of existence properties. A *pure* strategy equilibrium is guaranteed in these games for standard strategy sets (like finite ones) for which in general the extension to mixed strategies or additional assumptions are needed:

Proposition 3 *Let $G = (A, S, z)$ be an attention game with absolute salience. Suppose that S is finite. Then G has an equilibrium in pure strategies.*

And:

Proposition 4 *Let $G = (A, S, z)$ be an attention game with absolute salience. Suppose that S is compact and e_i and p_i are continuous functions of σ_i for all i. Then G has an equilibrium in pure strategies.*

Propositions 3 and 4 are particular cases of more general ones. The existence results below holds for a large class of attention games - including games of absolute salience - namely those satisfying the following condition

Worse Alternative Independence : For all $\sigma, \sigma' \in S^n$, for all i: $\sigma_j = \sigma'_j$ for all $j \leq i \Rightarrow p_i(\sigma) = p_i(\sigma')$.

Worse Alternative Independence says that the probability of being noticed for an alternative a_i depends only on the salience of the alternatives which are better than a_i. Note that this is a strictly weaker requirement than absolute salience. To avoid confusion, we stress that in *any* salience game it is true that the probability of being *chosen* depends only on the choice probability (hence on the salience) of better alternatives. What is further asserted by Worse Alternative Independence is that a similar structure holds for the probability of being *noticed* of an alternative in relation to the salience of the other alternatives.

Proposition 5 *Let $G = (A, S, z)$ be an attention game such that Worse Alternative Independence holds. Suppose that S is finite. Then G has an equilibrium in pure strategies.*

Proof. Define $p_1^1 : S \to (0,1)$ by $p_1^1(\sigma_1) = p_1(\sigma_1, \sigma_{-1})$ for all σ_{-1}. The function p_1^1 is well-defined by Worse Alternative Independence. At a pure strategy equilibrium, alternative a_1 simply solves the one-alternative problem

$$\max_{\sigma_1 \in S} p_1^1(\sigma_1) - e_1(\sigma_1)$$

Given the assumption on S, a solution to this problem exists. Let σ_1^* denote such a solution. We construct an equilibrium recursively. For all $k \leq n$ and $\sigma \in S^n$ denote $\sigma^k = (\sigma_1, ..., \sigma_k) \in S^k$. For $i \leq k$ let

$$p_i^k : S^k \to (0,1)$$

be given by $p_i^k(\sigma^k) = p_i(\sigma)$ for all $\sigma \in S^n$ (Worse Alternative Independence ensures that p_i^k is well-defined). Suppose that we have defined the components $\left(\sigma_1^*, ..., \sigma_{j-1}^*\right)$ of a pure strategy equilibrium for the first $j-1$ alternatives. Then the j^{th} component σ_j^* is defined by selecting a solution to the problem

$$\max_{\sigma_j \in S} p_j^j\left(\sigma_1^*, ..., \sigma_{j-1}^*, \sigma_j\right) \prod_{k<j}\left(1 - p_k^j(\sigma_k^*)\right) - e_j(\sigma_j) \tag{4}$$

(a solution obviously exists). No alternative a_j can profitably deviate at $\sigma^* = (\sigma_1^*, ..., \sigma_n^*)$ thus constructed. In fact, if it were

$$p_j\left(\sigma_j', \sigma_{-j}^*\right) \prod_{k<j} (1 - p_k(\sigma^*)) - e_j\left(\sigma_j'\right) > p_j(\sigma^*) \prod_{k<j} (1 - p_k(\sigma^*)) - e_j\left(\sigma_j^*\right)$$

for some $\sigma_j' \in S$, then by the definition of the p_i^k also

$$p_j^j\left(\sigma_1^*, ..., \sigma_{j-1}^*, \sigma_j'\right) \prod_{k<j} \left(1 - p_k^j(\sigma_k^*)\right) - e_j\left(\sigma_j'\right) > p_j^j\left(\sigma_1^*, ..., \sigma_{j-1}^*, \sigma_j^*\right) \prod_{k<j} \left(1 - p_k^j(\sigma_k^*)\right) - e_j\left(\sigma_j^*\right)$$

so that σ_j^* would not solve problem 4, a contradiction. ■

By an analogous reasoning and standard facts about the existence of maxima of a function, we also have:

Proposition 6 *Let $G = (A, S, z)$ be an attention game such that Worse Alternative Independence holds. Suppose that S is compact and that e_i and p_i are continuous functions of σ_i for all i. Then G has an equilibrium in pure strategies.*

Compared to standard existence results, note that in the statement of Proposition 6 no convexity assumptions are made - so that Kakutani-type fixed point arguments do not apply - and that the best reply functions could be non-monotonic - so that Tarsky-type fixed point theorems do not apply either. Notice also that p_i is only required to be continuous in σ_i, not necessarily in any σ_j for $j \neq i$. Finally, observe that Own Monotonicity is not used in the result.

