LIMITED ATTENTION, COMPETITION AND WELFARE*

By Andreas Hefti†

Abstract

We present a framework that embeds the allocation of limited attention into competition and market theory. Motivated by the vast literature in psychology and marketing, our core model concentrates on the case, where firms have means to influence consumers' attention. We show that the resulting equilibrium predictions differ fundamentally from standard competition theories of perfect or scarce information, such as informative advertising, both at the positive and normative level.

JEL Classification: D01, L11, L13, L15, D43, D62, M31, M37, C72

Keywords: Limited Attention; Salience; Bounded rationality; Imperfect Competition; Product Differentiation; Diversity; Advertising; Contest Theory

---

*My special thanks go to Armin Schmutzler and Josef Falkinger. I wish to thank Andrei Shleifer, Klaus Schmidt, Ran Spiegler, Steve Tadelis, Xavier Gabaix, Mark Schankermann, Larry Samuelson, Joel Sobel, Paul Heidhues, Martin Peitz, Glenn Ellison, Pedro Bordalo, David Laibson, Ernst Fehr, Carmen Astorne-Figari, Susan Athey, Dimitri Taubinsky and Annie Liang for their most valuable comments. Further, I thank participants of seminars at Zurich, Innsbruck, Harvard, MIT, Memphis as well as participants of the EEA 2012, Midwest Theory Meeting 2013, IIOC 2013, the 2014 ESSET Gerzensee and the Copenhagen 2014 meeting on limited attention for many nice discussions.

†Corresponding author: Andreas Hefti, Bluemlisalpstrasse 10, CH-8006 Zurich, econhefti@gmail.com. Author affiliations: Department of Economics, University of Zurich; School of Management and Law, Zurich University of Applied Sciences

1 Introduction

An edition of the Sunday New York Times contains more factual information than was available to a reader in all the written material of the fifteenth century (Davenport and Beck, 2001). A typical supermarket shelve features more than 70 different soaps, and Google search queries, such as "Hawaii holiday", generate more than 100m links in the blink of an eye. The common theme of such observations is that superabundance of information is an omnipresent, perhaps even characteristic phenomenon of the modern, digitized economy. A comprehensive literature in psychology emphasizes the importance of limitations on the ability to perceive multiple stimuli for storing information, planning actions and making decisions (see Pashler (1998) for an overview). In presence of cognitive capacity constraints the "wealth of information" creates a "poverty of attention" (Simon, 1955). That is, by paying attention to certain objects I miss other, possibly superior alternatives.

How should we expect markets, as the central institutions of exchange, to cope with the superabundance of information that results from attention-constrained decision-makers? What causes markets to become information-thick relative to the potential of consumers to absorb the information? How could the presence of scarce attention affect standard predictions of competition theory with respect to strategic pricing, advertising or the prevailing degree of diversity in a market? Can we not just use existing theories of scarce information, such as informative advertising, and treat scarce attention as a special case? Put differently, does the change from scarce to superabundant information impose a new paradigm – requiring a genuine theory of limited attention? If you ask businessmen this last question, you almost surely get a strongly affirmative answer. As Davenport and Beck (2001) put it in their guidebook on prosperous business: "If you want to be successful in the current economy, you've got to be good at getting attention".

In this article we address such constitutive questions by studying a rigorous, yet tractable setting that embeds the allocation of limited attention (LA) into competition theory. In essence, attention as a scarce resource means that agents are forced to form a simplified mental representation of the situation at hand when reaching their decisions. If someone thinks about a certain object this means that, by scarcity, less spare resources are left for other alternatives. Information senders, understanding that choice occurs conditional on the market as perceived, therefore have incentives to influence mental dispositions to their favor. The evidence of psychology and marketing on attention congregate essentially in two stylized facts.[1]

[1] See Hefti and Heinke (2014) for a detailed literature review.

1. People have LA: Given vast information only a *subset* of all information items (e.g. physical products, links,...) is considered.

2. With vast information, the *relative salience* of an information item (e.g. shelf-level placement, display ads, on-screen ranking) influences its *chance* of perception.

We propose a simple but powerful attention framework, consisting of two natural building blocks: Capacity constraints R_i and an attention allocation rule P_i. Capacity constraints quantify the information load someone can handle. The psychologically feasible set consists of all possible choice sets that could be perceived, subject to capacity constraints. This set is singleton if and only if attention is not scarce (the capacity constraint is not binding). With scarce attention a non-trivial allocation problem exists: which part of the perceivable world is perceived, and which inhibited from perception. In our framework, the attention allocation rule P_i, technically a probability law over the psychologically feasible set, determines the actual choice set (the "perceived world").

The baseline model incorporates the two empirical facts from above into our simple attention framework in the context of a market economy, and studies the resulting strategic game between the different information-senders (firms). Consumers are boundedly rational in the sense that they make optimal choices conditional on what they perceive, but their perception is limited, and influenced by the *relative stimuli* of the attention-seeking firms, where those firms to whose messages a consumer's sensorium is most exposed have the highest chance of being on a consumer's mind. The resulting compound competition for consumer attention and budget has a natural interpretation in terms of a multi-prize contest: Being perceived corresponds to seizing one of R_i prizes, where the value of a prize to a firm (expected earnings from consideration), the number of competitors, the number of prizes and the efforts to seize a prize are endogenously determined in the strategic equilibrium.

One of our main objectives is to study how a theory of LA and competition might differ from standard competition theory with perfect or scarce information but unbounded capacity constraints. We show that a much richer set of possible equilibrium patterns (pricing, attention-seeking, market structure) than predicted by standard theory results if attention becomes a scarce resource. Our setting unravels that the potential of a firm to influence the consumers' mental model of the market rests on two central aspects of attention: i) consumers' capacity constraints R_i (attentiveness) and ii) how much the attention process can be biased by firms (responsiveness). The high tractability of our setting allows for a detailed analysis of how the peculiarities of attentiveness and responsiveness in interaction with traditional choice or technological fundamentals (preferences, costs) determine the equilibrium patterns, resulting in sharp, context-specific predictions.

A standard prediction of competition theory is that limited information on the consumer side gives market power to firms, which leads to dispersed price equilibria in case of homogeneous products (e.g. Butters (1977); Stahl (1989)), or to markups reflecting scarce information alongside with heterogeneous preferences (Grossman and Shapiro, 1984). The common insight of these models is that the resulting market imperfections vanish as technology improves, e.g., when switching from brick-and-mortar to online shops, or because digitization increases the number of potential customers. As markets become more competitive due to increased consumer-side information, consumer welfare increases, and the natural limit as setup costs approach zero is the case of perfect competition with marginal-cost pricing. By taking our setting to the standard circular model of differentiated products (Salop (1979)) we show that *none* of these conclusions hold if attention is the scarce resource – which it naturally becomes as information in the economy approaches infinity. In fact, LA bounds equilibrium prices away from marginal costs despite massive firm entry, as firms realize that only markets as perceived by consumers are relevant for their pricing strategies. Moreover, with all firms shouting equally loud for attention in the unique strategic equilibrium, consumers increasingly fail to find suitable products as more products are offered – despite better products actually being available! That is, consumers get lost in diversity, and pay non-competitive prices not reflecting the actual market size (a traditional measure of competition). This inversion of the positive diversity-utility relation predicted by standard theory has its roots in the *inhibitive* nature of mental capacity constraints: Perceiving some items leaves less spare resources for other items. With equalized perception chances, the generic feature of the symmetric price-attention equilibrium in the game, this implies that consumers on average fail to find better alternatives as the number of items increases, because there are then more ways of doing relatively worse. We show that the main normative result – under LA the standard pro-competitive, consumer welfare enhancing effect of increased diversity (firm entry) ceases to exist – extends to the case of a general random utility model.

In an extension we examine how the equilibrium prediction changes with "more rational" consumers, where attentiveness is a part of the consumer choice problem. A striking finding is that the negative diversity-welfare relation still applies, and may even be intensified. The reason is that consumers who rationally choose their attentiveness are aware of their own cognitive imperfection. As consumers realize that it becomes increasingly harder to find suitable products, it is individually optimal to "look less carefully". If many consumers respond in such a way, this affects average perception, which makes *higher* equilibrium prices sustainable. This causes a negative diversity-attentiveness spiral: Higher profits attract more firms, the increased diversity makes rational consumers more inattentive, which increases prices further, and so on. It follows that lower setup costs or increased market audience – two alterations compatible with technological progress and digitization – induce a surge in diversity, but with lower consumer welfare caused by higher prices *and* an increased product mismatch. This heavily disagrees with standard pro-competitive predictions from competitive models with perfect or scarce information, but unbounded attention. We find a state-dependence of the indirect externality consumers (heterogeneous tastes) impose on each other: As average attention converts from abundant to scarce, the consumer market externality changes from positive to negative. This negative consumer externality also is important if firms try to influence the allocation of attention and the cognitive load (complexity) of their messages: If firms have the ability to manipulate attentiveness – by resorting to obfuscation techniques – then a larger market audience or lower information costs imply higher equilibrium markups, again contrary to standard wisdom.

In a further extension we analysis an abstract model of imperfect competition, including non-discrete choice. We use this generalized setting to challenge some of the predictions obtained from discrete choice. In a first step, we investigate the possible pricing patterns in a market with limited, but potentially heterogeneous attention. If market diversity is so rich that even the most attentive consumers cannot perceive the entire market, prices are completely rigid against further firm entry. In the intermediate case, where some consumers perceive only a submarket but other consumers perceive the entire market, prices may decrease or increase as additional firms (products) become available. The unconventional case of entry leading to higher prices results, e.g., if the fully attentive consumers also have comparably strong brand preferences.

In a second step, we explore how the equilibrium market structure – the diversity of alternatives – and attention efforts depend, in general, on attention, preferences, and technology. We show that the contest-nature of attention-seeking generates a non-monotonic pattern of attention (marketing) expenditure as a function of diversity, which contradicts the standard monotonic advertising patterns suggested by the theory of informative advertising. With LA, the aggregate "noise" level in the market is decisive for the individual choice of attention activities, attention-seeking being of an offensive nature at relatively low volume, but defensive at high levels, because it then becomes much harder to "being heard". The discrete choice model unambiguously predicts a negative relation between attentiveness and firms profits (or market entry in the long run). It turns out that, in general, this is not the only possible prediction of an attention theory. The equilibrium market structure depends on preferences, attentiveness and firms' ability to attract attention. The stronger substitutes the products are, the fiercer economic competition becomes, and the more firms benefit if consumers become less attentive.[2] If at the same time it is hard for firms to influence the attention allocation to their favor (e.g. consumer attention is highly inertial), market rents from inattention are note competed away by market forces, which is the generic case in the discrete choice model. This means that, contrary to standard wisdom, markets with weakly differentiated tastes may actually generate a huge amount of diversity as a pure consequence of consumer inattention. Vice-versa, markets may become thin despite strongly differentiated tastes, if products are mild substitutes and consumers are inattentive but highly responsive to firms' efforts to attract attention. The observation that LA possibly detaches market structure from preferences implies that ignoring the possibility of LA might lead to biased conclusions when evaluating data e.g. concerning brand-loyalty or the so-called Long-Tail effect.

Related literature Psychology views attention, the main gating mechanism of cognition, as working either "goal-driven, controlled in a top-down fashion, or stimulus-driven, controlled in a bottom-up fashion" (Yantis (1998)). While both types matter for explaining individual performance during visual search experiments, stimulus-driven attentional control frequently is regarded as faster and more potent (see Pashler (1998)). It is perhaps not surprising that economics, centered around the benchmark notion of a knowledgeable, optimizing decision-maker, has mainly pursued the goal-driven perspective (e.g. Sims (2003), Reis (2006), Gabaix et al. (2006)), where the allocation of attention is an active part of the choice problem, and under full control of the decision-maker. Similarly, in models of consumer search (e.g. Stahl (1989)), consumers rationally decide themselves when to stop their search. While our attention framework includes goal-driven attention allocation as a special case,[3] we pursue the stimulus-driven approach in this article, following the vast evidence from marketing science and psychology. Further, our attention model also features limited sampling, but one important difference to search models is that sampling chances are not exogenous, but endogenously determined by strategic considerations of the information-senders.

The idea that consumers base their decisions on a subset of all possible alternatives (the evoked set or consideration set) has a tradition in marketing, dating back to Howard and Sheth (1969).[4] Marketing studies concentrate on estimating consumer response behavior with respect to ad-

[2]In fact, rational consumers may well be less attentive, once they realize that products are very similar.

[3]See appendix B for an illustration.

[4]Hauser and Wernerfelt (1990) provide an overview of such marketing studies, and emphasize that these sets usually are small. For example, the median shopper considers about 4 shampoos out of more than 30, or $2-5$ autos out of more than 160. Notably, the average consideration set size seems not to hinge on the size of the grand set and varies only little between different classes (Cars, Food, Soaps...).

vertising or salience levels, and care less about the strategic equilibrium and the corresponding positive or normative predictions, which is the focal point of our article.

Other contributions have considered competition with boundedly rational agents. For example, Eliaz and Spiegler (2011a,b) consider the case, where homogeneous duopolists compete for market shares using various marketing or product design techniques. In their models, consumers are initially randomly assigned to a default firm, but full equilibrium consideration is always a possibility (in fact, this makes their setting interesting). In contrast, we study the equilibrium causes and consequences of the interaction of preferences (differentiated products), capacity constraints and (information) technology. When does attention become a scarce resource, and what are the resulting pricing and attention patterns, or the prevailing degree of product diversity? Incorporating LA into competition theory naturally requires us to leave a fixed duopoly and consider the case of an endogenous market structure. Our setting shows that the survivor of attention-seekers (profitability) depends non-trivially on the psychological and preferential constellation in the market, which is a major difference to the homogeneous goods case, where market forces drive down profits to the rational consumer benchmark.

A number of other papers in economics have dealt with consumer capacity constraints. Anderson and De Palma (2009) consider a model, where consumers draw a random sample of firms, which are described by an exogenous profit and surplus distribution. Their model neither features active competition for attention nor economic competition, whereas our core model builds around the interdependence of the two.[5] Falkinger (2008) considers a deterministic model with capacity-constrained consumers, where non-strategic firms need to pay an endogenous fixed cost to assure, with certainty, their perception. The zero-measure setting of his paper makes it less appealing to competition theory, because essential interactions between LA, attention-seeking, pricing strategies, profits and market structure are absent. Finally, a different literature considers how salient attributes of a product may influence its evaluation (Bordalo et al. (2013); Köszegi and Szeidl (2013)). A central difference to these studies is, that they focus on how relative differences in product attributes lead to different decision weights within an *exogenously* given choice set, while we study how firms' abilities to attract attention influences the formation of the choice set itself, and the resulting market equilibrium.

Article structure In section 2 we present the general attention framework, and show how to embed the allocation of LA into competition theory. While this exposition is more general than

[5]In an extension, Anderson and De Palma (2012) consider the case of a multi-sector model with capacity-constrained consumers. There is no competition for attention among firms, nor competition between sectors (products are independent) in that model.

the core model of stimulus-driven attention that we analyze later, it provides a helpful way of organizing thoughts on the subject. Section 3 analyzes our setting in the context of discrete choice, and in section 4 we proceed to the case of an abstract model of imperfect competition. In order to maintain a swift, legible character, we have postponed the rigorous analysis of our model to the main appendix.[6]

2 An economic model of limited attention

We now present a framework that integrates the allocation of LA into competition theory. Proofs of formal claims are in appendix B.

2.1 The allocation of LA

There is a mass Δ of consumers, distributed over $[0, \Delta]$ according to the measure μ, where $\int_{[0,\Delta]} \mu(di) = 1$.[7] X is a finite set of n information items, and $X_i \subset X$ are those items available to consumer i. An item $j \in X$ is available to i, if j is part of i's information technology, e.g. an advertised product in a newspaper i regularly reads, or a Google-indexed web page related to some search query i might use. We formalize attention as the allocation of bounded mental capacities over X_i. This logically requires the specification of two components: i) A *quantification* of information, making the notion of a bound precise, and ii) an *allocation rule*, characterizing the attentional selection process.

Attention sets For any nonempty $A \subset X_i$ let $r_i(A) > 0$ quantify the mental resources required to *perceive* ("decode") this collection of items.[8] By saying "A is perceived" we mean that all items in A are perfectly recognized or, in psychological terms, passed on to the recognition network (Pashler (1998)). That is, perception of a multi-dimensional item is a binary event in our setting. For example, if j is a product with the attributes "price" (p_j) and "type" (b_j), perceiving j means that these attributes are learned.

We assume that perceiving any nonempty $A \subset X_i$ requires some minimal attention, which can be normalized to unity: $A \neq \emptyset \Leftrightarrow r_i(A) \geq 1$, and set $n_i \equiv r_i(X_i)$. LA means that only item sets up to a measure of $R_i \in [1, \infty)$ are perceived, where R_i is i's level of *attentiveness*. The

[6] We have moved further interesting, but technically more involved results, such as the inexistence of asymmetric attention-pricing equilibria under reasonable conditions, to the online appendix.

[7] Formally, $([0, \Delta], \mathcal{B}, \mu)$ is a measure space, where $\mathcal{B}$ is the Borel field and $\mu([0, \Delta]) = 1$. This gives us one compact notation encompassing "atomistic" consumers ($\mu(i) = 1/\Delta$ for $i \in \{1, ..., \Delta\}$), as well as the "continuous" case (where μ is the Lebesgue-measure with density $g(i) = 1/\Delta$ on $[0, \Delta]$).

[8] Formally, r_i is a measure on $(X_i, \mathcal{P}_i)$, where $\mathcal{P}_i$ is the power set of X_i.

psychologically feasible set $\mathcal{A}_i(r_i, R_i) \equiv \mathcal{A}_i$ is the set of all subsets in X_i that maximally exhaust mental resources: $A \in \mathcal{A}_i$ iff $r_i(A) \leq R_i$ and $r_i(A') > R_i$ whenever $A' \supsetneq A$. The elements $A_i \in \mathcal{A}_i$ are the perceivable item sets, accordingly we call $A_i \in \mathcal{A}_i$ an *attention set.*

In the main part of our analysis we consider symmetric information items, in such that $r(j) = r(j'), \forall j, j' \in X_i$. Then, without loss of generality, r_i can be replaced by the count measure, noting that differences in perceptive abilities r_i can be encoded in R_i.[9] One reason why we introduced the more general specification (r_i, R_i) is to allow for endogenous product complexity. A complex product typically consists of many attributes, which might impede its evaluation relative to other alternatives. In section 4.5 we show that product complexity, as a strategic choice of firms, can be encoded in terms of the information measure r_i, which implies an endogenous, firm-influenced level of effective attentiveness in a market.

Attention-demand profiles If X_i involves a processing capacity $n_i \leq R_i$, i has sufficient spare capacity to consider all available items, hence $\mathcal{A}_i = X_i$. If $n_i > R_i$, then i is overloaded with information, thus $r_i(A_i) \leq R_i, \forall A_i \in \mathcal{A}_i$. Then attention, as a gating mechanism, needs to select which items are perceived, and which drop out. We capture the general gate-keeping role of attention in terms of a probabilistic account. Specifically, let P_i be a probability law with support $\mathcal{A}_i$. P_i formalizes i's *attention allocation*, where $P_i(A)$ is the probability (frequency), with which i observes attention set A_i. Perfect perception is the special case of non-overwhelmed receivers: X_i is perceived iff attention is not scarce, i.e. $P(\mathcal{A}_i) = 1$ iff $R_i \geq n_i$. It should be noted that, at this stage, P_i is merely a descriptive primitive of the attention process. Goal- or stimulus-driven attention amount to putting more structure on the allocation rule P_i.