The following example of a finite game that only has an equilibrium in mixed strategies shows that Worse Alternative Independence is necessary in the statements.

Example 2 Let $\sigma_i \in \{l, h\}$ with $i \in \{1, 2\}$ and $h > l$, with

$$\begin{aligned} p_1(l,l) &= \frac{1}{2} = p_1(l,h) \\ p_1(h,l) &= \frac{2}{3}, \ p_1(h,h) = \frac{3}{5} \end{aligned}$$

and

$$\begin{aligned} p_2(l,l) &= \frac{1}{3} = p_2(h,l) \\ p_2(l,h) &= \frac{5}{12}, \ p_2(h,h) = \frac{2}{3} \end{aligned}$$

The effort functions are given by $e_1(l) = e_2(l) = 0$, $e_1(h) = \varepsilon$ and $e_2(h) = \eta$. Worse Alternative Independence fails since $p_1(h,l) \neq p_1(h,h)$. The matrix below, in which a_1 plays rows and a_2 plays columns, summarises the payoffs:

	l	h
l	$\frac{1}{2}, \frac{1}{2}\frac{1}{3}$	$\frac{1}{2}, \frac{1}{2}\frac{5}{12} - \eta$
h	$\frac{2}{3} - \varepsilon, \frac{1}{3}\frac{1}{3}$	$\frac{3}{5} - \varepsilon, \frac{2}{5}\frac{2}{3} - \eta$

It is easy to check that for the game has no pure strategy equilibrium for the parameter values $\frac{1}{10} < \varepsilon < \frac{1}{6}$, $\frac{1}{24} < \eta < \frac{7}{45}$.

It would be nice to obtain similar general existence results for all games considered in the characterisation result of proposition 2, but we have so far been unable to do so. Note, however, that there are many examples of games satisfying all conditions in proposition 2 as well as standard existence conditions, so that the characterisation result is not vacuous. To briefly illustrate, the following is a class of symmetric, weakly supermodular and cross monotonic games:

Example 3 Let $\sigma_i \in [0,1]$ with $i \in \{1,2\}$. Let $p_i(\sigma_i, \sigma_j) = \sigma_i^\alpha \sigma_j^\beta$, $\alpha, \beta > 0$, and $e_i = e$ for some continuous and convex function e, $i = 1,2$.

In this example, each alternative's contribution of its own salience to its own visibility is measured by α and the contribution to the rival's visibility is measured by β. It is easy to check that the payoff function of a_1 is concave in σ_1 and that the payoff function of a_2 is quasi-concave in σ_2 for all values of the parameters.[9] Then the game has a pure strategy equilibrium by the Debreu-Glicksberg-Fan theorem.[10]

[9]To this effect, check that $\partial \frac{\pi_i(\sigma_i,\sigma_j)}{\partial \sigma_2} > 0$ for σ_2 sufficiently small and that the derivative changes sign at most once on the interval $[0,1]$. This implies that π_2 is quasi-concave in $\sigma_2 \in [0,1]$ even if it is necessarily convex near zero.

[10]One notable such equilibrium for $e = 0$ is $\sigma_1 = 1$ and $\sigma_2 = \left(\frac{\alpha}{\alpha+\beta}\right)^{\frac{1}{\beta}}$ so that the salience ordering of proposition 2 holds strictly: in spite of the absence of costs the inferior alternative refrains from becoming more salient because in doing so it would also raise the visibility of the rival, to the point of harming its own chances of being chosen.