Given individual profiles $(\mathcal{A}_i, P_i)$, aggregate attention can be found by setting $P_i(A) = 0 \ \forall A \notin \mathcal{A}_i$. Then, $\tilde{\mathcal{A}} \equiv \bigcup\limits_i \mathcal{A}_i$ is the aggregate support of attention and, for $A \in \tilde{\mathcal{A}}$, $\tilde{P}(A) \equiv E_i[P_i(A)] = \int P_i(A)\mu(di)$, gives a probability distribution $\tilde{P}$ on $\tilde{\mathcal{A}}$, where $\tilde{P}(A)$ is the fraction of consumers perceiving A. Other measures can be derived, which are of importance later. For example, if $\pi_{ij} = \sum\limits_{A \in \mathcal{A}_i} P_i(j \in A)$ is consumer i's probability to perceive alternative j, the fraction of consumers perceiving j is $\pi_j = \int \pi_{ij}\mu(di)$.

In the following we think of the information items $j \in X$ as products. Let $|X| = n$ and $p \in \mathbb{R}^n_+$ be the price vector of all existing products. We connect the allocation of attention over products with choice theory by specifying an appropriate demand function, the (random) mapping $d^i : \mathcal{A}_i \times \mathbb{R}^n_+ \to \mathbb{R}^n_+$. Hence $d_i(A, p)$ is i's demand vector given attention set A

[9] If r_i is the count measure and $R_i \in \mathbb{N}$ with $R_i \leq n_i$, then every $A \in \mathcal{A}_i(R_i)$ contains exactly R_i items, and $\mathcal{A}_i(R_i) = r_i^{-1}(R_i)$.

and price p, and $d^i_j(A,p)$ is i's demand for product j. Formally the *attention-demand profile* $(\mathcal{A}_i, P_i, d^i)$ ties together the cognitive and economic aspects of consumer choice. If $(\mathcal{A}_i, P_i, d^i)$ is known, one may derive (expected) consumer demand $D^i(p) = \sum_{A\in\mathcal{A}_i} P_i(A) d^i(A,p)$, or average per-consumer market demand $D(p) = \int D^i(p)\mu(di)$, where $D_j(p)$, the j-th component of D, is average per-consumer demand for product j. Sections 3 and 4 contain microfoundations for such attention-demand profiles, where d^i will be derived from constraint utility maximization, and the attention law P_i follows a bottom-up attention process.

2.2 The attention seeker's problem in a market economy

We now adopt our setting to the case, where firms compete for consumer attention and budget. We set $X_i = X$, where X consists of n goods produced by n single-product firms.[10] Further, r_i is the count measure, and $r_i(X) = n$. Firm j's value of attracting consumer i's attention is described by a function $V_i^j : \mathcal{A}_i \times \mathbb{R}^n_+$, $(A,p) \mapsto \mathbb{R}$, with $V_i^j(A,p) = 0$ if $j \notin A$. In the market context, V_i^j is j's (expected) profit from selling to i: $V_i^j(A,p) = (p_j - c_j) d^i_j(A,p)$, where c_j is j's unit cost of production. In the following we generally allow for consumer-side heterogeneity in preferences, budgets or attention capacities, but assume that firms cannot perfectly identify consumer types. One interpretation of our model is that firms have imperfect targeting abilities, allowing them to sort out consumers without any interest in their product, but cannot learn in greater detail the preferences of a consumer with interest. In such a case we could think of the market as consisting only of targeted consumers, who revealed their baseline interest, e.g. by entering a certain type of keyword in a web search engine.[11] Given $(\mathcal{A}_i, P_i)$ and V_i^j, (expected) firm revenue is:

$$E(V^j) = \int \sum_{A\in B_{ij}} P_i(A) V_i^j(A,p)\mu(di)\Delta \qquad B_{ij} = \{A \in \mathcal{A}_i : j \in A\} \tag{1}$$

Separability Generally, a firm's attention rent depends on its ability to influence attention-demand profiles, and (1) summarizes the various theoretical possibilities. Pricing decisions directly affect the demand for perceived items, but could also influence perception by affecting the psychologically feasible set (if R_i depends on p) or the attention allocation (if P_i depends

[10] If $X_i \subsetneq X$ and X_i is "small" relative to X, i faces the problem of scarce information rather than scare attention. Setting $X_i = X$ appears natural in a digitized economy, where availability of information is not a problem. It also is natural for mature product categories. In a separate article we suppose that X_i is endogenously determined by firms' choice how far to spread out its messages, and discuss the resulting trade-offs between information-density (competition for attention) and information-spread (reaching new consumers).

[11] We explore the exact consequences of LA if firms have the ability to target their messages with possibly infinitesimal precision in a separate paper.

on p). According to the theory of persuasive advertising firms can use marketing devices to directly manipulate the willingness to pay, i.e. consumer preferences (see e.g. Bagwell (2007)). In our framework persuasive advertising would mean that conditional demand d^i depends on advertising activities.

Our core analysis concentrates on the case, where competition for attention and budget are *separable*: The *function* d^i is independent of attention activities, and P_i is independent of prices. Note that "unconditional" demand (observed choice behavior) is not generally independent of attention activities despite separability. In fact, this has been one central motivation for recent research on choice theory (Masatlioglu et al. (2012)). We will return to this important point, when we discuss the positive and normative consequences of LA for market equilibria. We impose separability because of two reasons.[12] First, separability gives a clear-cut form of bounded rationality: Consumers make stable choices conditional on what they perceive. Their perception depends on what is brought to their attention. This endows us with a tractable baseline structure, allowing for a distinction between the competition for attention and the economic competition within attention sets ("the market as represented in a consumer's head"). Despite separability, price and attention competition will be *interdependent*: How intensely firms compete for attention depends on extractable attention rents and, conversely, pricing depends on consumer attentiveness. Second, many examples suggest separability to be adequate. Separability is consistent with sponsored-link advertising e.g. as used by Yahoo! (Edelman et al. (2007)), because relatively higher per-click bids increase the chance of obtaining a higher-ranked on-screen position and, given the pertinence of top-list clicking behavior, a higher chance of obtaining a click (but not necessarily a sale) independent of product prices. Many online studies find that rank (salience) predominates relevancy, for example in case of information search tasks (Pan et al. (2007)) or research paper downloads (Novarese and Wilson (2012)). Newspaper advertisements provide an intuitive non-internet example: Acquiring a better placement (e.g. on the front-page) or a larger, or graphically more salient ad may increase the chance of catching a reader's eye relative to smaller or weaker placed ads - independent of the advertised product price itself.[13]

[12] As we will see later, obfuscation strategies or endogenous consumer attentiveness are important special cases that may violate separability.

[13] The marketing science has accumulated vast evidence, starting with Dickson and Sawyer (1990), showing that shoppers frequently have only limited price awareness, or that consumers purchase a low quality brand "simply because they fail to notice the prices of the other brands", but at the same time the product is "the most feature-advertised and displayed brand" (Mehta et al. (2003)). Also see Van Nierop et al. (2010) for scanner panel data evidence suggesting that, in low-involvement categories, prices indeed are a competitive instrument, but only *conditional* on consideration, whereas display ads and shelf level placement are highly relevant determinants of consideration.

Firm payoffs We assume that firms may use a variety of devices ("stimuli") to influence the allocation of attention P_i in their favor. There is strong evidence for positional effects, e.g. on Internet search pages, where outstanding links tend to grab relatively more attention (clicks).[14] Using web traffic data Dreze and Zufryden (2004) show that a site can increase its visibility by altering its linkage in the web, thereby also improving search engine indexing. Similar findings are documented with sponsored search advertising (Ghose and Yang (2009)), PDA's (Baye et al. (2009)), computer memory modules (Ellison and Ellison (2009)) or online bookstores (Smith and Brynjolfsson (2001)). Various forms of salience or "reminder" advertising, such as feature or display ads in supermarkets, constitute further examples.

We also assume that attention-seeking has no direct effect on attentiveness R_i. This is reasonable for low involvement products, where marketing evidence indicates that consideration set *sizes* vary only little among consumers (Hauser and Wernerfelt (1990); Mehta et al. (2003)), and appear to be resilient against advertising (Mitra (1995)), while consideration itself depends on situational or marketing factors (shelf placement, display ads, package design). The assumption of an exogenous R_i is relaxed in section 3.3 (rational attentiveness) and section 4.5 (obfuscation).

Let $f_j \in \mathbb{R}_+$ quantify firm j's efforts to attract attention; f_{-j} is the vector of its competitors' efforts. Firms need to pay a variable attention costs $C(f_j)$ and a fixed setup cost $F > 0$, summarizing the sunk costs of production and attention technology. The cost function satisfies $C(0), C'(0) = 0$ and $C'(f), C''(f) > 0$ for $f > 0$.[15] A useful example is given by

$$C(f) = \theta f^{\eta} \qquad \theta > 0 \quad \eta > 1 \tag{2}$$

Summarizing, firm j's payoff are:

$$\Pi^j = \int \left(\sum_{A \in B_{ij}} P_i \left(A, (f_j, f_{-j})\right) V_i^j \left(A, (p_j, p_{-j})\right) \right) \mu(di)\Delta - F - C(f_j) \tag{3}$$

2.3 The symmetric price-attention game

We analyze the above model as a symmetric game. Symmetry means that while there could be consumer-side heterogeneity in preferences, attention capabilities, budgets or attentional defaults (some consumers are more likely to perceive certain attention sets), such heterogeneity

[14] Novarese and Wilson (2012) find that not only the top positions attract attention, but so do the bottom positions, whereas interior positions are attention-weak, which makes sense from a salience perspective, but less from the viewpoint of rational, sequential search (as e.g. in Athey and Ellison (2011)).

[15] Efforts costs could be the overall costs of a marketing campaign, including e.g. the costs of obtaining a top position in an online sponsored ad setting (search engine marketing), or the costs associated with obtaining a better Internet search rank by increasing the site's linkage in the Web.

does not lead to a *systematic* bias in attention or preferences at the aggregate level towards any firm. In any case, symmetry greatly simplifies the analysis and constitutes an important benchmark. Suppose that $1 \leq R_i < \infty$, and let $S_p \equiv [c, \infty)$, $S_f \equiv [0, \infty)$, where $S \equiv S_p \times S_f$ denotes the strategy space for an individual firm. Active firms simultaneously choose their strategy, the pair $(p_j, f_j) \in S$, to maximize (3), taking (p_{-j}, f_{-j}) as given. We concentrate on symmetric equilibria, as this is the generic equilibrium type in symmetric games. In the online appendix we demonstrate that asymmetric equilibria are unlikely to exist. To find symmetric equilibria, we fix an arbitrary firm j (we later drop the j-index), called the *indicative firm*, and consider (3) under the restriction that all opponents choose an identical strategy $(\bar{p}, \bar{f})$. Symmetry implies that i's probability to perceive j is $\pi_{ij} = |B_{ij}| \, P_i \left(A, (f, \bar{f})\right)$ for arbitrary $A \in B_{ij}$. As $|B_{ij}|$, the set of all attention sets containing j, depends on R_i and n we write $\pi_{ij} \equiv \pi^i \left(f, \bar{f}, n, R_i\right)$, where $\pi^i \left(f, \bar{f}, n, R_i\right) = 1$ whenever $R_i \geq n$. Defining $V_i^j(p, \bar{p}, z_i) \equiv \sum\limits_{A \in B_{ij}} \frac{V_i^j(A,(p,\bar{p}))}{|B_{ij}|}$, $z_i = min\{R_i, n\}$, (3) can be restated as

$$\Pi^j \left(p, f\right) = \int \pi^i(f, \bar{f}, n, R_i) V_i^j(p, \bar{p}, z_i) \mu(di) \Delta - F - C(f) \tag{4}$$

(4) exposes how LA affects firm behavior. $V_i^j \left(p, \bar{p}, z_i\right)$ is the conditional attention rent earned from i, given that j is perceived together with $z_i - 1$ competitors, and $\pi^i V_i^j$ is j's expected attention rent. Attentiveness R_i affects *perceived* price competition (V_i^j), but also the chance of perception π^i, because R_i determines $\mathcal{A}_i$, and thus how many competitors j is compared to. Expression (4) reveals a *contest-theory flavor* of the competition for attention. For each i, the firms play a simultaneous R_i-prize contest, where being perceived corresponds to seizing a prize worth of V_i^j. The value of the prize itself is endogenous and depends, inter alia, on strategic pricing considerations. This is a central difference to the stochastics produced by informative advertising models, which ignore the possibility of attention capacity constraints. In these models perception is like an urn with unbounded capacity, but the firms, throwing out their ads, may sometimes miss the urn, and it plays no role how often they actually hit an urn. This setting naturally lead to a Binomial (asymptotically to a Poissonian) law of recognition (Butters (1977); Grossman and Shapiro (1984)). This is fundamentally different from the contest-type of attention competition under capacity constraints, where the consumer draws R_i balls out of a potentially heavy-filled urn (see below for a specific example). As we will see this apparently simple difference has huge implications in a market economy. The empirical relevance of attention rather than pure information concerns is well-documented in marketing, and explains why "companies spend a significant amount of their budget on reminder-advertising", attempting to

"keep a product top-of-mind" (Iyer et al. (2005)).

2.4 Relative salience

It is a well-known fact from perceptual psychology that with binding capacity constraints the *relative* strength of competing stimuli is decisive for which signals are processed on to the recognition network. Visual search experiments typically show that to generate a pop-up or salience effect the motion, color or luminance of an object matters *relatively* to the local and global surrounding of the object (Pashler (1998); Nothdurft (2000)) and, accordingly, attention works as a spotlight (Kahneman and Henik (1981), Maunsell and Treue (2006)). In our model the relative nature of salience is appropriately represented by assuming that the perception function π^i is zero-homogeneous in attention efforts. This means that if all firms double their attention efforts, individual perception chances do not change. In the following π^i_m denotes the partial derivative w.r.t. the m-th argument.

Assumption 1 *If $n > R_i$ and $\bar{f} > 0$ then $\pi^i(f, \bar{f}, n, R_i) \in [0, 1)$. Moreover: i) If $f > 0$, then $\pi^i\left(\lambda f, \lambda \bar{f}, n, R_i\right) = \pi^i\left(f, \bar{f}, n, R_i\right)$ for $\lambda > 0$. ii) π^i is C^2 in $\left(f, \bar{f}, n\right)$, $\pi^i_1 > 0$ and $\pi^i_3, \pi^i_{11} < 0$ whenever $f > 0$. Also, π^i is strictly increasing in R_i and $\pi^i(0, \bar{f}, n, R_i) = 0$.*

Assumption 1 states that a firm can influence its perception chances up to the facts that relative salience matters, and attracting attention is subject to diminishing returns. A technical justification for the latter is that it leads to interior solutions. A psychological justification is that salience of a feature map is of a subadditive nature, i.e. it gets more difficult to generate a pop-up effect the more salience features are already active (Nothdurft (2000)). It is a consequence of relative salience and diminishing returns that $\pi^i_2(f, \bar{f}, n, R_i) < 0$, i.e. attracting attention imposes a negative externality on other attention-seekers (see lemma 1 in appendix A.1). Further, a negative relation between a firm's perception chance and the number of alternatives n, and a positive relation between attentiveness R_i and perception chances both seem natural assumptions.

Attention Contest Function (ACF) What could a reasonable example for π^i, satisfying assumption 1, look like? We now derive a useful example, building on the idea that attracting attention follows a stochastic process similar to making random draws in an urn model. Let $n > R$, where each firm's attention effort is represented as a ball in an urn. A consumer draws R balls from the urn without repetition, which corresponds to selecting R different alternatives. Firm j is perceived iff it is among the R draws. Suppose that obtaining a draw depends on the mass of j's ball (f_j) versus the aggregate mass of all remaining balls (each has mass $\bar{f}$). Accordingly, we get

$$\pi(f, \bar{f}, n, R) = 1 - \prod_{i=1}^{R} \left(1 - \frac{f}{f + (n-i)\bar{f}}\right) \tag{5}$$

We refer to (5) as the *Attention Contest Function*, and show in appendix A.1 that it satisfies assumption 1.[16] For the equilibrium analysis the marginal perception probability $\pi_1(f, \bar{f}, n, R)$, evaluated at $f = \bar{f}$, shall play a central role. With the ACF it takes on a simple form:[17]

$$\pi_1(f, f, n, R) = \frac{n-R}{nf} \sum_{i=1}^{R} \frac{1}{1+n-i} \cong \frac{n-R}{n^2 f} R \tag{6}$$

We show in appendix A.1 that the suggested approximation is excellent especially if n is large compared to R. Moreover, this approximation corresponds *exactly* to the marginal perception chance obtained from an urn model assuming random draws *with* replacement, in which case:

$$\pi\left(f, \bar{f}, n, R\right) = \frac{fR}{fR + \bar{f}(n-R)} \tag{7}$$

(7) also satisfies assumption 1. Moreover, (7) is a special case of a general way of deriving attention allocations P_i from simple sampling chances, which we elaborate in appendix B. Examples (5) and (7) are important to our analysis, because these examples generate clear comparative-static patterns.

3 LA in a discrete choice model

We now apply our attention framework to the well-known circular model of product differentiation (Salop (1979)). We focus on how LA affects equilibrium pricing and product diversity. The patterns of attention-seeking are discussed later within an abstract competitive framework (section 4).

Ideal-Variety model Consumers and firms are located uniformly around the unit circle. The location of a consumer represents his ideal variety (IV); consuming at different locations causes some inconvenience. Consumers have unit ("discrete") demand, and choose optimally conditional on their attention set. A consumer at location i perceiving A solves: $\max_{j \in A} U_i(j) =$

[16] In the special case $R = 1$ (5) collapses to the Tullock contest success function.

[17] See lemma 2, in appendix A.1. While (6) has the inconvenience of not being differentiable in R, there exists a formal generalization of (5) using Gamma function expansions, which delivers a fully differentiable version of the ACF (online appendix C.1).

$V - p_j - tw_j$, where w_j is the smallest arc distance between i and j. The taste parameter t quantifies the disutility incurred by consuming at non-ideal locations. We assume that all consumers acquire some product in equilibrium. One interpretation is that firms have limited targeting abilities, allowing them to identify consumers with a baseline interest in the varieties offered by the market. That is, while firms can identify consumers with a sufficient willingness-to-pay for any of the existing varieties, the precise consumer location remains unknown.

Without LA it is a standard equilibrium result that competition is localized between neighbors, and firms sell only to their prime segments. With LA a firm might sell to non-prime consumers, provided that no superior alternative is perceived. The fact that the joint distribution of preferences (locations) and attention profiles $(\mathcal{A}_i(R_i), P_i)$ determines strategic firm behavior can make the model hard to analyze. A natural simplification is obtained by setting $R_i = R$ and $P_i = P$.[18] This greatly increases tractability, and is not unrealistic even with heterogeneity in attentiveness, because firms might not know the R_i-distribution over consumers, but have a good sense of *average* attentiveness $R \equiv \int R_i \mu(di)$, and of the "typical" attention rule P. With $R_i = R$ and $P_i = P$ we obtain the following special form of the payoff function (4):[19]

$$\Pi = \pi(\cdot) \underbrace{(p-c)\left(\frac{\bar{p}-p}{t} + \frac{1}{z}\right)}_{V(p,\bar{p},z)} \Delta - F - C(f) \qquad , z = \min\{R, n\} \tag{8}$$

3.1 Equilibrium

A symmetric free-entry equilibrium (p, f, n) in this model is characterized by optimal firm behavior (pricing and attention efforts) and the zero-profit condition:

$$\begin{array}{cc} p = c + \frac{t}{z} & z = \min\{R, n\} \\ \pi_1(f, f, n, R)\frac{t\Delta}{z^2} = C'(f) & \pi(f, f, n, R)\frac{t\Delta}{z^2} = F + C(f) \end{array} \tag{9}$$

We refer to a solution of (9) with $n > R$ as an *attention equilibrium* (AE), and call a solution with $n \leq R$ a *conventional equilibrium* (CE). In the following we assume that $\pi(\cdot)$ is given by the ACF or by (7) (both imply similar results). The equilibrium (p, f, n) depends on preferences (t), attentiveness (R), production costs c and the state of *technology* $\{F, C(\cdot), \Delta\}$. We give the measure of consumers Δ with access to the market a technological interpretation, having in mind that more and more consumers can learn about the market as the Internet expands,

[18] One can show that equilibrium prices exactly reflect average attentiveness $E[R(i)]$ if $R(i)$ is iid with support $[R_0, R_1]$ over consumers, provided that $n > R_1$. In fact, if additionally all consumers have the same π-function given by (7), then the overall equilibrium is determined by system (9), where $R = E[R(i)]$.

[19] We derive market demand in appendix B.

independent e.g. of physical location or time.[20] The following theorem shows that a unique equilibrium exists, and characterizes when an AE occurs.

Theorem 1 *Suppose that $t\Delta \geq 4F$. Then (9) has a unique solution (p, f, n), $n \geq 2$. An AE occurs endogenously iff $\frac{t}{R^2} > \frac{F}{\Delta}$, $f > 0$ iff an AE occurs, and n decreases in R in an AE.*

For given R, an AE results if the market is sufficiently profitable, such that there is ample firm entry.[21] This tends to occur with strong preferences (t high), or low per-capita setup costs. Comparing AE to CE three results are noteworthy. First, equilibrium prices behave rigidly with respect to technology iff an AE has occurred: Changes in technology have no effect on the prevailing prices in an AE (as long as we do not switch to a CE). Such changes affect equilibrium prices in a CE, because technology determines the size of the perceived market (n), and thus influences price competition. The paradigm of localized competition inherent in the IV model is present also with LA, but the "localization" takes place in the market *as perceived by consumers* rather than in the effective market: To the average consumer the market looks like a circle with R (rather than n) equidistant firms. It follows that prices never converge to unit costs as $\frac{F}{\Delta} \to 0$, contrary to the standard result with unbounded attention, where this convergence occurs despite product differentiation (Perloff and Salop (1985)). This result is especially noteworthy in view of the fact that one of the main theories about scarce information – informative advertising – yields qualitatively the *same* prediction as a CE about how technology and pricing are related (Grossman and Shapiro (1984); henceforth GS): Improved technology (lower $\frac{F}{\Delta}$), including cheaper ad costs, induces entry, which fosters competition and decreases markups. In contrast, LA protects a firm's demand from the increased downward pressure on prices resulting from firm entry.[22] Second, attention competition is characteristic for an AE, i.e. $f > 0$ iff an AE has occurred. This follows because efforts are costly, and is not confined to discrete choice. We analyze the general comparative-static patterns of attention efforts and expenditures in section 4.3. Third, consumer inattention is an unambiguous source of market power in the IV model. This follows because $n'(R) < 0$ is equivalent to the fact that $\Pi'(R) < 0$ assuming that n is an exogenously fixed number of competitors. That is, the markup firms gain by setting higher prices to inattentive consumers outweighs the costs they need to bear to compete for attention. We show below that $n'(R) < 0$ holds more generally in discrete choice models. Section 4.3

[20] Of course, a change in Δ could have many other non-technological reasons, e.g. international integration by a reduction of trade-barriers between two countries.

[21] To find cut-off parameters one can analyze an auxiliary model with the assumption of unbounded attention capacities. An AE results iff the hypothetical $\tilde{n}$ exceeds R, see appendix A.

[22] We show in section 4.2 that the set of equilibrium pricing patterns can be even richer if R_i is heterogeneous in the population.

shows that the relation between profitability (or diversity) and attentiveness can be richer than captured by discrete choice models, and generally depends on consumer preferences and on how easy it is for firms to change attention to their favor.

Brand-loyalty and attention The equilibrium markup, $p - c = t/R$, or j's price elasticity, $1 + \frac{2c}{(\bar{p}-c)+t/R}$, show that attentiveness works against the preference parameter t. This is a vindication of a general pattern (see section 4.3), showing that attention - a psychological concept - and preferences - an economic concept - are tightly tied together. Empirical studies finding low demand elasticities frequently interpreted those as capturing a preferential brand-loyalty effect. Because consumer inattention can imitate brand-loyalty patterns by lowering demand elasticities, such an interpretation is challenged. A market with diverse products, or a market with little brand switching, need not originate in strong taste heterogeneity (a true "need for differentiation"), but rather in constrained attention.[23] The fact that only an attention-adjusted taste parameter correctly describes demand elasticities provides an alternative and possibly *taste-independent* explanation of the Amazon.com price premium puzzle documented by many authors in online bookstores. Many studies have attributed the observation that market dominators (Amazon and Barnes and Noble) charge higher prices compared to smaller retailers to systematic biases in consumer preferences (e.g. Chevalier and Goolsbee (2003) or De los Santos et al. (2012)). While both articles acknowledge that Amazon.com has perceptional advantages (in terms of online clicks), this is not a part of the econometric model. Our model raises concerns about empirical identification of such preference effects as these could be disguised attention effects. In light of the positive and normative consequences of LA shown below, empirical research on strategic pricing and brand-loyalty might want to seriously account for attention effects. In fact, a recent article in "The Economist" strongly advocated that for brands such as Kellogg's much of its consumer loyalty comes from "stuff they can easily find in shops and in their memory banks" (The Economist, 08.30. 2014, p. 50).

3.2 Lost in diversity?

Consumer welfare depends on prices and average disappointment from consuming non-ideal products. In the symmetric equilibrium, average welfare is $\int U_i \mu(di) = V - p - T$, where $T = t\bar{w}$, and $\bar{w} = \int \bar{w}_i \mu(di)$ is average consumer-firm mismatch. Treating n as a parameter rather than an equilibrium variable, our next result establishes a fundamental difference in the

[23] A related point is that LA reinforces the positive effect of stronger brand preferences on prices, because the competitive effect of entry on prices is absent: Prices increase proportionally in t in any AE, while they increase at a diminishing rate in a CE.

diversity-welfare relation between CE and AE.

Theorem 2 *In the symmetric equilibrium average transportation costs T are:*

$$T = t\bar{w} = t \cdot \begin{cases} \frac{n+1}{2n(R+1)} - \frac{1}{4n} & n > R \\ \frac{1}{4n} & n \leq R \end{cases} \tag{10}$$

Then $T'(n) > 0$ iff $n > R$ and $T'(n) < 0$ iff $n < R$. Thus in a free-entry equilibrium: $T'(\Delta) > 0, T'(F) < 0$ iff $n > R$, and $T'(\Delta) < 0, T'(F) > 0$ iff $n < R$.

A CE features a negative relation between diversity n and transportation costs T, because higher diversity leads to a better firm-consumer match (matching effect). As more diversity also implies lower prices, expected utility depends *positively* on n over two interrelated channels. This result extends to the case of scarce information and informative advertising: In GS an exogenous increase of n improves average consumer-side information, which implies lower prices and a better firm-consumer match similar to a CE. Theorem 2 proves a critical *reversal* of this result if attention, rather than information, is the scarce resource: Then, consumers increasingly fail to find a suitable product as diversity expands, *despite* better varieties becoming available! That is, consumers get lost in diversity.

This diversity-utility reversal is the source of substantial differences between free-entry AE and CE. The following illustration assumes that $C(f)$ is given by (2), because equilibrium diversity n then depends on per capita costs F/Δ,[24] where $n'(F/\Delta) < 0$ holds for both an AE and a CE. In a CE, lower F/Δ implies happier consumers, as firm entry increases diversity and thereby also fosters price competition: Consumers obtain *better* varieties at *lower* prices. This also is true in the scare-information setting of GS. In an AE, lower F/Δ leads to a surge in diversity, but *without* putting downward pressure on prices (which reinforces the expansive effect), and *higher* average product disappointment. Because prices reflect the perceived market, perfect competition (marginal-cost pricing) cannot be the natural limit of a market economy with scarce attention, and figure 1 exemplifies the normative consequences of this fact. Assuming unbounded attention capacities ($R = \infty$), and letting $F/\Delta \downarrow 0$, the standard result that $T \downarrow 0$, $p \downarrow c$ and $U \uparrow V - c$ is obtained. This limit result extends to the scarce-information setting by GS. With LA ($R < \infty$) the attention constraint must become binding as $F/\Delta \downarrow 0$, prices then remain bounded away from marginal costs and average transportation costs remain bounded away from zero, despite the firms crowding the entire circle![25] Moreover, the fact that $sign\, T'(\Delta)$ depends

[24] $n(F, \Delta) = n(F/\Delta)$ already holds if $C(f)$ is given by (2) and π satisfies zero-homogeneity (see the proof of theorem 4).

[25] This result persists if there is heterogeneity in attentiveness R_i but $R_i < \infty\ \forall i$.

Figure 1: Transportation costs and consumer utility in dependence of F/Δ

on whether a CE or an AE results (theorem 2) reveals a state-dependence of the externality between different consumers:[26] Loosely spoken, an increase of the overall consumer budget in the market attracts additional firms, which is beneficial to individual consumers iff attention is not scarce.

Inhibition effect Why does LA *cause* a diversity-utility reversal? First, one should realize that the uniform sampling chances $\pi = R/n$ of each firm are a consequences of the strategic equilibrium (equal attention efforts) and symmetry. Second, with LA consumers do not benefit from lower prices as firms price to the market as perceived by consumers rather than to the effective market. Given the symmetry of the model, it is only average perceived distance between adjacent firms that determines equilibrium pricing, and this distance is independent of diversity in an AE. Third, consumers experience a welfare loss as diversity increases. This is striking, because the standard channel by which consumers benefit from increased variety – the matching effect – also is present with LA, as "better" firms are principally available in the market. The reason why welfare declines in diversity is the existence of a second, dominant effect, which we call the *inhibition effect.* This in its essence ordinal effect says that whenever we add to a finite list of successive integers the next higher integer, then the expected first order statistics of any fixed-size random subsample of R *different* integers must increase. Specifically, if varieties are ordered by $1 \succ 2 \succ ... \succ n$ for some consumer, and $R < n$ products are considered, on average the $\frac{1+n}{1+R}$-favorite product is chosen. This index is increasing in n.[27] The inhibition effect means that consumers choose relatively worse products (in an ordinal sense) as diversity increases. This effect naturally works against the matching effect. By theorem 2, the decrease of the distance between varieties caused by an increase of diversity n (matching effect) does *not*

[26] This argument assumes that additional consumers enter the circle in an iid uniform way.

[27] To see this more intuitively, consider someone who currently consumes his k-th favorite variety. Then, there are $k-1$ superior products, but also $n-k$ inferior products. The chance of consuming an inferior variety is $\frac{n-k}{n}$, which is increasing in n, as there are more ways to do relatively worse if n increases.

compensate the loss in favority (inhibition effect). We further explore the tension between the two effects in section 3.4 in the context of a general random utility model.

Conceptually, the inhibition effect originates from the fact that LA implies an "inverse pigeonhole principle" structure of perception,[28] because being perceived reduces perception chance of others. Such an inhibitory effect of attention is a known phenomenon in psychology. In presence of multiple stimuli attention operates as a gating mechanism by restricting the amount of information that is processed at once. Therefore some signals naturally are inhibited, while other signals are processed on to the recognition network (e.g. Milliken and Tipper (1998)).[29] It is this inhibitory nature of LA that explains the welfare divergence between attention and conventional equilibria (including scarce information equilibria).

3.3 Rational attentiveness

Is the welfare reversal under LA (theorem 2) robust with more rational consumers who also optimally choose their attentiveness R_i? Perhaps surprisingly, the answer to this question is yes, and the main reason is the inhibition effect. Suppose that besides choosing the variety, consumers also optimize on R_i, trading of the expected gains of increased attentiveness against an increasing, convex and twice continuously differentiable attentiveness cost $h(R_i)$.[30] Specifically, suppose that i solves

$$\max_{R_i \in [1,n]} \; E(U_i(R_i)) - h(R_i) \tag{11}$$

where $E(U_i(R_i)) \equiv E[U_i(A) : A \in \mathcal{A}(R_i)]$, and $U_i(A) = \max\limits_{j \in A} V - p_j - tw_j$. This setup is generally compatible with the idea of a phased decision process – a recurring theme in psychology and the marketing science – where consumers first select a decision rule by weighting costs and benefits, and then act according to this rule.[31] Rational consumers anticipate symmetry in prices and attention efforts, and the law of iterated expectations gives:

$$E(U_i(R_i)) = V - p - tE_k \left[E_w \left[w|\, k\right]\right] = V - p - T(R, n)$$

Problem (11) has a unique solution $R_i \in [1, n]$ (see appendix B), and $-tT'(R_i) = h'(R_i)$

[28]In mathematics, the pigeonhole principle states that if n pigeons are to be divided over R holes, and $n > R$, then at least one hole must be occupied by at least two pigeons.

[29]Marketing has recognized inhibition as an important empirical fact. An early example is Alba and Chattopadhyay (1986), who find that perceiving a brand inhibits recall of other brands.

[30]In case of a multi-market model with LA, opportunity costs of increased attentiveness in a single market would be the utility costs of forgone attention in the other markets. Such a model of attentional spillovers would be an interesting extension to this setting.

[31]See Payne (1982), Gensch (1987), Hauser and Wernerfelt (1990), Bronnenberg and Vanhonacker (1996) and many more. Also, Anderson and De Palma (2009) consider a related consumer decision problem in case of an exogenously given surplus distribution.

whenever it is optimal to set $R_i < n$. Obviously, if attention were free, then $R_i = n$. We refer to a solution with $R_i < n$ as a solution with *rational attentiveness* (RA). RA implies $R_i'(t) > 0$, $R_i'(n) < 0$. A consumer has a propensity to be attentive if opportunity costs of inattention are high, which is the case with strong taste effects (high t) or low diversity n. The reason for the latter is the inhibition effect: If a consumer realizes that it becomes increasingly hard to find a suitable variety, it is optimal for him to decrease own attentiveness as diversity increases. If a single individual behaves in this way, there is no impact on equilibrium, but if many consumers adopt such a behavior, the equilibrium will be affected. In a symmetric equilibrium with RA the equilibrium variables (R, p, f, n) are determined by ($z = \min\{R, n\}$):

$$\begin{array}{cc} p = c + \frac{t}{z} & R = \arg\min\limits_{R \in [1,n]} T(R,n) + h(R) \\ \pi_1(f,f,n,R)\frac{t\Delta}{z^2} = C'(f) & \pi(f,f,n,R)\frac{t\Delta}{z^2} = F + C(f) \end{array} \tag{12}$$

For simplicity, we assume that π is given by (7), $C(f) = f^\eta$, $\eta > 1$, and $h'(1) = 0$.[32]

Theorem 3 (RA equilibria) *A unique equilibrium (R, p, f, n) with $R > 1$ exists. An equilibrium with RA occurs iff $\frac{t}{2\tilde{n}(1+\tilde{n})} < h'(\tilde{n})$, where $\tilde{n} = \sqrt{\frac{t\Delta}{F}}$. Any RA equilibrium satisfies $n'\left(\frac{\Delta}{F}\right), p'\left(\frac{\Delta}{F}\right), T'\left(\frac{\Delta}{F}\right) > 0$ as well as $R'\left(\frac{\Delta}{F}\right) < 0$.*

The implications of RA are illustrated in figure 2. In the figure $\tilde{\frac{\Delta}{F}}$ is the regime switching point: Iff $\frac{\Delta}{F} > \tilde{\frac{\Delta}{F}}$ an equilibrium with RA results. The dotted line in the right figure represents a

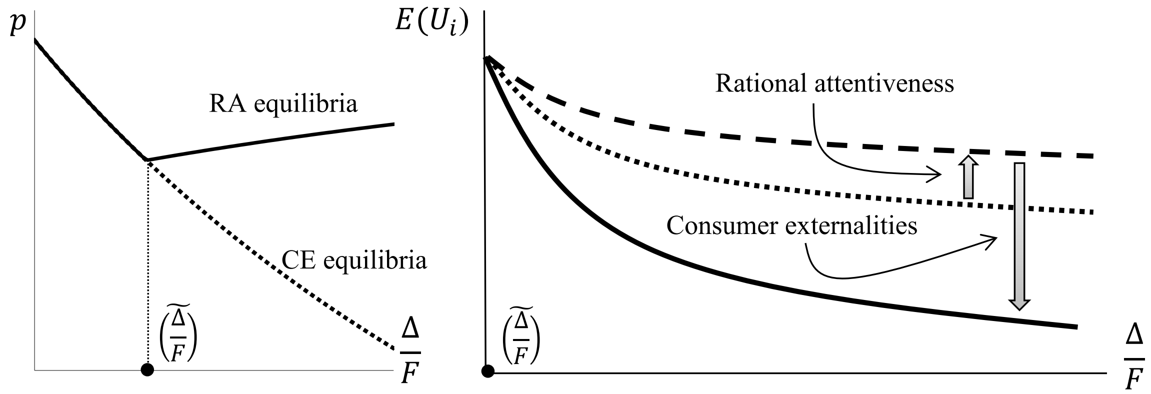

Figure 2: Attentiveness-diversity spiral

consumer's expected utility if his attentiveness were exogenously fixed at $R_i = \tilde{n}$. If $\frac{\Delta}{F} > \tilde{\frac{\Delta}{F}}$ then an individual consumer *reduces* her attentiveness (moving from the dotted to the dashed line in the right figure). While this is individually optimal, the resulting lower *collective* attentiveness

[32]As can be seen in the proof, these assumptions simplify the representation of the result as only per capita setup costs matter, and a unique optimal attentiveness level exists.

enables firms to increase prices, as they face lower demand elasticities. Theorem 3 thus suggests a negative attentiveness-diversity spiral under technological progress: If $d\frac{\Delta}{F} > 0$, then market-level attentiveness decreases but prices and profits increase. This leads to firm entry, which further reduces attentiveness and increases prices and so on, ultimately converging to a new equilibrium (solid line) with *higher* prices, more diversity and *lower* net consumer surplus compared to the hypothetical case of a collectively fixed attentiveness (dotted line). Notably, the negative externality consumers impose on each other in an AE becomes more pronounced with "more rational" consumers, who individually adjust their attentiveness. Overall, theorems 1 - 3 impressively show that standard positive and normative predictions from competition theory are far from being robust predictions with rational but attention-constrained consumers.