7 Related literature

The literature that deals with various aspects of attention is getting extensive - here we focus only on a subset of works that is most closely related to ours. For studies of attention *within* a consideration set (where attention is selectively applied to specific features of the alternatives), see for example the recent approach in Bordalo, Gennaioli and Shleifer [3] and the works cited therein. Eliaz and Spiegler ([7] and [8], henceforth ES) study the strategic aspects of the competition between firms to make products (or the firm's brand name as a whole) enter the consideration sets of consumers. Beside the broad conceptual relation, the focus of the analysis there is quite different from ours. The heart of the ES models is the mechanism of attracting attention via the choice of offered product *menus* (or more in general, of marketing strategies) on the part of firms. Firms deploy marketing strategies devoted to attention grabbing, and one main result of the analysis is that in equilibrium firms will include pure attention grabbers in the menu. Our emphasis has been rather on the relationship between (exogenous) quality on the one hand, and equilibrium *salience* and its effectiveness (the probability of being chosen) on the other. In order to pose and answer the questions of this paper we need to be able to talk about the degree of salience and the probability of being chosen. These variables are not identifiable as part of the elements of the ES models. Precisely because marketing strategies are more complex objects in ES than in this paper, there is no notion of marketing 'intensity' equivalent to our notion of salience. Moreover, the choice model and the description of the chooser's psychology at the heart of the ES' work is different from ours, and related to the one axiomatised in an abstract context by Masatlioglu, Nakajima and Ozbay [13] using revealed preference techniques. In these models choice is deterministic rather than stochastic. In ES an attention function determines *whether* other products are paid attention to, depending on a default set of products, but not *how much* attention is paid to any object. Our device of stochastic attention is the key for obtaining a more nuanced ranking of salience effectiveness. Stochastic choice in a consideration set model has been axiomatised in Manzini and Mariotti [12]: from this perspective, the present work offers a mechanism to endogenise, as equilibrium values, the salience parameters of our previous work.

There is of course a vast literature specific to advertising competition in retail markets, and some forms of informative advertising can be interpreted as placing a product in the consideration set. Our model is more stylised than the models in this literature (we refer the reader to Bagwell [1] for an extensive survey), where the issue of 'attracting attention' is intertwined with that of pricing strategies and other features of the market, and where different questions from the ones we have posed are addressed. We illustrate this with the advertising model which is conceptually the closest, namely Butters [4]. In this model, there are many sellers and buyers. Producers of a homogeneous good can pay a fixed cost to inform a random number of buyers of their price and location. If a (potential) buyer is not informed, he cannot buy the product: in this sense, Butters' can be seen as a model based on consideration sets. The focus is the limit equilibrium price distribution (as the number of buyers and sellers grows to infinity). A first major difference between Butters' and our analysis is that the latter, but not the former, fits situations where there is only a single chooser: this is because the probability of noticing an alternative is a feature of the psychology of the individual chooser, whereas the statistical attention effects in models à la Butters depend on a 'mailbox' model of information transmission, in which k out of n agents are randomly informed. There is no information processing by an individual agent: if he gets the ad in the mailbox he is informed, otherwise not. Aside from this, the relationship between quality and salience, on which we have focused, cannot be studied in Butters' model, because what is being traded is a single homogeneous good for which all buyers have the same willingness to pay (which in terms of our model can be interpreted as an assumption that alternatives are not differentiated by quality). In a conceivable variation in which producers sell products differentiated by quality, price would remain a key strategic variable: what this class of models addresses is the trade-off between the price and the information variables, whereas we have sought to study an equilibrium relationship between salience and quality that holds also in environments in which prices are not relevant, such as political competition (as noted by ES, who also eschew prices, this may apply even in some market contexts, such as media markets).

While motivated very differently, it may be worth noting that attention games bear some superficial structural similarity with (separable) 'All-Pay Contests' (henceforth APCs, see Siegel [18] for a very general treatment). In a separable APC, the payoff function has the form $u_i(\sigma) = P_i(\sigma) V_i - c_i(\sigma_i)$, where $\sigma = (\sigma_1, ..., \sigma_n)$ is interpreted as a 'score' (e.g. lobbying intensity) profile, P_i is the probability of winning for player i at profile σ, V_i is the value of the prize for i , and c_i captures the (increasing) cost of a score. However, the 'benefit' part of this expression does not include that of attention games, since in an APC $P_i(\sigma) = 1$ if i is on the m agents with the highest score, and $P_i(\sigma) = 0$ if it is not. In particular, increasing one's own score can only have negative spillovers on the rivals. Moreover, in attention games the e_i functions are essentially unrestricted, whereas it is a defining feature of APCs that increases in the probability of winning come at a cost. So there is in fact almost no relation between attention games and APC, and our results are independent of the existing results in the theory of APCs.