3.4 Random utility

We next explore whether the main conclusions from the IV model change if we consider a random utility model instead. Suppose that $U_i(j) = V - p_j - X_{ij}$ is i's net utility of consuming product j. The match value $X_{ij} \geq 0$ is a continuous random variable, say with support $[0, a]$. Realizations of X_{ij} are iid between products and consumers. If all opponents of j play the same strategy $(\bar{p}, \bar{f})$ and $z = \min\{R, n\}$, then the chance of j to sell to i is: $d_j^i(A, p) = \pi_{ij} P(X_{ij} \leq \bar{X} + p_j - \bar{p})$, where $\bar{X} = \min\left\{X_{i1}, ..., X_{i(z-1)}\right\}$. For example, if X_{ij} is uniform on $[0, a]$, then j's chance of selling to i conditional on being perceived can be linearly approximated by $D(p, \bar{p}, z) = \frac{\bar{p}-p}{a} + \frac{1}{z}$ at $p = \bar{p}$, which gives the same demand system as the IV model, and thus offers the same predictions on the positive side.[33] Turning to the normative side (for arbitrary X_{ij}), note that in the random utility model average consumer disappointment $\bar{w}$ is the expected value of the first order-statistics of a random z-sample from $X_i = \{X_{i1}, ..., X_{in}\}$. In the IV model, (10) showed that average consumer disutility $\bar{w}$ depends positively on n and negatively on R in an AE. The latter applies also to the random utility model, as reducing the sampling size increases expected first order-statistics. Similarly, an increase of n always lowers expected disappointment in a CE because there are more firms to sample. The main difference between the IV and the random utility model is that in the latter $\bar{w}$ is independent of n in any AE.[34] This generally holds, because every R-subsample of X_i has the same marginal distribution, and sampling is iid between consumers. Thus the matching and inhibition effects exactly offset each other in the random utility model. It follows that the main welfare prediction from the IV model carries

[33] While the unapproximated demand system is not linear in p, it is possible to show that the same symmetric equilibrium as in the IV model results.

[34] E.g. the uniform distribution implies that $\bar{w} = \frac{1}{1+z}$, $z = \min\{R, n\}$.

over to the random utility model: The standard claim that increased diversity positively affects consumer welfare breaks down if attention is scarce, as both prices and expected disappointment are rigid against n. The IV model even predicted a pattern *reversal* in the diversity-welfare relation. This difference to random utility seems surprising as, with the uniform distribution, the two models imply essentially the same demand system! The crucial difference, responsible for the more dramatic welfare prediction in the IV model, is that the *effective* market X_i is iid between consumers with random utility, but perfectly correlated in the IV model, because a change in diversity affects the underlying market in the same way for all consumers. While the random utility model might be convenient especially in econometric applications, the idiosyncrasy in the underlying market may appear less realistic. While this is not problematic in a CE – both models then predict essentially the same welfare effects – this crucially changes if attention becomes a scarce resource.

4 LA and competition: The general case

In section 3 we studied the effects of LA in the context of a specific discrete choice model, and obtained a clear set of predictions. We now present a comprehensive analysis of the equilibrium patterns implied by LA using an abstract model of product differentiation, which includes the previous discrete choice model and a CES-demand system as special cases. We first focus on the implications of LA for strategic pricing with possibly heterogeneous attentiveness, and then turn to the comparative-statics of the free-entry equilibrium. As before, firms first decide on entry, entrants pay F, and then all entrants simultaneously compete in prices and attention-seeking.[35] We continue to interpret n, the measure of active firms, as equilibrium diversity. We reconsider the diversity-attentiveness relation from section 3, and show that the equilibrium diversity generally depends on preferences, attentiveness and on how much control firms have over the attention process. We also extend our attention framework to the case of firm-level obfuscation.

4.1 Equilibrium conditions

The indicative firm optimizes payoffs in (4). We set $1 \leq \underline{R} \leq R_i \leq \bar{R} < \infty$, and impose the following assumption on V_i^j:

Assumption 2 *For any $z > 1$ and $p > c$ the value function $V_i^j(p, \bar{p}, z)$ is C^2 in $(p, \bar{p})$, (i) $V_i^j(p, p, z)$ is non-increasing in z and (ii) $V_i^j(p, \bar{p}, z)$ is non-decreasing in $\bar{p}$.*

[35]The symmetric equilibrium is the same if active firms first choose attention efforts and then prices.

(i) and (ii) convey basic competition principles. (i) means that, for equal prices, the attention rents never decrease if less alternatives are perceived, and (ii) means that a higher opponent price never decreases attention rents from those consumers, who perceive the indicative firm. These assumptions appear natural in a context, where products are imperfect alternatives (substitutes); standard models of oligopolistic competition, such as the discrete choice model of section 3 have both features. A well-known example of V_i with non-discrete demand is obtained from quadratic preferences:

$$V_i(p, \bar{p}, z_i) = \max\{\frac{(p-c)}{1+(z_i-1)\gamma_i}\left(1 - \frac{1+(z_i-2)\gamma_i}{1-\gamma_i}p + \frac{\gamma_i(z_i-1)}{1-\gamma_i}\bar{p}\right), 0\} \tag{13}$$

where $\gamma_i \in (0,1)$ quantifies the degree of substitutability[36] between perceived commodities and $z_i = min\{R_i, n\}$. (13) satisfies assumption 2 on the relevant domain ($c \in [0,1)$). Similarly, CES-demand[37]

$$V_i(p, \bar{p}, z_i) = (p-c)\frac{v_i p^{-\sigma_i}}{p^{1-\sigma_i} + (z_i-1)\bar{p}^{1-\sigma_i}} \qquad , c > 0 \tag{14}$$

satisfies assumption 2, where $\sigma_i > 1$ is the elasticity of substitution, and $v_i > 0$ is i's budget. A symmetric free entry equilibrium (p, f, n) is a solution to[38]

$$\begin{aligned} &\textstyle\int \pi^i(f, f, n, z_i)\, V_{i1}(p, p, z_i)\, \mu(di) = 0 \\ &\textstyle\int \pi^i_1(f, f, n, z_i)\, V_i(p, p, z_i)\, \mu(di)\Delta = C'(f) \\ &\textstyle\int \pi^i(f, f, n, z_i)\, V_i(p, p, z_i)\, \mu(di)\Delta = F \end{aligned} \tag{15}$$

The first two equations are the first-order conditions pertaining to (4), evaluated at $(\bar{p}, \bar{f}) = (p, f)$. We call a symmetric equilibrium unique, if there is only one vector (p, f, n) that solves (15). As before, we refer to a solution of (15) with $n > \underline{\mathrm{R}}$ as an AE, and call a solution with $n \leq \underline{\mathrm{R}}$ a CE. A first, intuitive consequence of (15) is that with costly attention-seeking positive attention efforts are characteristic for an AE.[39]

4.2 Equilibrium pricing under LA

By (15) the joint distribution of preferences, budgets and attentiveness pins down equilibrium pricing. To analyze the general implications of LA for equilibrium pricing, given possibly hetero-

[36] $\gamma_i \to 1$ means that commodities are perfect substitutes whereas $\gamma_i \to 0$ means that commodities are independent.

[37] More generally, assumption 2 holds under additive separable utility with strictly concave subutility, as long as marginal utility is not too elastic ($-xu''(x)/u'(x) \leq 1$).

[38] We ignore the integer-value problem of n, and assume that (4) has a unique optimizer (p, f) for any given $(\bar{p}, \bar{f})$. The formal details for such an assumption can be found in appendix A.

[39] See proposition 2, appendix A.

geneous attentiveness, we consider the case of a symmetric unique regular price equilibrium,[40] holding diversity n exogenously fixed. As $\pi^i = min\{\frac{R_i}{n}, 1\}$ the allocation rule P_i is irrelevant for prices in a symmetric equilibrium. Hence equilibrium pricing is solely characterized by the first equation of (15), which can be decomposed as

$$\underbrace{\int_{R_i \leq n} \frac{R_i}{n} V_{i1}(p, p, R_i)}_{A} + \underbrace{\int_{R_i > n} V_{i1}(p, p, n)}_{B} = 0 \tag{16}$$

Pricing depends on two components, which can depend in a conflicting way on attentiveness. First, attentiveness R_i determines the probability of being perceived by the various consumers, and thus their (relative) importance to a firm's pricing decision. Because perception chances depend positively on attentiveness R_i, less attentive consumers matter less for pricing decisions (perception effect). Second, a firm's pricing depends on the consumers' individual demand functions conditional on being perceived. Price pressure might be lower from less attentive consumers, as these compare less (comparison effect). In the special case $R_i = R$ only the comparison effect is present, which has previously simplified the analysis. In general, the possible tension between the two effects may cause non-standard equilibrium pricing patterns to emerge. In the following we assume that the marginal value of being on i's mind is always higher, the less i thinks about other alternatives ($V_{i13}(p, p, z) < 0$).[41]

Diversity effects on pricing If $n < \underline{R}$, then $p'(n) < 0$: If perceived and effective market coincide, the conventional pro-competitive effect of market entry results. At the other extreme nobody perceives the entire market if diversity is so rich that a pure AE ($n > \bar{R}$) results. In such a case prices are perfectly rigid ($p'(n) = 0$), as in section 3. What about the "transition", where $n \in (\underline{R}, \bar{R})$? If the attention-constrained consumers display a higher marginal revenue ($A > 0 > B$ in (16)), then the conventional sign pattern results ($p'(n) < 0$). This holds if heterogeneity in V_i is solely rooted in heterogeneous attention capacities. For example, with CES-demand (14) this would mean that $(v_i, \sigma_i) = (v, \sigma)$.[42] If however $B > 0 > A$, then $p'(n) > 0$ is possible. This can happen if the relatively attentive consumers also are the ones with a comparably strong brand preference. Such a situation is depicted in figure 3 (right) for

[40] Formally, this requires that $\int \pi_i \frac{\partial V_{i1}(p,p,z_i)}{\partial p} \mu(di) < 0$ at equilibrium points.

[41] This is verified, e.g., by (13) and by the discrete choice model of section 3.

[42] In the discrete choice model with heterogeneous attentiveness we would require that both preferences (location) and attentiveness are independently and uniformly distributed over the circle.

a two-type version of (13).[43] The intuition for this result is simple, and illustrated in the left figure. If n increases to n', then the marginal revenue from the inattentive types (A-type) rotates

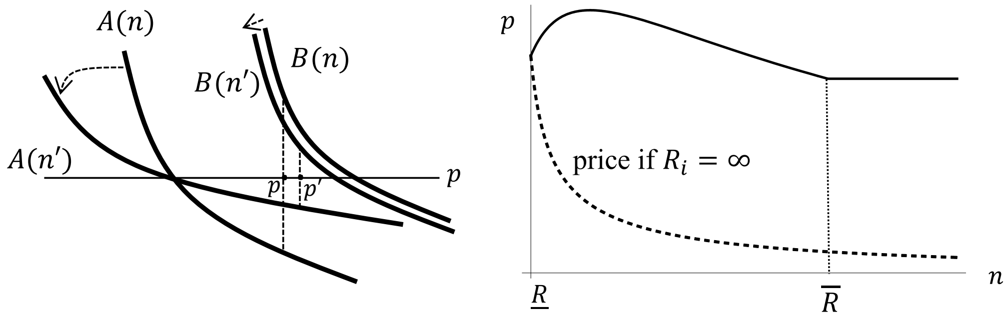

Figure 3: Pricing in a semi-constrained attention economy

to the left, because it becomes less likely that the firm is perceived (perception effect). At the same time marginal revenue from the fully attentive types (B-type) shifts to the left, as these consumers now face more choice options (comparison effect). Optimality requires the pricing firm to balance the two effects. If the shift of marginal revenue of the B-type is sufficiently small, then it is optimal to increase the price as diversity increases.

Finally, the right figure suggests that the hypothetical equilibrium under unbounded attention (dotted line) is a lower bound of LA pricing. A sufficient condition therefore is that $g_i(k) = kV_{i1}(p,p,k)$ is a decreasing function of k, i.e. if the comparison effect dominates the perception effect.[44]

Attentiveness effects on pricing To see how heterogeneity in attentiveness may affect equilibrium pricing in a pure AE, suppose that there are two consumer types a and b. (16) then reduces to to $\alpha R_a V_{a1}(p,p,R_a) + (1-\alpha) R_b V_{b1}(p,p,R_b) = 0$, where α is the fraction of a-types. If attentiveness of the type with a relatively lower marginal revenue *decreases*, equilibrium prices must *increase*, because i) the relative weight of these consumers decrease and ii) their marginal revenue shifts up. If attentiveness of the high value type decreases, the effects are countervailing: ii) implies higher prices, but i) lower prices. A more detailed inspection of examples (13) and (14) suggests dominance of ii), implying that $p'(R_i) < 0$, is the more common case. Moreover, if firms consider *average* attentiveness $R = \alpha R_a + (1-\alpha) R_b$, then $p'(R) < 0$ because i) is absent. In sum, lower attentiveness likely leads to higher prices, and the presence of comparably inattentive consumers usually encumbers more attentive consumers with a negative externality, as they are

[43]This particular figures has $R_1 = 2$, $\gamma_1 = 0.95$ $R_2 = 25$, $\gamma_2 = 0.2$, $c = 0$ and assumes an equal measure of both types.

[44]This condition is satisfied if $R_i = R$, or if V_i is given by (14), but not necessarily if V_i is given by (13). It also is satisfied in the discrete choice model.

forced to pay higher prices.

4.3 Equilibrium patterns under LA

We now turn to the question how firms' abilities to influence both the allocation of scarce attention and the allocation of consumer budget affect the resulting market equilibrium. For simplicity we set $R_i = R$ and $\pi^i = \pi$. Then, setting $V(p, \bar{p}, z) \equiv \int V_i(p, \bar{p}, z)\mu(di)$, (15) reduces to

$$\begin{aligned} &V_1(p, p, z) = 0 \\ &\pi_1(f, f, n, z)\, V(p, p, z)\, \Delta = C'(f) \\ &\pi(f, f, n, z) V(p, p, z)\, \Delta = F + C(f) \end{aligned} \tag{17}$$

The discrete choice model (9) is a particular case of (17), so we might wonder which patterns of section 3 are special to that model, and what the general predictions of competition under LA might be. In the following we concentrate on the equilibrium consequences of LA on pricing, attention-seeking and diversity. A thorough analysis of equilibrium existence and uniqueness is done in appendix A. We spice (17) up by adding a demand shifter ξ to V, where $V_4(p, \bar{p}, R, \xi)$, $V_{14}(p, \bar{p}, R, \xi) > 0$, as well as a cost shifter α to C, where $C_2(f, \alpha), C_{12}(f, \alpha) > 0$. We present a comparative-static analysis for the case, where π is given by the ACF or by (7), assuming that (17) has a unique and regular solution (p, f, n). One can find sufficient conditions such that indeed there is a unique symmetric and regular solution to (17), and these conditions amount to natural regularity and boundary assumptions on $V(\cdot)$ and $\pi(\cdot)$. In particular, (13), (14) and the discrete choice model each imply $V(\cdot)$ that satisfies these assumptions. In order not to impede the exposition, we have postponed all formal details to appendix A.

Theorem 4 *Consider an equilibrium* (p, f, n) *with* $n > R$*. The comparative statics of* (17) *are* $p(\underset{+}{\xi}, \underset{-}{R})$*,* $f = f(\underset{+}{\xi}, \underset{+}{\Delta}, \underset{?}{F}, \underset{?}{\alpha}, \underset{?}{R})$ *and* $n = n(\underset{+}{\xi}, \underset{+}{\Delta}, \underset{-}{F}, \underset{?}{\alpha}, \underset{?}{R})$

In general, the parameters $\{F, \alpha, R\}$ can generate non-monotonic equilibrium patterns of (f, n), which we discuss below. Figure 4 illustrates a particular comparative static pattern, using (7) for π and (14) for V. To understand the comparative-statics on intuitive grounds, it is useful to note the following general rent-expenditure trade-off. An exogenous increase of attention rents or a decrease of attention costs incentivizes firms to increase their attention-seeking. This implies higher attention expenditures which, in equilibrium, could principally offset the initial gain. Whether or not equilibrium forces compete away the initial gains depends non-trivially on the firms' ability to manipulate attention (responsiveness) as well as on consumer willingness to substitute between perceived alternatives, i.e. on the intensity of the economic competition

Figure 4: (p, f, n) as a function of F. The bold dashed lines represent the (hypothetical) unconstrained equilibrium ($R = \infty$).

in the perceived market.[45]

Responsiveness An important insight is that a firm's ability to influence the attention allocation P can be encoded in the attention cost function $C(\cdot)$. To see this, consider attention technology (7), and let $C(f) = f^\eta$, $\eta > 0$. Then, payoff (4) is $\Pi = \frac{fRV(\cdot)\Delta}{fR+(n-R)\bar{f}} - F - f^\eta$. The monotonic transformation $(e, \bar{e}) = (f^\eta, \bar{f}^\eta)$ leads to an equivalent game with payoff $\Pi = \frac{RV(\cdot)\Delta}{R+(n-R)(\bar{e}/e)^{1/\eta}} - F - e$. The elasticity parameter η quantifies how random (noisy) consumer attention is to the firm, i.e. how much control it has over the attention allocation process. Accordingly, we interpret η as capturing consumer *responsiveness* to attention-seeking activities. The larger η, the harder it is for an individual firm to influence its equilibrium perception chances. Intuitively, responsiveness captures how volatile consumer attention allocation is, i.e. how sensitive perception responds to a change of stimulus. For example, if consumers have strong attentional defaults or habits (they always look at the same web sites or regions in a shopping shelve), responsiveness should be low. If $\eta \to 0$ in the above example, then attention allocation becomes quasi deterministic, and our contest model of attention approaches an all-pay auction. Therefore our model identifies responsiveness as a decisive element for whether the quest for attention rather resembles an all-pay auction as suggested by the homogeneous duopoly model of Eliaz and Spiegler (2011a) or a multi-prize contest. In our view, the empirical evidence of a certain fuzziness in individual (and aggregate) clicking behavior (e.g. not all people deterministically follow a top-down clicking strategy) rather advocates the contest setting.

The function $C(f) = f^\eta$ represents the case of *effort-independent responsiveness*, and it provides us with a useful special case to examine how the equilibrium patterns may vary with different levels of consumer responsiveness. It turns out that responsiveness η may be decisive

[45] The comparative statics of n reflect those of Π and vice-versa as $sign\,(\Pi'(\chi)) = sign\,(n'(\chi))$ for $\chi \in \{R, \Delta, F, \alpha, \xi\}$: Whatever increases diversity n in the free-entry equilibrium also increases firm profits for a fixed given n.

for the resulting equilibrium patterns, because it determines how strongly efforts f respond to exogenous shocks, i.e. how "elastic" equilibrium attention competition behaves. To see this note that in an AE the 2nd equation of (17) has the form $\frac{\zeta(n,R)}{f}V(\cdot)\Delta = \eta f^{\eta-1}$ (see lemma 1, appendix A). Hence

$$f = \left(\frac{\tau}{\eta}\right)^{1/\eta} \qquad \tau = (\zeta(n,R)V(\cdot)\Delta)^{1/\eta} \tag{18}$$

showing that $f'(\tau)\tau/f(\tau) = 1/\eta$. Thus a large value of η implies f to react inelastically towards changes of τ.

We now discuss how variations in $\{F, \Delta, \xi, \alpha\}$ affect the equilibrium (p, f, n). The equilibrium effects of more or less attentive consumers are the main topic of section 4.4.

Setup costs (F): Changes in setup-costs (e.g. switching from brick-and-mortar to online shops) have no effects on prices iff attention is scarce, as only then perceived and effective markets differ. We now focus on how F affects attention efforts; the locus $n(F)$ is examined in section 4.4. Figure 4 suggests that f depends on F in a non-monotonic way, which occurs naturally with (7) (or the ACF). The reason therefore is that f depends non-monotonically on n, which in turn originates from the fact that attention efforts are strategic complements at low levels of attention-seeking, but substitutes at high levels.[46] Thus the noise level nf, itself depending on how scarce a resource attention is, qualitatively influences behavior; attention-seeking being rather of an offensive nature at low noise levels, but defensive at high levels.[47]
The way how individual attention efforts depend on firm entry distinguishes the LA prediction on advertising efforts from the classical theory of informative advertising. An exogenous increase in the number of competitors unambiguously decreases firm-level advertising (e.g. Grossman and Shapiro (1984)), because with substitute products advertising efforts are strategic substitutes. The LA model predicts a more complex equilibrium pattern of advertising, which depends, inter alia, on the aggregate advertising volume. This difference has its origin in the distinct stochastics generated by a model with capacity constraints (see section 2.3).
Demand shifters (Δ, ξ): If market demand shifts up, either because the measure of budget sets in the market ($d\Delta > 0$) or the value of individual budget sets to firms ($d\xi > 0$) increase, this augments the market value of attention, hence firms beef up their attention-seeking

[46] It is easy to see that for (7) $sign(f'(\bar{f})) = sign(f'(n))$.

[47] If $n \approx R$ (attention almost not scarce), then attention-seeking incentives are weak since the firm is perceived almost surely. There is only little attention-seeking, and entry ($dF < 0$) causes firms to become louder. If however $n \gg R$, then eventually there is so much noise that further entry discourages individual attention-seeking.

$(f'(\Delta), f'(\xi) > 0)$. Theorem 4 shows that the additional expenditure on attention-seeking caused by such changes in the economic fundamentals is covered by the rents earned from competition. Hence demand shifts rooted in non-attentional consumer characteristics benefit firms for given n, and induce entry in the long run. As we will see below this is not true, in general, for changes in attentiveness or responsiveness.

Attention costs (α): A positive shock to marginal costs decreases, ceteris paribus, attention efforts. As cutting back attention-seeking reduces expenditures, the total effect on profits (diversity) may be ambiguous. Let $\beta \equiv \frac{C_{12}f}{C_2}$, $\mu \equiv \frac{C_{11}f}{C_1} > 0$. Then $n'(\alpha) > 0$ iff equilibrium attention expenditures decrease in α, which is the case iff $(\beta - 1) - \mu > 0$. For equilibrium efforts: $f'(\alpha) < 0$ if $\beta \geq 1$, showing that $n'(\alpha) > 0$ only if $f'(\alpha) < 0$, which is intuitive. In appendix A we prove that $C(f, \alpha) = f^{\alpha}$ unambiguously implies that $n'(\alpha) > 0$. Hence responsiveness and profits are *negatively* related: The *less* influence the firms have on the attention process, the *more* profitable LA becomes in equilibrium. Moreover, $f(\alpha)$ is strictly decreasing, if α enters the cost in some multiplicative way: For $C(f, \alpha) = m(\alpha)c(f)$ we get $\beta = \frac{c'(f)f}{c(f)} > 1$, which implies that $f'(\alpha) < 0$. If also $c(f) = f^{\eta}$, we have $\beta = \eta$ and $\mu = \eta - 1$, implying that $n'(\alpha) = 0$. The only effect of scaling attention costs is a change in equilibrium attention efforts, which exactly offsets the initial change, such that equilibrium attention expenditures remain unchanged.

Discrete choice Because the discrete choice model of section 3.1 is a special case of (17), the general results of theorem 4 with respect to technology parameters $\{F, \Delta\}$ and the taste parameter t $(= \xi)$ also apply to that model. Vice-versa, the facts that technology parameters influence equilibrium pricing iff attention is not scarce, and (f, n) increases in resources Δ and tastes t $(= \xi)$ are robust predictions. We may also conclude that attention efforts will be a hump-shaped function of F also in the discrete choice model.

4.4 Attention and diversity: The general case

A central prediction of the discrete choice model was that a decline of attentiveness in a market is unambiguously profitable for active firms. This in turn was closely linked to the welfare-diversity reversal under LA. In this section we reconsider the attentiveness-diversity relation, and show that $n'(R) < 0$ is not a universal result, but depends in an intuitive and intimate way on attentiveness, responsiveness and consumer preferences.

Let $\hat{V} = \frac{R}{n} V(R)\Delta$ denote expected equilibrium revenue $(V(R) \equiv V(p(R), p(R), R))$. In the following we let $V_3(p, p, R), V_{13}(p, p, R) < 0$ and $V_{11}(p, p, R) + V_{12}(p, p, R) < 0$. These assumptions are consistent with assumption 2 and the conditions for existence and uniqueness of the symmetric equilibrium.[48] Expected revenue is composed of the two components π, V, and both depend on attentiveness in a conflicting way. First, lower attentiveness reduces perception chances π and thus expected revenue. Second, firms can extract more money from perceivers with lower attentiveness, because these have fewer spare mental capacities for other alternatives. More precisely, the second channel involves two components. First, lower attentiveness means that, for given, symmetric prices, firms earn more revenue because consumption expenditure is shared with less competitors ($V_3 < 0$). Second, lower attentiveness triggers strategic pricing effects, as a smaller perceived market makes higher equilibrium prices sustainable ($p'(R) < 0$). These two effects work in the same direction, and it follows that $V'(R) < 0$ in an equilibrium. Hence $\hat{V}$ increases in inattention ($\hat{V}'(R) < 0$) iff the increase in revenues from less attentive consumers compensates for the loss in perception chances. Formally, LA is beneficial to firms iff $-\frac{V'(R)R}{V(R)} > 1$. Here, the story becomes interesting. If LA indeed is beneficial to firms, then scarcer attention makes perception more profitable, which induces firms to invest more resources (costs) to compete for the rare, spare mental capacities. Therefore, the firms' ability to influence attention, i.e. consumer *responsiveness*, and the intensity of price competition jointly determine to what extent competitive forces drive down potential profits from less attentive consumers. The next theorem pins down the exact tension between price and attention competition, and the resulting consequence for equilibrium diversity and attention-seeking.

Theorem 5 *Suppose that* $\varepsilon_V \equiv -\frac{\hat{V}'(R)R}{\hat{V}(R)} > 0$ *and* π *is given by* (7). *Then* $f'(R) < 0$ *and* $sign\,(n'(R)) = sign\left(\frac{R}{n} - \varepsilon_V(\mu + \frac{R}{n})\right)$, *where* $\mu \equiv \frac{C''(f)f}{C'(f)}$. *If* V *is given by* (13) *or* (14), *then* $f'(R) < 0$ *and the stronger substitutes the goods are* (σ, γ *increase*) *the more likely diversity is to increase as attentiveness declines.*

In case of effort-independent responsiveness ($C(f) = \theta f^{\eta}$) we have $\mu = \eta - 1$, and the theorem shows that the competing firms benefit from inattention if inattention is sufficiently profitable, and responsiveness is sufficiently low (large μ). That is, firms are more likely to benefit from inattention (and markets to become thick) if economic competition is *tough* (perceived products are close substitutes), as then a reduction of attentiveness significantly relaxes price competition. Specifically, if a 1% decrease of R increases V by at least 2% (hence $\varepsilon_V \geq 1$), the inattention effect is dominant for *any* $\mu > 0$. This is exactly the case in the discrete choice model, where $\varepsilon_V = 1$.

It follows that whether or not attention rents are competed away by market forces depends

[48]See assumption 3 in appendix A.

crucially on the economic and psychological characteristics of a market. This is a substantial difference to the homogeneous duopoly by Eliaz and Spiegler (2011a), where competitive forces drive down firm profits to the rational consumer benchmark. Theorem 5 shows that an expanding product range and a simultaneous increase in attention activities is characteristic for markets featuring strong price competition and unresponsive consumers if attentiveness declines. Importantly, this is not the unique possible prediction of the attention theory developed in this article. In fact, theorem 5 shows that perceived and effective diversity change in *opposite* ways iff the effect of attentiveness on expected revenues is dominant.

Anderson (2004) formulated the famous prediction that the Internet will make product diversity more profitable than ever. Restricted storage capacities and costly product distribution limit supplied diversity by cost minimization, and mainly popular products are offered, despite the prevalence of "love of variety" in the population. By improving the economics of storage and distribution the expansion of the Internet therefore should imply a surge in diversity – the Long-Tail effect.

Using CES-demand (14) we now demonstrate how LA may change the relationship between tastes, technology and diversity compared to standard theory.[49] Standard models of imperfect competition, ignoring the possibility of LA, account for a positive diversity effect: A decline in setup costs F (or an increase in audience Δ) induces entry, and this effect becomes more pronounced the stronger the taste for variety in the population is.[50] Taking into account LA may entirely change this picture. Markets with attention-constrained consumers may produce a huge product range *without* preferences featuring a strong taste for diversity (this happens in the discrete choice model). Vice-versa, markets might become thin under LA despite a strong love-of-variety at the population level. To see the latter, consider an elasticity parameter σ close to one. Then products are weak substitutes, and standard theory predicts diversity to respond sharply to a change in F. The right panel of figure 5, depicting theorem 5, illustrates that with a small σ it is more likely to enter the $n'(R) > 0$ area, especially with strong responsiveness (low μ). In such a case there are too little gains of attention to compensate for a tough attention competition. Therefore diversity expands less as $dF < 0$ (or $d\Delta > 0$) compared to the standard case without LA. Notably, if consumers are highly responsive, and thus the competition for attention very expensive, the discouragement to enter may even imply that the

[49] The Long-Tail effect not only consists of the diversity effect as discussed here, but also comprises redistribution from hits to niches (Brynjolfsson et al. (2006)). Further, the Long-Tail literature also considers demand-side effects (e.g. recommender systems) but strongly puts forth supply-side effects.

[50] In the CES case: $\frac{n'(F)F}{n(F)} = -\frac{1}{1+F(\sigma-1)}$, where lower σ means a stronger love-of-variety, which implies more local monopoly power for firms.

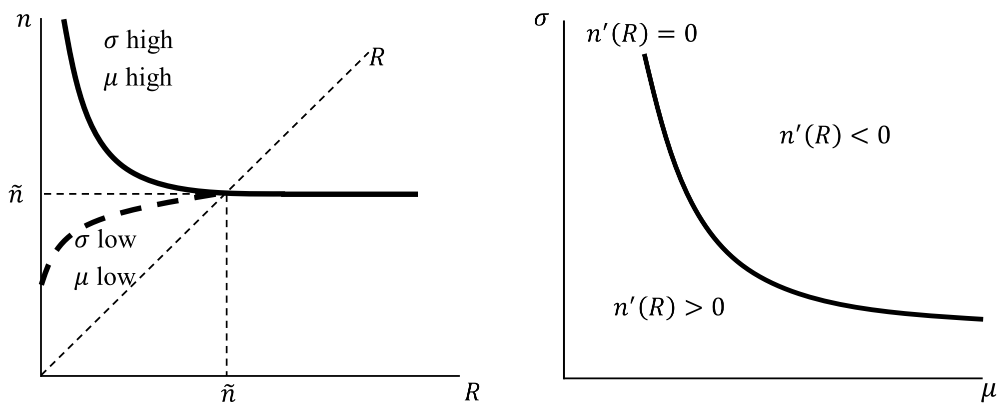

Figure 5: Diversity and attention with CES-demand

standard divergence of diversity in the limit breaks down: Wile $F \to 0$ implies $n \to \infty$ if the possibility of LA is ignored, it may well happen that $\lim_{F \to 0} n(F) < \infty$ with LA and very strong responsiveness ($\mu < 1$).

If products are strong substitutes (σ large), LA thwarts the standard prediction that only little diversity shall prevail in the market, provided that μ is not too low. With strong substitutes competition puts great downward pressure on prices. In such a case LA substantially may shield firms from competitive pressure, which in the long run attracts more competitors to the market.[51] We have seen this to be the unique prediction of the discrete choice model, which is a difference to the case of CES-demand. This is an interesting result in view of a recent empirical challenge to the Long Tail effect in case of the movie or music industry. Elberse (2008) presents evidence suggesting that the Long Tail is weaker than presumed, in part because of a love-of-variety effect: Niche consumers also consume mainstream goods and vice-versa. Such behavior is broadly in line with CES, because consumers share their budget over *all* perceived items, whereas they consume inelastically in the IV model. In view of the rich attention-diversity relation suggested by the CES-demand, an attempt to fit such a model to choice data, taking into account consumer attentiveness and responsiveness, could provide novel insights into the true nature of the Long Tail.

In sum, taking into account LA yields a much richer set of possible equilibrium patterns than predicted by standard theory. In general, the prevailing paradigm in a market depends in a non-trivial but intuitive way on preferences, attentiveness and responsiveness.

[51] Numerical evaluations show that the effect of LA indeed may be quantitatively strong: If $R = 2$, $\frac{F}{v\Delta} = 0.0001$ and $\sigma = \eta = 100$ the effective diversity exceeds the level as predicted without LA roughly by factor 2.

4.5 Obfuscation and limited attention

An observation in e-commerce is that firms "harness" the power of the Internet to protect or increase profits by resorting to obfuscation strategies (Ellison and Ellison (2009)). Obfuscation involves both attracting attention to a product or a website and, perhaps more characteristic, limiting the attention left for others once it has been successfully attracted. The versatility of the attention framework presented in this article allows to incorporate obfuscation into the model, and we now tackle the equilibrium consequences of obfuscation in a market. Ellison and Wolitzky (2012) model obfuscation as the possibility to increase consumer search costs in Stahl's two-type sequential search model (Stahl (1989)). One difference is that a decrease in per-capita information costs implies *higher* equilibrium prices in our setting. Another difference is that our contest setting produces more accessible comparative static results compared to the mixed strategy equilibria of search models.

We incorporate obfuscation in the framework of section 2 by assuming that the information measure r depends on firm actions, besides the attention allocation P. There are n active firms, where $n > R_0 > 1$, and R_0 quantifies a consumer's attention constraint. Let $r \equiv r(\{j\}) \geq 1$ and $\bar{r} \equiv r(\{g\}) \geq 1$, $g \neq j$, denote the information weights of the indicative firm j and its competitors. Conditional on being perceived, a consumer perceives $\frac{R_0 - r}{\bar{r}}$ competitors, and j's expected revenue is $\tilde{\pi}\left(f, \bar{f}, n, r, \bar{r}\right) V\left(p, \bar{p}, 1 + \frac{R_0 - r}{\bar{r}}\right) \Delta$, where $\tilde{\pi}_4 < 0$, and $\tilde{\pi}\left(f, \bar{f}, n, r, r\right) = \pi\left(f, \bar{f}, n, \frac{R_0}{r}\right)$. This captures the tension an obfuscating firm faces. Increasing own complexity r, e.g., by designing webpages to "trap" consumer attention, comes at the benefit of higher extractable attention rents once perceived (assuming $V_3 < 0$ as before). If all firms use similar obfuscation tactics, such that $\bar{r} = r$, this comes at the costs of lower perception chances due to a higher cognitive load at the market level. We note a positive externality between obfuscating firms: Higher competitor complexity is favorable to j, provided that j is perceived, because fewer spare mental resources remain to analyze competing products, which generally points towards a collusive potential of obfuscation.

In the following we assume that efforts to attract and keep consumer attention are summarized by f, i.e. $r = r(f)$ with $r'(f) > 0$, and overall effort costs are $\theta c(f)$.[52] For given n, a symmetric

[52] A more detailed setting would perhaps differentiate explicitly between attention (salience) efforts and obfuscation levels. It can be shown that in a model featuring separate, costly attention efforts and obfuscation levels, a change affecting both cost components in a similar way (e.g. a change of Δ) have a similar consequence for equilibrium prices as stated in proposition 1 below.

equilibrium (p,f) solves

$$\begin{aligned} V_1\left(p,p,\tfrac{R_0}{r}\right) &= 0 \\ \pi_1\left(f,f,n,\tfrac{R_0}{r}\right)V + \left(\tilde{\pi}_4\left(f,f,n,r,r\right)V - \tfrac{R_0}{nr^2}V_3\right)r'(f) &= \tfrac{\theta}{\Delta}c'(f) \end{aligned} \tag{19}$$

Higher equilibrium obfuscation implies higher prices because limited comparability reduces competitive pressure. Compared to (17), the additional term in the 2nd equation of (19) captures the countervailing forces pertaining to obfuscation. Limiting comparison allows to increase the attention rents earned, which reinforces attention-seeking, but increased obfuscation also diminishes the chance of perception in the first place. If $V(p,p,z)$ is sufficiently elastic with respect to z, the net effect is positive. If e.g. $\tilde{\pi} = \pi(f,\bar{f},n,1+\frac{R_0-r}{\bar{r}})$, then $\tilde{\pi}_4(f,f,n,r,r) = -(nr)^{-1}$, and the net effect is non-negative iff[53] $-\frac{V_3 R_0/r}{V} \geq 1$. We take this condition to hold. Assuming a regular solution to (19), lower per-capita information costs $\frac{\theta}{\Delta}$ then imply *higher* equilibrium prices:

Proposition 1 *If $V_{13} < 0$, then $p'(\theta/\Delta) < 0$ and $f'(\theta/\Delta) < 0$.*

If $\frac{\theta}{\Delta}$ decreases, obfuscation increases, and the market becomes less competitive. This is another vindication of the paradigm, characteristic for this article, that the presence of mental capacity constraints may severely alter the equilibrium predictions compared e.g. to models of scarce information, where a decrease of information costs on the consumer or firms side increases the equilibrium information level, thereby working in a pro-competitive way. Also notable is the observation that two very different ways of endogenous attentiveness – obfuscation or rational attention choices – share a common insight: The type of externality that the consumer mass Δ imposes on itself via the market mechanism is reverted compared to the standard case of unlimited mental capacities.

5 Conclusion

The classical theories of informative advertising or consumer search share a common prediction: Improved technology that allows, e.g., to move from brick-and-mortar to cheaper online shops, to reach more consumers or to transmit at lower information costs, should lead to more competitive markets, with perfect competition as its natural limit. Strong evidence indicating that frictionless e-commerce did not occur to the extent suggested by such theories appears puzzling, especially in view of the abundance of information available on the Web (e.g. Chevalier and

[53] If $V(p,p,z) = \frac{v(p)}{z}$, such as in the CES or the IV case, this condition holds.