Differential attention to alternatives is not necessarily tied to a consideration set interpretation. One interesting example is the recent work by Echenique, Saito and Tserenjigmid [6] (henceforth EST), who study a Luce-type (Luce [10]) model of stochastic choice in which the *perception* of alternatives by the decision maker is hierarchical. An alternative can only be chosen if the alternatives that precede it in a perception priority ranking is (randomly) not chosen. Given that in our setting the set of alternatives is fixed, it turns out that our model is also consistent with the EST interpretation. More precisely, in a slightly simplified version of their choice model, the primitives are a utility function $u : A \to \mathcal{R}$ and a perception ordering $\succ_p$.[11] The probability that $a \in A$ is chosen is

$$\frac{u(a)}{\sum_{b \in A} u(b) + u(A)} \left(\prod_{c \in A : c \succ_p a} \left(1 - \frac{u(c)}{\sum_{b \in A} u(b) + u(A)} \right) \right)$$

where $u(A)$ is the (menu dependent) value of choosing some default option. Defining, for all $a \in A$

$$\mu(a, A) = \frac{u(a)}{\sum_{b \in A} u(b) + u(A)}$$

[11]In the general version the perception ordering is a weak order, rather than a strict order, thus allowing for "perception ties".

this choice model falls within the class of choice models we are considering. To see this, set $\succ = \succ_p$and set $\sigma_i = u(a_i)$, so that $p_i(\sigma) = \mu(a_i, A)$. In this interpretation, we assume that alternatives cannot change the perception ranking, just as in the consideration set interpretation we assumed that they cannot change the quality ranking. However, they can compete by striving to raise their quality (utility) and thus increase the probability of being selected in the event that no alternative that is higher ranked in the perception ordering is itself selected. Observe that $\mu(a_i, A)$ is increasing in the utility of a_i and decreasing in the utility of the other alternatives, so that Own Monotonicity for an attention game is satisfied. The model is not one of absolute salience, and not even one satisfying Worse Alternative Independence, since $p_i(\sigma)$ depends on the utility of all other alternatives including inferior ones. Moreover it fails Weak Supermodularity. However, as shown by Claim 2, in the two alternative version the the-showiest-is-the-best feature holds under mild assumptions on the cost function.

8 Concluding remarks

While in many models of choice with errors (such as the logit) the choice frequency of an alternative is an indicator of its quality, this is not necessarily the case when the source of error is imperfect attention. We have shown however that, when salience is set through a competitive process rather than being a given, there are plausible circumstances in which the showiest (and most chosen) is also the best.

This result is especially interesting in light of some recent laboratory evidence suggesting that, when strategic aspects are absent, there is a lack of correlation between alternatives' quality and the probability of being noticed (Reutskaya et al. [16], Krajbich and Rangel [9]). In our previous work [12] on individual decision making we also made no assumption of correlation between quality and visibility. Our current result indicates that when strategic factors do operate, they might induce a sharp departure from this baseline. This is particularly relevant in a context such as supermarket choice, which is the environment that Reutskaya et al. [16] replicate experimentally. This study supports the hypothesis of our model that consumers optimise within the consideration set (called there the "seen set"). However it also shows that consumers appear to search randomly with respect to product quality. High quality products are not more likely to be considered. Because in real supermarkets producers invest heavily to increase the salience of their products, it is hard to assume that strategic forces do not operate, and thus our model suggests the possibility that the lack of correlation observed in the experiment might not continue to hold in the market, or that it may be due to yet unspecified countervailing factors.

More generally, we have highlighted several features of mechanisms for attracting attention. Roughly speaking we have identified the key variables at work as the existence and the nature of *externalities* in the technology of salience. Increases in own salience may or may not have spillovers on the attention grabbed by rivals. And such spillovers when they exist may or may not be benign, that is, increasing own salience may (i) increase or decrease the attention devoted to others, and (ii) increase or decrease the effectiveness of others' salience for getting noticed. It is only the situation with non-benign spillovers (in either of the two senses (i) and (ii)) in which it *may* happen that the worst alternative strictly attracts the most attention and is the most chosen. Otherwise, equilibrium forces always push the best alternative to attain weakly more salience; its identity is revealed by the frequency of choice; and if no externalities occur then frequencies of choice perfectly correlate with quality.

While endogenising salience, we have considered quality as a fixed characteristic of the alternatives. Also, we have assumed that the chooser responds to salience in an unsophisticated way, rather than actively thinking through the consequences of any information contained in the alternatives' salience. These issues should be addressed in future research.