Goolsbee (2003), Ellison and Ellison (2005)). By embedding the allocation of limited attention (LA) into competition theory, this article provides a powerful rationale for such findings. In particular, scarce attention is not a special case of scarce information and vice-versa, hence the different scarcity regimes require different theories. We have applied our attention framework to the empirically well-documented case of stimulus-driven attention allocation, where firms jointly compete for attention and in prices. Our setting offers a solid path through the "wilderness of bounded rationality" (Sims (1980)) in such as the notion of an optimal consumer decision over a given choice set is maintained. Yet we have to recognize that in an information-loaded society but a fraction of the true world is perceived, and it may depend on extra-individual aspects which parts make it into individual perception windows.

We showed that once LA is taken into account, traditional patterns of competition theory may break down both at the positive and the normative level. The symmetric model we study allows for a tractable and rich analysis. Strategic firms price to the market as perceived by consumers. This makes consumer attentiveness a crucial determinant of the firms' ability to "make the price" in a market, alongside with preferences. This has important consequences. Marginal cost pricing is not the natural limit of such an economy, and firms may earn positive attention rents in the short run. Scarcer attention allows firms to retrieve higher markups, and the market response is increased competition for attention. Whether attention rents are competed away depends on traditional fundamentals (preferences) and on attentional characteristics (responsiveness): If markets are very competitive in the sense that products are strong substitutes and consumer attention is sticky, then firms can secure positive profits from less attentive consumers. We have established that in discrete choice models, such as the circular model of product differentiation, this is the unique prediction. The inhibitive nature of LA then causes the standard pro-competitive effects of firm entry to break down: More diversity (e.g. because of lower setup costs) hurts consumers, because prices become rigid and perception chances are equated in the symmetric equilibrium, which implies that consumers increasingly fail to find suitable products despite such products being available. This welfare reversal also exists with more rational consumers, where attentiveness is part of their choice problem. In fact, things may get worse, because attentiveness-adjusting consumers realize that higher diversity makes it more difficult for them to retrieve a suitable variety. Their individually optimal behavior then is to reduce their attentiveness, which allows firms to set higher prices, and the increased profits induce further entry in the long run.

In general, the firms' ability to influence the attention process and traditional economic characteristics (e.g. tastes) matter jointly in a non-trivial way, and substantially different equilibrium patterns may emerge in dependence of their interrelation. This clearly distinguishes our contributions from other standard work on competition with boundedly rational agents, that assumes homogeneous products and a fixed market size. For example, our theory predicts that a market may produce a large diversity despite highly homogeneous consumer tastes, if attention is sufficiently sticky. Conversely, markets may become thin despite strongly differentiated tastes, if consumers' attention is highly volatile. This versatility of our predictions may help to explain why empirical evidence on the Long Tail effect is mixed at best (Elberse (2008), Brynjolfsson et al. (2011)). Our results also suggest that one should beware of jumping to preferential explanations of observed brand-loyalty patterns to quickly, as these could be caused by attention effects disguised as preference effects.

Our baseline setting introduced the attention-demand profiles as the central descriptive primitive. Our main analysis has separated between the competition for attention and the competition "within attention sets". While such a separation is reasonable in many situations and leads to a tractable structure of the model, important exceptions exist, e.g. price-filter search pages. Within our framework price-filters could be accommodated by assigning salience value to prices, thus making the attention allocation P_i price-dependent. An interesting extension could introduce product quality as an additional endogenous product attribute, which cannot be sufficiently assessed by price filters. How does then the possibility to quality-differentiate cope with the double role of pricing as attention-grabbing and a source of revenues? Another modification could seek to combine the competition for LA with attribute salience theory (Bordalo et al. (2013); Köszegi and Szeidl (2013)), where the relative salience of product attributes may affect the choice set formation process. Further, it would be interesting to analyze a setting, where besides salience-driven attention P_i, individual information access X_i is sender-dependent, capturing a trade-off between informing widely and densely. Moreover, modifying our baseline setting to include attention processes of a correlated, dynamic nature, e.g by relaxing non-independence of attention allocation between consumers ("attention generates attention") or between products ("consumers who bought X also bought..."), may provide a highly useful tool to analyze the formation of fashion trends, hypes and other collective attention phenomena. Finally, the interdependence between competition for attention and competition among perceived alternatives matters beyond market environments, and the framework in this article may provide valuable guidelines on how to introduce LA to other settings, such as international trade or the political economy.

Appendix

Appendix A provides a rigorous analysis of the abstract symmetric price-attention game from section 2. Appendix B contains proofs of the claims from the main text. In the online appendix we discuss the possibility of asymmetric attention equilibria, and show that these are unlikely to occur in our main setting.

A Symmetric price-attention game: Formal analysis

We first derive some important properties of the general perception function π as well as of its special versions (ACF and (7)). These are important for the subsequent equilibrium analysis. Next, we derive general conditions asserting existence and uniqueness of symmetric price-attention equilibria, and then show that the examples for V and π in the main text satisfy these conditions. Figure 6 organizes the results.

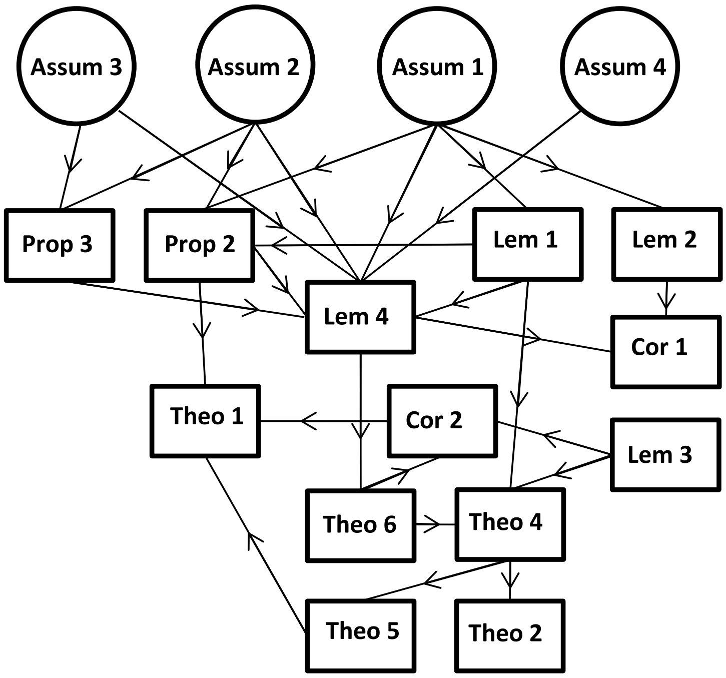

Figure 6: Overview

A.1 The perception function π: Properties

Lemma 1 *For $n > R_i$ and $f, \bar{f} > 0$, assumption 1 implies that $\pi_2(f, \bar{f}, n, R_i) < 0$, $\pi_1(f, f, n, R_i) = \frac{\zeta(n,R_i)}{f}$, $\frac{\frac{\partial}{\partial f}\pi_1(f,f,n,R)f}{\pi_1(f,f,n,R)} = -1$, $\lim_{f\to 0} \pi_1(f, f, n, R) > 0$ and $\lim_{f\to\infty} \pi_1(f, f, n, R) = 0$. Moreover, if $n > R_i$ and $f = \bar{f}$ then $\pi = \bar{\pi} = R_i/n$.*

Proof: Zero-homogeneity of $\pi(f,\bar{f},n,R_i)$ in $(f,\bar{f})$, and $\pi_1(f,\bar{f},n,R_i)>0$ imply that $\pi_2(f,\bar{f},n,R_i)<0$ and $\pi_1(f,f,n,R_i)=\frac{\zeta(n,R_i)}{f}$, where $\zeta(n,R_i)>0$, from which the other claims follow. ■

Lemma 2 *The ACF and* (7) *satisfy assumptions 1 and 4. For $n>R$ and $f>0$ we obtain* (6)*, where the suggested approximation is reasonable if n is large and R is small.*

Proof: It is easy to verify that (7) satisfies assumptions 1 and 4, so we concentrate on the non-obvious parts in case of the ACF. Define

$$G(f)\equiv\prod_{i=1}^{R}g(f,i)\qquad g(f,i)\equiv\frac{(n-i)\bar{f}}{f+(n-i)\bar{f}}\tag{20}$$

Hence $\pi(f,\bar{f},n,R)=1-G(f)$. With$-G'(f)=-G(f)\sum_{i=1}^{R}\frac{\partial g(f,i)}{\partial f}\frac{1}{g(f,i)}$ and $\frac{\partial g(f,i)}{\partial f}\frac{1}{g(f,i)}=-\frac{1}{f+(n-i)\bar{f}}$ we obtain

$$\pi_1(f,\bar{f},n,R)=-G'(f)=G(f)\sum_{i=1}^{R}\frac{1}{f+(n-i)\bar{f}}>0\tag{21}$$

An identical argument reveals that $\pi_3(f,\bar{f},n,R)<0$ because $\frac{\partial g(f,i)}{\partial n}=\frac{f\bar{f}}{\left(f+\bar{f}(n-i)\right)^2}>0$. Suppose that $n\geq R+1$. The strong monotonicity of π in R can be seen from $G(f,R+1)-G(f,R)=-G(f,R)\left(\frac{f}{f+\bar{f}(n-(R+1))}\right)<0$. Taking derivatives and rearranging gives

$$\begin{aligned}\pi_{11}=-G(f)\Bigg(&\left[\left(\sum_{i=1}^{R}\frac{\partial g(f,i)}{\partial f}\frac{1}{g(f,i)}\right)^2-\sum_{i=1}^{R}\left(\frac{\partial g(f,i)}{\partial f}\frac{1}{g(f,i)}\right)^2\right]\\&+\sum_{i=1}^{R}\left(\frac{\partial^2 g(f,i)}{\partial f^2}g(f,i)\right)\frac{1}{g(f,i)^2}\Bigg)\end{aligned}$$

But as $\left(\sum_{i=1}^{R}a_i\right)^2>\sum_{i=1}^{R}a_i^2$ if all a_i have the same sign and $\frac{\partial^2 g(f,i)}{\partial f^2}>0$, $\pi_{11}(f,\bar{f},n,R)<0$ follows. To obtain (6) note that (20) gives $G(f)=\frac{n-R}{n}$ if $f=\bar{f}>0$ and plug this into (21). Finally, we show why the suggested approximation works. Note that

$$\sum_{i=1}^{R}\frac{1}{1+n-i}=\sum_{j=0}^{R-1}\frac{1}{n-j}=H(n)-H(n-R)\tag{22}$$

where $H(x)$ is the x-th harmonic number. The sequence $(H(x)-Ln(x))_{x\in\mathbb{N}_+}$ monotonically converge from above to the Euler-Mascheroni constant γ. Hence $H(n)-Ln(n)=\gamma+e(n)$, where $e(n)$ is decreasing in n. Thus the maximal error occurs at $n=1$ and has $e(1)=1-\gamma<1/2$. Further $e(n-R)$ is larger the closer R is to n. Replacing $H(x)$ by $Ln(x)+\gamma+e(x)$ in (22) gives $\sum_{i=1}^{R} \frac{1}{1+n-i} = Ln\left(\frac{n}{n-R}\right) + e$, where the overall error $e \equiv e(n) - e(n-R)$ is bounded in absolute value by one, and gets small if n large and R small. By defining $x \equiv R/n$ we have $Ln\left(\frac{n}{n-R}\right) = Ln\left(\frac{1}{1-x}\right)$. A Taylor expansion around $x = 0$ gives $Ln\left(\frac{1}{1-x}\right) = x + r_1(0,x)$ with $|r_1(0,x)| \leq \frac{x^2}{1-x}$. Taking all these facts together shows that $\sum_{i=1}^{R} \frac{1}{1+n-i} \cong Ln\left(\frac{n}{n-R}\right) \cong \frac{R}{n}$. Hence the approximation works well if R is small compared to n. ■

Lemma 3 *Let $n > R > 1$ and $\bar{f} = f > 0$. The ACF and* (7) *verify the following properties: a)* $\frac{\pi_1}{n} + \pi_{13} > 0$, *b)* $\frac{\pi_{1f}}{\pi_1} < \pi_{13}\frac{n^2}{R}$ *and c)* $\frac{\pi_{1f}}{\pi_1} < -\frac{n}{R}\pi_1$.

Proof: We show the result only for the non-obvious case of the ACF.

a) As $\pi_{13} = \frac{1}{nf}\left(\frac{R}{n}\sum_{i=1}^{R}\frac{1}{1+n-i} - (n-R)\sum_{i=1}^{R}\left(\frac{1}{1+n-i}\right)^2\right)$ we have $\frac{\pi_1}{n} + \pi_{13} > 0$ iff $\sum_{i=1}^{R}\left(\frac{1+R-i}{(1+n-i)^2}\right) > 0$, and the last inequality is true.

b) As $\frac{\pi_{1f}f}{\pi_1} = -1$ (lemma 1) we have $\frac{\pi_{1f}}{\pi_1} < \pi_{13}\frac{n^2}{R}$ iff $\frac{R}{n} + \pi_{13}nf > 0$. Now, a) implies that $\pi_{13}nf > -\frac{n-R}{n}\sum_{i=1}^{R}\frac{1}{1+n-i}$. Therefore, we need only show that $R - \sum_{i=1}^{R}\frac{n-R}{1+n-i} > 0$, which is true because $R - \sum_{i=1}^{R}\frac{n-R}{1+n-i} = \sum_{i=1}^{R}\left(1 - \frac{n-R}{1+n-i}\right) = \sum_{i=1}^{R}\left(\frac{1+R-i}{1+n-i}\right) > 0$.

c) This follows from the proof of b) as $\frac{\pi_{1f}}{\pi_1} < -\frac{n}{R}\pi_1$ iff $R - (n-R)\sum_{i=1}^{R}\frac{1}{1+n-i} > 0$. ■

A.2 Equilibrium analysis

Suppose that the set of consumers can be partitioned into a finite number of attention types. Let $1 < \underline{R} = R_1 < ... < R_m = \bar{R} < \infty$, and $I_1, ..., I_m \subset [0, \Delta]$, with $\mu(I_i) > 0$, denote the groups of similar attention types. We assume, without loss of generality, that the sets I_i are left-closed and disjoint intervals.

Emergence of attention equilibria Analyzing (15) reveals the causes of an AE. The conditions for an AE to occur endogenously can be found by investigating the hypothetical equilibrium system with $R_i = \infty$.

Proposition 2 *Suppose that $V_i(p,p,z_i) > 0$ holds at any solution (p,f,n) of* (15)*. Equilibrium attention efforts f are positive iff the solution constitutes an AE. Moreover, any solution to* (15) *must be an AE iff any solution $(\tilde{p}, \tilde{n})$ to*

$$\begin{aligned} &\int V_{i1}(p,p,n)\,\mu(di) = 0 \\ &\int V_i(p,p,n)\,\mu(di) - \frac{F}{\Delta} = 0 \end{aligned} \tag{23}$$

satisfies $\tilde{n} > \underline{R}$.

Proof:

If $n \leq \underline{\mathrm{R}}$, then $\pi_i = 1\ \forall f, i$ and $f > 0$ cannot be optimal as f is costly. If $n > \underline{\mathrm{R}}$ then

$$\sum_{I_k : R_k < n} \int_{I_k} \pi_1^i(f, f, n, R_k) V_i(p, p, R_k) \mu(di) = C'(f)$$

in any equilibrium. As $\pi_1^i(f, f, n, R_i)$ is decreasing in f (lemma 1), Levi's monotone convergence theorem together with assumption 1 imply that

$$\lim_{f \to 0} \int_{I_k} \pi_1^i\left(f, f, n, R_k\right) V_i\left(p, p, R_k\right) \mu(di) > 0$$

As $C'(0) = 0$, $f = 0$ cannot be part of an equilibrium. If $\tilde{n} \leq \underline{\mathrm{R}}$, then $(\tilde{p}, 0, \tilde{n})$ is a solution to (15). Conversely, if (15) has a solution (p, f, n) with $n \leq \underline{\mathrm{R}}$, then $f = 0$ and hence (p, n) also is a solution to (23). ■

Generally, attention is more likely to become scarce ($\tilde{n} > \underline{\mathrm{R}}$) the more profitable the market is, which depends on technology (costs), demographics, preferences (intensity of competition) and budget (wealth in the market). If (23) has only regular solutions,[54] then an AE becomes more likely as per-capita setup costs fall ($\tilde{n}'\left(\frac{F}{\Delta}\right) < 0$). Similarly, if $V_i(\cdot), V_{i1}(\cdot)$ depend (positively) on some parameter ξ_i (a demand shifter), then $\tilde{n}'(\xi_i) \geq 0$. For example, ξ_i could be (individual) wealth allocated to the market (v_i) in example (14) or individual willingness to substitute (σ_i).

Conventional equilibria The previous argument shows why the auxiliary system (23) plays an important role in the equilibrium analysis. We continue by introducing general conditions asserting (23) to have a unique and regular solution $(\tilde{p}, \tilde{n})$. Let $V\left(p, \bar{p}, n, \xi_i\right) \equiv \int V_i\left(p, \bar{p}, n, \xi_i\right) \mu(di)$, where ξ_i is a demand-shifter with $V_4\left(p, p, n, \xi_i\right) > 0$ and $V_{14}\left(p, p, n, \xi_i\right) \geq 0$ at equilibrium points. We impose a standard regularity and boundary assumption on $V(\cdot)$, which is satisfied by examples (13), (14) and the discrete choice model.

Assumption 3 *The functions* $V(\cdot), V_1(\cdot)$ *are* C^1 *in* (p, n, ξ_i) *and satisfy:*

(i) For any $n > 1$ *we have* $V_1(c, c, n, \xi_i) > 0$*,* $\exists\ \hat{p}$ *with* $V_1(p, p, n, \xi_i) < 0$ *for any* $p \geq \hat{p}$ *and*

$$V_1\left(p, p, n, \xi_i\right) = 0 \quad \Rightarrow \quad \frac{\partial V_1}{\partial p} < 0, \frac{\partial V_1}{\partial n} \leq 0 \tag{24}$$

[54] The Jacobian of (23) with respect to (p, n) has a positive determinant.

(ii) $V(\tilde{p}(n), \tilde{p}(n), n, \xi_i) > 0$, $\lim_{n\downarrow 1} V(\tilde{p}(n), \tilde{p}(n), n, \xi_i)\Delta > F$ *and* $\lim_{n\to\infty} V(\tilde{p}(n), \tilde{p}(n), n, \xi_i)\Delta < F$, *where* $\tilde{p}(n)$ *is a solution to* $V_1(p, p, n, \xi_i) = 0$. *Moreover,* $V_2(\tilde{p}(n), \tilde{p}(n), n, \xi_i) < 0$ *for* $V(\tilde{p}(n), \tilde{p}(n), n, \xi_i) = \frac{F}{\Delta}$.

A solution $(\tilde{p}, \tilde{n})$ to (23) is regular, if the determinant of (23) with respect to (p, n) is positive.

Proposition 3 *Under assumptions 2 and 3 system* (23) *has a unique and regular solution* $(\tilde{p}, \tilde{n})$, *and* $\tilde{p} \in (c, \infty)$ *as well as* $\tilde{n} \in (1, \infty)$. *Moreover:* $\tilde{n}'(F) < 0$, $\tilde{n}'(\Delta) > 0$ *and* $\tilde{n}'(\xi_i) \geq 0$.