References

[1] Bagwell, K. (2007) "The Economic Analysis of Advertising", in Mark Armstrong & Robert Porter (eds.), *Handbook of Industrial Organization*, Elsevier, edition 1, volume 3: 1701-1844.

[2] Brady, R. and J. Rehbeck (2014) "Menu-Dependent Stochastic Feasibility", mimeo, University of California at San Diego

[3] Bordalo, P., N. Gennaioli and A. Shleifer (2014) "Competition for Attention", working paper, http://scholar.harvard.edu/shleifer/publications/competition-Attention.

[4] Butters, G.R. (1977) "Equilibrium Distribution of Sales and Advertising Prices", *The Review of Economic Studies* 44(3): 465-91.

[5] de Clippel, G., K. Eliaz and K. Rozen (2014) "Competing for Consumer's Inattention", *Journal of Political Economy*, 122 (6): 1203-34.

[6] Echenique, F., K. Saito and Tserenjigmid (2013) "The Perception-Adjusted Luce Model", discussion paper, California Institute of Technology.

[7] Eliaz, K. and R. Spiegler (2011) "Consideration Sets and Competitive Marketing", *Review of Economic Studies* 78: 235-262.

[8] Eliaz, K. and R. Spiegler (2011) "On the strategic use of attention grabbers", *Theoretical Economics* 6: 127–155.

[9] Krajbich, I. and A. Rangel (2011) "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions", *Proceedings of the National Academy of Sciences*, 108, 13852-13857.

[10] Luce, R. D. (1959) Individual Choice Behavior; a theoretical analysis. Wiley: New York.

[11] Luce, R.D. and Suppes, P. (1965) "Preference, Utility, and Subjective Probability". In R.D. Luce, R.R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology* (Vol. III, pp. 235–06). New York: Wiley.

[12] Manzini, P. and M. Mariotti (2014) "Stochastic choice and consideration sets", *Econometrica* 82 (3): 1153–1176.

[13] Masatlioglu, Y., D. Nakajima and E. Ozbay (2012): "Revealed Attention", *American Economic Review*, 102(5): 2183-2205.

[14] McFadden, D.L. (1974) "Conditional Logit Analysis of Qualitative Choice Behavior", in P. Zarembka (ed.), Frontiers in Econometrics, 105-142, Academic Press: New York.

[15] van Nierop, E., B. Bronnenberg, R. Paap, M. Wedel & P.H. Franses (2010) "Retrieving Unobserved Consideration Sets from Household Panel Data", *Journal of Marketing Research* XLVII: 63–74.

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

[17] Roberts, J. H. and J. M. Lattin (1997) "Consideration: Review of research and prospects for future insights", *Journal of Marketing Research*, 34(3): 406–410.

[18] Siegel, R. (2009) "All-Pay Contests", *Econometrica* 77(1): 71-92.

[19] Shocker, A. D., M. Ben-Akiva, B. Boccara and P. Nedungadi (1991) "Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions", *Marketing Letters*, 2(3): 181–197.

[20] Sovinsky Goeree, M. (2008) "Limited Information and Advertising in the US Personal Computer Industry", *Econometrica* 76: 1017-1074.

[21] Wilson, C. J. (2008) "Consideration Sets and Political Choices: A Heterogeneous Model of Vote Choice and Sub-national Party Strength", *Political Behavior* 30: 161–183.

[22] Wright, P., and F. Barbour (1977) "Phased Decision Strategies: Sequels to an Initial Screening." In: Studies in Management Sciences, Multiple Criteria Decision Making, ed. Martin K. Starr, and Milan Zeleny, 91-109, Amsterdam: North-Holland.

[23] Yellot, J.I., JR. (1977) "The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution", *Journal of Mathematical Psychology*, 15: 109-144.

9 Appendix: Proof of Claim 3

We consider the case of three alternatives and show that the position profile $\sigma^* = (0,0,1)$ is a Nash Equilibrium with the desired property. For a generic profile $\sigma = (\sigma_1, \sigma_2, \sigma_3)$, the choice probabilities are given by

$$z_1(\sigma) = \left(\sigma_1 - \tfrac{\sigma_2+\sigma_3}{2}\right)^2 \alpha_1$$
$$z_2(\sigma) = \left(\sigma_2 - \tfrac{\sigma_1+\sigma_3}{2}\right)^2 \alpha_2 \left(1 - \left(\sigma_1 - \tfrac{\sigma_2+\sigma_3}{2}\right)^2 \alpha_1\right)$$
$$z_3(\sigma) = \left(\sigma_3 - \tfrac{\sigma_1+\sigma_2}{2}\right)^2 \alpha_3 \left(1 - \left(\sigma_1 - \tfrac{\sigma_2+\sigma_3}{2}\right)^2 \alpha_1\right)\left(1 - \left(\sigma_2 - \tfrac{\sigma_1+\sigma_3}{2}\right)^2 \alpha_2\right)$$

so that

$$z_1(0,0,1) = \tfrac{1}{4}\alpha_1$$
$$z_2(0,0,1) = \tfrac{1}{4}\alpha_2\left(1 - \tfrac{1}{4}\alpha_1\right)$$
$$z_3(0,0,1) = \alpha_3\left(1 - \tfrac{1}{4}\alpha_1\right)\left(1 - \tfrac{1}{4}\alpha_2\right)$$

and thus

$$z_3(\sigma^*) > z_2(\sigma^*) > z_1(\sigma^*) \tag{5}$$

for suitable values of α_i, e.g. provided that $\alpha_3 > \min\left\{\frac{\alpha_2}{4-\alpha_2}, \frac{4\alpha_1}{(4-\alpha_1)(4-\alpha_2)}\right\} \in (0,1)$. To check that the above is an equilibrium, observe that first derivatives of the payoff functions with respect to own salience are:

$$\frac{\partial(z_1(\sigma))}{\partial\sigma_1} = (2\sigma_1 - (\sigma_2+\sigma_3))\,\alpha_1$$

$$\frac{\partial(z_2(\sigma))}{\partial\sigma_2} = (2\sigma_2 - (\sigma_1+\sigma_3))\,\alpha_2\left(1 - \left(\sigma_1 - \tfrac{\sigma_2+\sigma_3}{2}\right)^2\alpha_1\right) + \alpha_1\alpha_2\left(\sigma_2 - \tfrac{\sigma_1+\sigma_3}{2}\right)^2\left(\sigma_1 - \tfrac{\sigma_2+\sigma_3}{2}\right)$$

$$\begin{aligned}\frac{\partial(z_3(\sigma))}{\partial\sigma_3} = {} & \alpha_3\left(2\sigma_3 - (\sigma_1+\sigma_2)\right)\left(1 - \left(\sigma_1 - \tfrac{\sigma_2+\sigma_3}{2}\right)^2\alpha_1\right)\left(1 - \left(\sigma_2 - \tfrac{\sigma_1+\sigma_3}{2}\right)^2\alpha_2\right) \\ & +\alpha_1\alpha_3\left(\sigma_3 - \tfrac{\sigma_1+\sigma_2}{2}\right)^2\left(\sigma_1 - \tfrac{\sigma_2+\sigma_3}{2}\right)\left(1 - \left(\sigma_2 - \tfrac{\sigma_1+\sigma_3}{2}\right)^2\alpha_2\right) \\ & +\alpha_2\alpha_3\left(\sigma_3 - \tfrac{\sigma_1+\sigma_2}{2}\right)^2\left(1 - \left(\sigma_1 - \tfrac{\sigma_2+\sigma_3}{2}\right)^2\alpha_1\right)\left(\sigma_2 - \tfrac{\sigma_1+\sigma_3}{2}\right)\end{aligned}$$

It is seen immediately that, with $\frac{\sigma_2+\sigma_3}{2} \in [0,1]$ being a minimum for $z_1(\sigma)$, alternative 1's best reply is a corner solution, i.e. either $\sigma_1 = 1$ (if $\frac{\sigma_2+\sigma_3}{2} \leq \frac{1}{2}$) or $\sigma_1 = 0$ (if $\frac{\sigma_2+\sigma_3}{2} \geq \frac{1}{2}$). In the candidate equilibrium $\frac{\sigma_2+\sigma_3}{2} = \frac{1}{2}$, so that alternative 1 cannot improve on its payoff by switching from $\sigma_1 = 0$ to $\sigma_1 = 1$. Turning now to alternative 2, we see that

$$\left.\frac{\partial\left(z_2(\sigma)\right)}{\partial\sigma_2}\right|_{\substack{\sigma_1=0\\ \sigma_3=1}} = \frac{1}{8}\alpha_2\left(2\sigma_2-1\right)\left(8-4\alpha_1\sigma_2^2-\alpha_1-5\alpha_1\sigma_2\right)$$

with three roots,[12]

$$\begin{aligned} r_1 &= \frac{1}{2} \\ r_2 &= \frac{1}{8\alpha_1}\left(-5\alpha_1+\sqrt{128\alpha_1+9\alpha_1^2}\right) \geq r_1 \\ r_3 &= -\frac{1}{8\alpha_1}\left(5\alpha_1+\sqrt{128\alpha_1+9\alpha_1^2}\right) < 0 \end{aligned}$$