Proof: As a consequence of the boundary conditions in assumption 3 it follows that $V_1(p, p, n, \xi_i) = 0$ has a unique solution $\tilde{p}(n)$, $\tilde{p} \in (c, \infty)$, and $\tilde{p}'(n) \leq 0$. As $V(\tilde{p}(n), \tilde{p}(n), n, \xi_i)$ is continuous in n the boundary conditions of assumption 3 also assert that a solution $\tilde{n}$ to $V(p(\tilde{n}), \tilde{p}(n), \tilde{n}, \xi_i)\Delta = F$ exists, and $\tilde{n} \in (1, \infty)$. Moreover, the solution $\tilde{n}$ is unique because $\frac{\partial V(\tilde{p}(\tilde{n}), \tilde{p}(n), \tilde{n}, \xi_i)}{\partial n} = V_p \tilde{p}'(n) + V_3 < 0$. Because $\frac{\partial V(\tilde{p}, \tilde{p}, \tilde{n}, \xi_i)}{\partial p} \geq 0$ and $\frac{\partial V_1(\tilde{p}, \tilde{p}, \tilde{n}, \xi_i)}{\partial n} \leq 0$ by assumption 2, the solution $(\tilde{p}, \tilde{n})$ is also regular. Remaining claims follow from the Implict Function Theorem. ■

Examples (13) (for $c \in [0, 1)$), (14) and the discrete choice model satisfy assumption 3 for conveniently chosen parameters Δ, F. Consequently, these examples generate unique solutions $(\tilde{p}, \tilde{n})$ to (23).

Symmetric free-entry equilibria: Existence and uniqueness We now derive general conditions asserting existence and uniqueness of a regular, symmetric equilibrium (p, f, n) for our baseline case, where $R_i = R$ and $\pi_i = \pi$. The indicative firm's profit function (4) is $\Pi(p, f) = \pi(f, \bar{f}, n, R)V(p, \bar{p}, z)\Delta - F - C(f)$, $z = min\{n, R\}$, leading to the equilibrium system (17), as stated in the main text.[55] A corollary of our main result, theorem 6, is that if attention competition is represented either by the ACF or by (7), and economic competition by (13), by (14) or by the discrete choice model, then a single symmetric equilibrium (p, f, n) results. We start our analysis treating n as a fixed parameter:

$$\begin{aligned} &V_1(p, p, z) = 0 \\ &\pi_1(f, f, n, z)\, V(p, p, z)\, \Delta = C'(f) \end{aligned} \tag{25}$$

For notational convenience we set $V_p \equiv \frac{\partial V(p,p,z)}{\partial p} \geq 0$, $V_{1p} \equiv \frac{\partial V_1(p,p,z)}{\partial p} < 0$ and $p_{1f} \equiv \frac{\partial p_1(f,f,n,R)}{\partial f} < 0$. We also impose the following regularity assumption:

[55] For notational simplicity we ignore the demand shifter ξ.

Assumption 4 *If $n > R$ then $\lim_{n \downarrow R} \pi_1(f, f, n, R) = 0$ whenever $f > 0$. Moreover, $V(p, \bar{p}, k)$ is strongly quasiconcave*[56] *in $p > c$, $V(c, \bar{p}, k) = 0$ and $\lim_{p \to \infty} V(p, \bar{p}, k) \leq 0$ for any $\bar{p} \geq c$ and any $k > 1$.*

Lemma 4 *Under assumptions 1 - 4 system* (25) *has a unique solution (p, f) for any $n > 1$, and (p, f) is a global maximizer of* (4)*. Moreover, $p \in (c, \infty)$ and $f > 0$ iff $n > R$. Finally, the solution $(p(n), f(n))$ is continuous in n, and differentiable in n except at $n = R$.*

Proof: If $R \geq n$ then propositions 2 and 3 imply that (25) has exactly one solution $(p(n), f)$, and $f = 0$. If $R < n$ then there exists unique solution $p = p(R)$ to $V_1(p, p, R) = 0$ by the proof of proposition 3. The same logic asserts that $\pi_1(f, f, n, R) V(p(R), p(R), R)\, \Delta - C'(f) = 0$ has a unique solution $f \in (0, \infty)$ because of lemma 1 and strict convexity of $C(f)$. Consequently, there exists a unique vector (p, f) that solves (17), and $f > 0$ iff $n > R$. In any case $p \in (c, \infty)$. Turning to the SOC we immediately note that if $R \geq n$ then $(p(n), 0)$ is the unique optimizer of (4) by strong quasiconcavity of $V(p, \bar{p}, n)$. If $n > R$ then $(p(R), f)$ is a local max of (4) by assumptions 1 and 4. But as $\bar{f} > 0$ $p = p(R)$ must be part of the global maximizer. Together with the strict concavity of (4) in f this implies that $(p(R), f)$ indeed is unique the global optimizer of (4). We now prove the remaining two claims and denote the solution to (25) by $(p(n), f(n))$. *Case 1*: $n < R$. Then $f = 0$, $f'(n) = 0$ and $p(n)$ is continuously differentiable by strong quasiconcavity of V. *Case 2*: $n > R$. As the Jacobian of (25) with respect to (p, f) has a positive determinant, (p, f) is continuously differentiable in n, and $p'(n) = 0$. *Case 3*: $n = R$. Clearly, $\tilde{p}(z)$ is continuous in z, where $\tilde{p}$ is implicitly defined by $V_1(\tilde{p}(z), \tilde{p}(z), z) = 0$. Because $z = min\{R, n\}$ is continuous in n, it follows that $p(n)$ is continuous in n also at $n = R$. As $f(n) = 0$ for any $n \leq R$, we also have $\lim_{n \uparrow R} f(n) = 0$. We also know that $f(n)$ is continuous if $n > R$, and $f(n) > 0$ follows from proposition 2. Now suppose that $\lim_{n \downarrow R} f(n) > 0$. Then because of assumption 4 and the continuity of $\pi_1(f, f, n, R)$ in (f, n), we have $\lim_{n \to R^+} \pi_1(f(n), f(n), n, R) = 0$, but $\lim_{n \to R^+} C'(f(n)) > 0$, contradicting (25). Consequently, $f(n)$ is right-continuous at $n = R$ and thus everywhere continuous. ■

A useful property of the ACF (or of (7)) is that, as long as the value function V satisfies all above assumptions, it generates a unique equilibrium (p, f) to (25).

Corollary 1 *If $\pi(f, \bar{f}, n, R)$ is given by the ACF or* (7)*, and V satisfies all previous assumptions, then there exists exactly one solution (p, f) to* (25)*. With* (7) *and $n > R$ the equilibrium*

[56] $V_1(p, \bar{p}, k) = 0 \Rightarrow V_{11}(p, p, k) < 0$.

f is determined by

$$\frac{n-R}{n^2 f} R\Delta V(p(R), p(R), R) = C'(f) \tag{26}$$

Proof: The first part follows from lemma 4 as both the ACF and (7) satisfy assumptions 1 and 4 (see lemma 2). Using (6) in (25) gives (26). ■

We now turn to the case of free-entry equilibria.

Theorem 6 *Under assumptions 1 - 4 a solution (p, f, n) to* (17) *exists. Moreover, if*

$$\frac{\pi_{1f}(f, f, n, R)}{\pi_1(f, f, n, R)} < \pi_{13}(f, f, n, R)\frac{n^2}{R} \tag{27}$$

is satisfied at solution points, (p, f, n) is uniquely determined, where $1 < n < \infty$, $p \in (c, \infty)$, $f > 0$ iff $n > R$ and (p, f, n) is a regular solution of (17).

Proof: For given (p, f, n), $\bar{f} = f$ and $\bar{p} = p$ profits (4) are

$$\Pi(p, f, n) = \begin{cases} V(p, p, n)\Delta - F - C(f) & n \leq R \\ \frac{R}{n}V(p, p, R)\Delta - F - C(f) & n > R \end{cases}$$

Note that $\Pi(p, f, n)$ is continuous in (p, f, n). From lemma 4 we know that for any $n > 1$ a unique solution (p, f) to (25) exists, and $f > 0$ iff $n > R$. Further, lemma 4 also implies that $\tilde{\Pi}(n) \equiv \Pi(p(n), f(n), n)$ is continuous in n Then $\lim_{n\downarrow 1} \tilde{\Pi}(n) > 0$ (assumption 3) and $\lim_{n\to\infty} \tilde{\Pi}(n) = -F - \lim_{n\to\infty} C(f) < 0$ together with continuity of $\tilde{\Pi}$ imply that $\exists\ n(1, \infty)$ such that $\tilde{\Pi}(n) = 0$, which proves existence. By lemma 4, $\tilde{\Pi}(n)$ is continuous and differentiable unless $n = R$. If $\tilde{\Pi}'(n) < 0$ holds for all $n \neq R$ that satisfy $\tilde{\Pi}(n) = 0$, then there can be only a single n that solves $\tilde{\Pi}(n) = 0$. Suppose that $n < R$ and $\tilde{\Pi}(n) = 0$. Then $\tilde{\Pi}'(n) = V_2(p(n), p(n), n)p'(n) + V_3(p(n), p(n), n) < 0$ under the assumptions imposed on V. We now show that if $n > R$ and $\tilde{\Pi}(n) = 0$, then (27) implies $\tilde{\Pi}'(n) < 0$. Differentiation and using (25) gives $\tilde{\Pi}'(n) = V\Delta\left(-\frac{R}{n^2} - \pi_1 f'(n)\right)$. As $f'(n) = -\frac{\pi_{13}\Delta V}{\pi_{1f}\Delta V - C''}$, we obtain $\tilde{\Pi}'(n) < 0$ iff $\pi_{13}\frac{n^2}{R} > \frac{\pi_{1f}}{\pi_1} - \frac{C''}{C'}$. Because $C', C'' > 0$, (27) implies that $\tilde{\Pi}'(n) < 0$. Finally, the Jacobian of (17) with respect to (p, f, n) is negative (see appendix B) showing that (p, f, n) is regular. ■

To understand the intuition behind (27) note that, for $n > R$, an increase of n has two effects on profits. First, expected attention rents are diminished as perception chances are smaller. Second, as the marginal probability of attention depends on n, equilibrium effort f and thus attention costs $C(f)$ change. If π_1 increases (weakly) in n ($\pi_{13} \geq 0$), then equilibrium attention costs do not decrease. Hence both the revenue and the cost effect unambiguously lead to lower profits (and hence smaller n). The opposite could occur only if equilibrium attention costs drop very strongly, requiring that $\pi_{13} < 0$. Condition (27) asserts that a negative equilibrium response of f to a change in n must always be sufficiently small (consumer responsiveness is limited), excluding the possibility that the cost effect dominates the revenue effect. In fact, dominance of the cost effect would imply the existence of multiple symmetric free-entry equilibria, given our boundary assumptions.

There is no natural assumption on the sign of π_{13}, in fact both signs are possible with the ACF (or with (7)). As the ACF and (7) both satisfy condition (27), these functions, combined with appropriate economic competition, naturally imply existence and uniqueness of a regular, symmetric equilibrium:

Corollary 2 *If $\pi(f, f, n, R)$ is given by the ACF or (7) and V satisfies all previous assumptions, then the free-entry game has exactly one symmetric equilibrium (p, f, n). Moreover, the ACF or (7) combined with either (13) or (14) implies the existence of a unique free-entry equilibrium.*

Proof: For (7) it is straightforward to verify that (27) is satisfied. Lemma 3 shows that the same is true for the ACF. The rest follows from theorem 6. ■

B Proofs of main claims

Path-independent attention allocation

A useful general way of obtaining specific attention allocation rules is to suppose that attention can be described by an urn model, and specify simple sampling chances. Suppose that r_i is the count measure, and let $\tilde{P}_i$ be a probability measure on $(X_i, \mathcal{P}_i)$, where $\tilde{P}_i(j) \equiv \hat{\pi}_{ij}$ is the probability of sampling item $j \in X_i$. Then $|A| = R_i$, and assuming independent draws with replacement gives

$$P_i(A) \equiv \tilde{P}_i\left(A \,|\mathcal{A}_i\right) = \frac{\tilde{P}_i(A)}{\tilde{P}_i(\mathcal{A}_i)} = \frac{\prod\limits_{j \in A} \hat{\pi}_{ij}}{\sum\limits_{A \in \mathcal{A}_i} \prod\limits_{j \in A} \hat{\pi}_{ij}} \qquad A \in \mathcal{A}_i \tag{28}$$

We refer to (28) as path-independent attention allocation as, by construction, the sampling order is irrelevant. The perception function (7) is a special case of (28). Let $\hat{\pi}_j \equiv \pi$ and $\hat{\pi}_h \equiv \bar{\pi}$,

$h \neq j$, be given by the single-draw contest success function $\pi = \frac{f}{f+(n-1)\bar{f}}$, $\bar{\pi} = \frac{\bar{f}}{f+(n-1)\bar{f}}$. Applying these functions to (28) gives (7).

Note that (28) is a general allocation rule in thus that it is not limited to a bottom-up attention process. De los Santos et al. (2012) consider the case of fixed sample size search, where consumers rationally decide which stores to visit prior to making their final consumption choices. With idiosyncratic iid Gumbel errors in indirect consumer utility of sampling a particular store j, (28) gives a simple logit-type sampling model close to their equation (4).

Derivation of (8)

We partition the set of all consumers into n favourity groups, where the k-th group is the subset of consumers to whom the firm would offer the k-th highest net utility under unlimited attention capacities. N_k denotes the fraction of consumers who would find the indicative firm to be their k-th best choice.[57] The marginal consumer of group $k = 1$ is a distance $w = \frac{\bar{p}-p}{2t} + \frac{1}{2n}$ away of our firm. Hence, accounting for the marginal consumers on both sides, $N_1 = 2w_1$. Similarly, one obtains for the fraction of consumers in group $k = 2, ..., n-1$ that $N_k = 2(w_k - w_{k-1}) = \frac{1}{n}$. Finally, the fraction in the n-th group is residually determined as $N_n = 1 - \sum\limits_{k<n} N_k = \frac{p-\bar{p}}{t} + \frac{1}{n}$. Now, j makes a sale to a member of group k iff i) j is perceived ($j \in A$) and ii) j offers the highest net surplus among all products in A. Let $P(A \cap k)$ denote the probability to make a sale to a member of group k. Hence expected market demand is $\sum\limits_{k=1}^{n} P(A \cap k) N_k \Delta$. Note that $P(A \cap k) = P(A)P(k|\, A)$, where $P(A) = \pi(f, \bar{f}, n, R)$ is the probability that $j \in A$ and $P(k|\, A)$ is the probability of j being the best perceived firm. Moreover, if $R \geq n$ then $P(1|\, A) = 1$ and $P(k|\, A) = 0$ for $k > 1$. For $n > R$, $P(k|\, A)$ is the Laplacian

$$P(k|\, A) = \begin{pmatrix} n-k \\ R-1 \end{pmatrix} \begin{pmatrix} n-1 \\ R-1 \end{pmatrix}^{-1} \tag{29}$$

Straightforward summation for $n > R > 1$ then yields the payoff function:

$$\begin{aligned} \Pi &= \pi(f, \bar{f}, n, R)(p-c) \sum_{k=1}^{n} P(k|\, A) N_k \Delta - F - C(f) \\ &= \pi(f, \bar{f}, n, R)(p-c) \left(\tfrac{\bar{p}-p}{t} + \tfrac{1}{z}\right)\Delta - F - C(f) \qquad z = \min\{R, n\} \end{aligned}$$

Proof of theorem 1

Existence and uniqueness follow from corollary 2, as V satisfies all required assumptions. The cut-off condition and $f > 0$ follow from propositions 2. If π is given by (7), $n'(R) < 0$ in any

[57] The following formula presume that $\bar{p} - t/n < p < \bar{p} + t/n$.

AE by theorem 5, as $\varepsilon_V = 1$ in the IV model and $C(f)$ is strictly convex. Using the Gamma-function representation in proposition 4 (online appendix), $n'(R) < 0$ can also be verified for the ACF, which is not surprising given approximation (6); we omit the details of this proof. ■

Proof of theorem 2

For $R \geq n$ the result is standard, so let $n > R$. From the perspective of an arbitrary consumer i there are n favourity groups ranking products in terms of net utilities. i purchases his k-th favorite variety iff this variety is the best perceived variety, which happens with probability $P(A \cap k)$. Hence a fraction of $P(A \cap k)$ consumers acquire their k-th favorite variety, and face an average transportation distance $E[w|\, k] = \frac{2k-1}{4n}$. The law of iterated expectations gives

$$\bar{w} = E_k\left[E[w|\, k]\right] = \sum_{k=1}^{n} P(k \cap A)\left(\frac{2k-1}{4n}\right) \tag{30}$$

Then (29) and (30) imply that

$$\begin{aligned}\bar{w} &= \sum_{k=1}^{n} P(A)P(k\,|A)\left(\tfrac{2k-1}{4n}\right) = \tfrac{R}{4n^2}\begin{pmatrix} n-1 \\ R-1 \end{pmatrix}^{-1} \sum_{k=1}^{n} \begin{pmatrix} n-k \\ R-1 \end{pmatrix}(2k-1) \\ &= \tfrac{R}{4n^2}\left(\tfrac{2n(n+1)}{R(R+1)} - \tfrac{n}{R}\right) = \tfrac{(n+1)}{2n(R+1)} - \tfrac{1}{4n}\end{aligned}$$

$T'(\Delta) > 0, T'(F) < 0$ hold because $n'(\Delta) > 0, n'(F) < 0$ by theorem 4. ■

Proof of theorem 3

Let $\tilde{n} = \sqrt{\frac{t\Delta}{F}}$ denote the (hypothetical) solution without LA. If $-T_1(\tilde{n}, \tilde{n}) \geq h'(\tilde{n})$, then full attention is optimal, i.e. $R = \tilde{n}$. In this case $(\tilde{n}, c + \frac{t}{\tilde{n}}, 0, \tilde{n})$, the CE, is the unique solution to (12), so from now on suppose that $-T_1(R, \tilde{n}) < h'(\tilde{n})$. Because $-T_1(R, n) = \frac{(1+n)t}{2n(1+R)^2}$ is strictly decreasing in n and $h'(1) = 0$ there exists a unique rational inattention rule $R(n) \in (1, \tilde{n})$ on $(\tilde{n}, \infty)$, and $R'(n) < 0$ by the inhibition effect. Using (7) and $C(f) = f^{\eta}$ in (12) gives $\frac{n(\eta-1)+R}{n^2 R} = \frac{F}{\Delta}\frac{\eta}{t}$. Solving for R: $\hat{R}(n) \equiv \frac{n(\eta-1)}{\eta F \frac{n^2}{t\Delta} - 1}$. The equilibrium n solves $R(n) = \hat{R}(n)$. As $\hat{R}(\tilde{n}) = \tilde{n} > R(\tilde{n})$ but $\lim\limits_{n\to\infty} \hat{R}(n) = 0$ and $R(n), \hat{R}(n)$ are continuous, a solution exists and $n > \tilde{n}$. It then can be verified that for $\eta > 1$ we have

$$0 > R'(n) = -\frac{(1+R)t}{2n(1+n)t + 2n^2(1+R)^3 h''} \quad > \quad -\frac{R\left(n + \frac{2R}{\eta-1}\right)}{n^2} = \hat{R}'(n)$$

which implies uniqueness of the solution. Hence, by (12), also R, p, f must be uniquely determined. The remaining claims are straightforward calculus results. ■