There are two candidate best replies[13] at $\sigma_2 = 0$ and $\sigma_2 = \min\{1, r_2\}$. Letting $\alpha_1 \in \left(0, \frac{4}{5}\right)$ ensures that $\sigma_2 = r_2$ is not a best reply.[14] The choice probabilities corresponding to the two remaining candidate best replies are:

$$\begin{aligned} z_2(0,0,1) &= \frac{1}{4}\alpha_2\left(1-\frac{1}{4}\alpha_1\right) \\ z_2(0,1,1) &= \frac{1}{4}\alpha_2\left(1-\alpha_1\right) < \frac{1}{4}\alpha_2\left(1-\frac{1}{4}\alpha_1\right) \end{aligned}$$

so that, regardless of the size of α_1 and α_2, alternative 2 cannot profitably deviate from σ^*.

Finally consider alternative 3:

$$\left.\frac{\partial z_3(\sigma)}{\partial\sigma_3}\right|_{\substack{\sigma_1=0\\ \sigma_2=0}} = \frac{1}{8}\alpha_3\sigma_3\left(-8\alpha_1\sigma_3^2-8\alpha_2\sigma_3^2+3\alpha_1\alpha_2\sigma_3^4+16\right)$$

---

[12] To check $r_2 \geq r_1$ observe that

$$\frac{1}{8\alpha_1}\left(-5\alpha_1+\sqrt{128\alpha_1+9\alpha_1^2}\right) \geq \frac{1}{2} \Leftrightarrow \frac{128}{\alpha_1 72} \geq 1$$

which holds true always.

[13] This holds because $\left.\frac{\partial(z_2(\sigma))}{\partial\sigma_2}\right|_{\substack{\sigma_1=0\\ \sigma_3=1}} < 0$ if $\sigma_2 \in \left[0, \frac{1}{2}\right)$ and $\left.\frac{\partial(z_2(\sigma))}{\partial\sigma_2}\right|_{\substack{\sigma_1=0\\ \sigma_3=1}} > 0$ if $\sigma_2 \in \left(\frac{1}{2}, \min\{1, r_2\}\right]$.

[14] Observe that $r_2 > 1$ if and only if $\alpha_1 \in \left(0, \frac{4}{5}\right)$.

Of the three distinct roots of the polynomial,[15] one is negative, one is larger than unity and one is $\sigma_3 = 0$, with $\left.\frac{\partial z_3(\sigma)}{\partial \sigma_3}\right|_{\substack{\sigma_1=0\\ \sigma_2=0}} > 0$ for $\sigma_3 \in (0,1)$. It follows that $z_3(0,0,\sigma_3)$ is maximised for $\sigma_3 = 1$, with corresponding choice probability

$$z_3(\sigma^*) = \alpha_3\left(1 - \frac{1}{4}\alpha_1\right)\left(1 - \frac{1}{4}\alpha_2\right) > 0 = z_3(0,0,0)$$

Then $\sigma^* = (0,0,1)$ is a Nash equilibrium in which, provided that $\alpha_1 \in \left(0, \frac{4}{5}\right)$ and $\alpha_3 > \min\left\{\frac{\alpha_2}{4-\alpha_2}, \frac{4\alpha_1}{(4-\alpha_1)(4-\alpha_2)}\right\}$, condition (5) holds, so that the worst alternative has the highest probability of being chosen ■

[15]The roots are 0 and $\pm 1.1547\sqrt{\frac{1}{\alpha_1\alpha_2}\left(\alpha_1 + \alpha_2 - \sqrt{\alpha_1^2 - \alpha_1\alpha_2 + \alpha_2^2}\right)}$, where the non zero roots are double roots.

Queen Mary
University of London

This working paper has been produced by
the School of Economics and Finance at
Queen Mary University of London

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/