Comparative statics: Calculations (Theorems 4 - 5)

We take assumptions 2 - 4 and (27) as satisfied and set $V_p \equiv \frac{\partial V(p,p,R,\xi)}{\partial p}$, $V_{1p} \equiv \frac{\partial V_1(p,p,R,\xi)}{\partial p}$ and $\pi_{1f} \equiv \frac{\partial \pi_1(f,f,n,R)}{\partial f}$. If (p,f,n) solves (17), then $V_{1p}, \pi_{1f} < 0$ and $V_p \geq 0$. The Jacobian corresponding to (17) with respect to (p,f,n) is

$$J = \begin{pmatrix} V_{1p} & 0 & 0 \\ \pi_1 \Delta V_p & \pi_{1f}\Delta V - C'' & \pi_{13}\Delta V \\ \frac{R}{n}\Delta V_p & -\pi_1 \Delta V & -\frac{R}{n^2}\Delta V \end{pmatrix} \tag{31}$$

$Det(J) < 0$ iff $\frac{\pi_{1f}}{\pi_1} - \pi_{13}\frac{n^2}{R} - \frac{\omega}{R\pi_1} < 0$, which holds by (27). Let $\chi \in \{R, \Delta, F, \alpha, \xi\}$. Then:

$$\begin{aligned}
&sign\,(p'(\chi)) = sign\,(V_{1\chi}) \\
&sign\,(f'(\chi)) = sign\left(V_{1\chi}\Delta\frac{RV_p}{n}\left(\pi_{13} + \frac{\pi_1}{n}\right) - V_{1p}\left(\frac{R\Pi_{f\chi}}{n^2} + \Pi_\chi \pi_{13}\right)\right) \\
&sign\,(n'(\chi)) = sign\left(V_{1p}\left(\left(\frac{\pi_{1f}}{\pi_1} - \frac{C''}{C'}\right)\Pi_\chi + \Pi_{f\chi}\right) - V_{1\chi}\Delta V_p\left(\left(\frac{\pi_{1f}}{\pi_1} - \frac{C''}{C'}\right)\frac{R}{n} + \pi_1\right)\right)
\end{aligned}$$

where $\Pi_\chi \equiv \frac{\partial}{\partial \chi}\left(\frac{R}{n}V(\cdot)\Delta - F - C(\cdot)\right)$ and $\Pi_{f\chi} \equiv \frac{\partial}{\partial \chi}\left(\pi_1(\cdot)\Delta V(\cdot) - C'(\cdot)\right)$. We now provide calculations for non-obvious cases. First, $sign\,(f'(\Delta)) = sign\,(f'(\sigma)) = sign\left(\frac{\pi_1}{n} + \pi_{13}\right)$, showing that f is increasing in both variables as a consequence of lemma 3. Similarly, $sign\,(n'(\Delta)) = sign\,(n'(\xi)) = sign(\frac{\pi_{1f}}{\pi_1} - \frac{C''}{C'} + \pi_1\frac{n}{R})$ which is also positive (lemma 3). We get $sign\,(f'(\alpha)) = sign\left(-\frac{C_{12}}{C_2} - \frac{\pi_{13}n^2}{R}\right)$ and lemma 3 implies that $\beta \geq 1$ gives $f'(\alpha) < 0$. For $C_2 > 0$ we get $sign\,(n'(\alpha)) = sign\left(\beta + \frac{\pi_{1f}}{\pi_1} - \mu\right)$, thus by lemma 1: $sign\,(n'(\alpha)) = sign\,(\beta - 1 - \mu)$. Similarly, if $C_2 < 0$, then $sign\,(n'(\alpha)) = sign\,(\mu + 1 - \beta)$, and if $C_2 = 0$, then $sign\,(n'(\alpha)) = sign\,(C_{12})$. From these results it follows that $n'(\alpha) > 0$ if $C(f,\alpha) = f^\alpha$, $\alpha > 1$. Next, using the definitions in the main text:

$$sign(f'(R)) = sign\left(-(1+\varepsilon_V)\left(\frac{\pi_1}{n} + \pi_{13}\right) + \frac{R}{n}\pi_{14} + \pi_{13}\right)$$

and $f'(R) < 0$ holds if π is given by (7). The above formula gives (use lemma 1)

$$sign\left(n'(R)\right) = sign\left(-\frac{R}{n}(1+\mu)\varepsilon_V - R\pi_{14}f + \pi_1(1+\varepsilon_V)f\right)$$

which can be reduced to the expression in theorem 5 if π is given by (7). Finally, if $C(f)$ is given by (2) then, by lemma 1, n is determined according to $\frac{R}{n}V(R) = \frac{F}{\Delta} + \frac{\zeta(n,R)V(R)}{\eta}$ in an

AE, showing that F and Δ affect diversity only in terms of per capita costs F/Δ. ■

References

Alba, J. W. and Chattopadhyay, A. (1986). Salience effects in brand recall. *Journal of Marketing Research*, 23:363–369.

Anderson, C. (October 2004). The long tail. *Wired Magazine*.

Anderson, S. P. and De Palma, A. (2009). Information congestion. *The RAND Journal of Economics*, 40(4):688–709.

Anderson, S. P. and De Palma, A. (2012). Competition for attention in the information (overload) age. *The RAND Journal of Economics*, 43(1):1–25.

Athey, S. and Ellison, G. (2011). Position auctions with consumer search. *Quarterly Journal of Economics*, 126:1213–1270.

Bagwell, K. (2007). The economic analysis of advertising. In Armstrong, M. and Porter, R., editors, *Handbook of Industrial Organization*, volume 3. Amsterdan: Elsevier B.V.

Baye, M. R., Gatti, J. R. J., Kattuman, P., and Morgan, J. (2009). Clicks, discontinuities, and firm demand online. *Journal of Economics Management Strategy*, 18:935–975.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121:803–843.

Bronnenberg, B. J. and Vanhonacker, W. R. (1996). Limited choice sets, local price response and implied measures of price competition. *Journal of Marketing Research*, 33:163–173.

Brynjolfsson, E., Hu, Y. J., and Simester, D. (2011). Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science*, 57(8):1373–1386.

Brynjolfsson, E., Smith, M. D., and Hu, Y. J. (2006). From niches to riches: Anatomy of the long tail. *MIT Sloan management review*, 47(4):67–71.

Butters, G. R. (1977). Equilibrium distribution of sales and advertising prices. *Review of Economic Studies*, 44:465–491.

Chevalier, J. and Goolsbee, A. (2003). Measuring prices and price competition online: Amazon.com and barnesandnoble.com. *Quantitative Marketing and Economics*, 1:203–222.

Davenport, T. H. and Beck, J. C. (2001). *The Attention Economy*. Boston, Massachusetts: Harvard Business School Press.

De los Santos, B., Hortacsu, A., and Wildenbeest, M. R. (2012). Testing models of consumer search using data on web browsing and purchasing behavior. *The American Economic Review*, 102(6):2955–2980.

Dickson, P. R. and Sawyer, A. G. (1990). The price knowledge and search of supermarket shoppers. *The Journal of Marketing*, pages 42–53.

Dreze, X. and Zufryden, F. (2004). Measurement of online visibility and its impact on internet traffic. *Journal of interactive marketing*, 18:20–37.

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

Elberse, A. (2008). Should you invest in the long tail? *Harvard Business Review*, 86(7/8):88–96.

Eliaz, K. and Spiegler, R. (2011a). Consideration sets and competitive marketing. *Reviews of Economic Studies*, 78:235–262.

Eliaz, K. and Spiegler, R. (2011b). On the strategic use of attention grabbers. *TheoreticalEconomics*, 6:127–155.

Ellison, G. and Ellison, S. F. (2005). Lessons about markets from the internet. *Journal of Economic Perspectives*, 19:139–158.

Ellison, G. and Ellison, S. F. (2009). Search, obfuscation and price elasticities on the internet. *Econometrica*, 77:427–452.

Ellison, G. and Wolitzky, A. (2012). A search cost model of obfuscation. *The RAND Journal of Economics*, 43(3):417–441.

Falkinger, J. (2008). Limited attention as the scarce resource in an information-rich economy. *The Economic Journal*, 118:1596–1620.

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96:1043–1068.

Gensch, D. H. (1987). A two-stage disaggregate attribute choice model. *Marketing Science*, 6:223–239.

Ghose, A. M. and Yang, S. (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science*, 55:1605–1622.

Grossman, G. M. and Shapiro, C. (1984). Informative advertising with differentiated products. *The Review of Economic Studies*, 51:63–81.

Hauser, J. R. and Wernerfelt, B. (1990). An evaluation cost model of consideration sets. *The Journal of Consumer Research*, 16:393–408.

Hefti, A. (2013). Equilibria in symmetric games: Theory and applications. *ECON working paper*.

Hefti, A. M. and Heinke, S. (2014). On the economics of superabundant information and scarce attention. *Forthcoming*.

Howard, J. A. and Sheth, J. N. (1969). *The Theory of Buyer Behavior*. New York: John Wiley.

Iyer, G., Soberman, D., and Villas-Boas, J. M. (2005). The targeting of advertising. *Marketing Science*, 24:461–476.

Kahneman, D. and Henik, A. (1981). Perceptual organization and attention. In Kubovy, M. and Pomerantz, J. R., editors, *Perceptual organization*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Konrad, K. (2009). *Strategy and Dynamics in Contests*. Oxford University Press.

Köszegi, B. and Szeidl, A. (2013). A model of focusing in economic choice. *The Quarterly Journal of Economics*, 128(1):53–104.

Masatlioglu, Y., Nakajima, D., and Ozbay, E. (2012). Revealed attention. *American Economic Review*, 102:2183–2205.

Maunsell, J. and Treue, S. (2006). Feature-based attention in visual cortex. *TRENDS in Neurosciences*, 29:317–322.

Mehta, N., Rajiv, S., and Srinivasan, K. (2003). Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing Science*, 22:58–84.

Milliken, B. and Tipper, S. P. (1998). Attention and inhibition. In In: Pashler, H., editor, *Attention*. Psychology Press.

Mitra, A. (1995). Advertising and the stability of consideration sets over multiple purchase occasions. *International Journal of Research in Marketing*, 12:81–94.

Nothdurft, H. C. (2000). Salience from feature contrast: additivity across dimensions. *Vision Research*, 40:1183–1201.

Novarese, M. and Wilson, C. M. (2012). Being in the right place: A natural field experiment on list position and consumer choice. *Unpublished Manuskript*.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., and Granka, L. (2007). In google we trust: Users decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, 12:801–823.

Pashler, H., editor (1998). *Attention*. Psychology Press.

Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin*, 92:382–402.

Perloff, J. M. and Salop, S. C. (1985). Equilibrium with product differentiation. *Review of Economic Studies*, pages 107–120.

Reis, R. (2006). Inattentive producers. *The Review of Economic Studies*, 73:793–821.

Salop, S. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, 10:141–156.

Simon, H. A. (1955). A behavorial model of rational choice. *The Quarterly Journal of Economics*, 69:99–118.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 48(1):1–48.

Sims, C. A. (2003). The implications of rational inattention. *Journal of Monetary Economics*, 50:665–690.

Smith, M. D. and Brynjolfsson, E. (2001). Consumer decision-making at an internet shopbot: Brand still matters. *Journal of Industrial Economics*, 49:541–558.

Stahl, D. O. (1989). Oligopolistic pricing with sequential consumer search. *American Economic Review*, 79:700–712.

Van Nierop, E., Bronnenberg, B., Paap, R., Wedel, M., and Franses, P. H. (2010). Retrieving unobserved consideration sets from household panel data. *Journal of Marketing Research*, 47(1):63–74.

Vives, X. (1999). *Oligopoly pricing.* Cambridge MIT Press.

Yantis, S. (1998). Control of visual attention. In Pashler, H., editor, *Attention.* Psychology Press.

C Online Appendix

In section C.1 we present a smooth version of the ACF using Gamma expansions. In section C.2 we discuss the possibility of asymmetric price-attention equilibria in our core model.

C.1 A smooth version of the ACF

Proposition 4 *For $\bar{f} > 0$, $f \geq 0$ and $n > R$,* (5) *can be represented by the smooth function*

$$\pi^C(f, \bar{f}, n, R) = 1 - \frac{\Gamma(n)}{\Gamma(n-R)} \frac{\Gamma\left(\frac{f}{\bar{f}} + n - R\right)}{\Gamma\left(\frac{f}{\bar{f}} + n\right)} \tag{32}$$

where $\Gamma(\cdot)$ denotes the Gamma-function and $n, R \in \mathbb{R}_+$. Further, for $f > 0$

$$\pi_1^C(f, f, n, R) = \frac{n-R}{nf} \left(\psi\left(1+n\right) - \psi\left(1+n-R\right)\right) \tag{33}$$

where $\psi(\cdot)$ is the digamma function. If $R \in \mathbb{N}$ then (33) *corresponds to* (6).

Proof: If $n, R \in \mathbb{N}$ and $n > R$ it can be verified that (32) corresponds exactly to (5) using the facts that $n! = \Gamma(n+1)$ and $\Gamma(x+1) = x\Gamma(x)$. Because $\Gamma(\cdot)$ is smooth on $\mathbb{R}_+$, π^C is smooth in $\left(f, \bar{f}, n, R\right)$ if $\bar{f} > 0$ and $n > R > 0$. Partial differentiation of (32) at $f = \bar{f} > 0$ then yields (33). Further, $\psi\left(1+n\right) = \psi\left(1+n-R_0\right) + \sum_{i=1}^{R_0} \frac{1}{1+n-i}$ (follows from $\psi(1+x) = \psi(x) + \frac{1}{x}$) implies that $\pi_1^C(f, f, n, R_0) = \frac{n-R_0}{nf} \sum_{i=1}^{R_0} \frac{1}{1+n-i}$ if $R_0 \in \mathbb{N}$. Examination of (33) shows that π_1^C is a strictly concave function of R as depicted in figure 7. ■

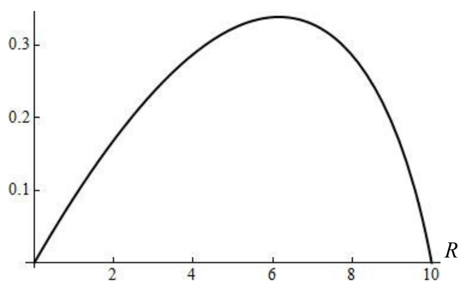

Figure 7: π_1^C as a function of R ($n = 10$)

C.2 Non-existence of asymmetric price-attention equilibria

From appendix A we know that under fairly general conditions (which all our examples satisfy) the price-attention game has exactly one symmetric equilibrium - but in principal there could be asymmetric equilibria. Checking for asymmetric equilibria in this game is far from trivial, as with its complicated two-dimensional structure with a "regime switch" standard tools (univalence or index theorem approach) are not applicable. We use some results developed in Hefti (2013) to obtain conditions excluding the possibility of asymmetric equilibria. We then illustrate that (13) together with the ACF or (7) satisfy these conditions. Throughout this section we take assumptions 2 - 3 as satisfied and consider the profit function

$$\Pi^j(p_j, f_j) = \sum_{A \in B_j} P\left(A, (f_j, f_{-j})\right) V^j\left(A, (p_j, p_{-j})\right)\Delta - F - C(f_j) \tag{34}$$

Uniqueness of pure price equilibria Suppose first that $\mathcal{A}_i = \mathcal{A}$ and $|A| = z > 1$ for all $A \in \mathcal{A}$ and take the attention allocation P to be *exogenously fixed* but non-degenerate in the sense that $\forall j\ \exists A \in B_j : P(A) > 0$. The revenue of attention of j is:

$$\hat{V}^j(p_j, p_{-j}) = \sum_{A \in B_j} P(A) V^j\left(A, (p_j, p_{-j})\right)\Delta \tag{35}$$

Consider the pure pricing game with payoffs $\sum\limits_{A \in B_j} P(A) V^j\left(A, (p_j, p_{-j})\right)\Delta - F$, where $P(A)$ is exogenously given. The individual first-order conditions are

$$\sum_{A \in B_j} P(A) \frac{\partial V^j\left(A, (p_j, p_{-j})\right)}{\partial p_j} = 0 \qquad \forall j \tag{36}$$

We first show that with assumptions 2 and 3 the pure pricing game has exactly one symmetric price equilibrium.

Lemma 5 *Under assumptions 2 and 3 system* (36) *has a unique symmetric solution* $p_j = p \in (c, \infty)$.

Proof: Set $p_j = p\ \forall j$ and note from (36) that, by symmetry of V, a symmetric equilibrium is a solution to $V_1(p, p, R) = 0$. This equation has a unique solution $p \in (c, \infty)$ by the proof of proposition 3. ■

The next step is to show that under an additional mild condition the symmetric equilibrium is in fact the unique equilibrium of the pricing game with heterogeneous attention allocation.

Assumption 5 *$\hat{V}^j(p_j, p_{-j}) \in C^2(S_p^n, \mathbb{R})$ is strongly quasiconcave in p_j, $\frac{\partial \hat{V}^j(c,p_{-j})}{\partial p_j} > 0$ and for any p_{-j} $\exists$ $\tilde{p} \in S_p$ such that $\frac{\partial \hat{V}^j(p_j,p_{-j})}{\partial p_j} < 0$ $\forall$ $p_j \geq \tilde{p}$. Finally, for any p_{-j} there is a p_j such that $\hat{V}^j(p_j, p_{-j}) > 0$ is possible.*

Assumption 5 essentially asserts existence and differentiability of the best-response function $p^j(p_{-j})$ and excludes boundary equilibria. An equilibrium $(p_1, ..., p_n)$ of the pricing game solves

$$\begin{pmatrix} \frac{\partial \hat{V}^1(p_1,p_{-1})}{\partial p_1} \\ \vdots \\ \frac{\partial \hat{V}^n(p_n,p_{-n})}{\partial p_n} \end{pmatrix} \equiv D_p(p_1, ..., p_n) = 0 \tag{37}$$

Proposition 5 *Suppose that for any j and any non-degenerate P the condition*

$$D_p(p_1, ..., p_n) = 0 \quad \Rightarrow \quad -\frac{\partial^2 V^j(A,p)}{\partial p_j^2} > (z-1)\left|\frac{\partial^2 V^j(A,p)}{\partial p_j \partial p_g}\right| \tag{38}$$

holds for any $g \neq j$ and $A \in B_j$ with $g \in A$, then (37) *has only one equilibrium, the symmetric equilibrium. Moreover, the equilibrium price p then is independent of the attention allocation P.*

Proof: The proof is an application of the index theorem (see e.g. Vives (1999)). If $D_p = 0$ implies $Det\,(J(-D_p)) > 0$, where $(J(-D_p))$ is the Jacobian of $-D_p$ with respect to $(p_1, ..., p_n)$, then one and only one equilibrium to the pure pricing game exists (for a given P) by the index theorem. But if this is true for any P, the unique equilibrium must be symmetric and independent of P by lemma 5. A sufficient condition for $Det\,(J(-D_p)) > 0$ is that $(J(-D_p))$ is a diagonally dominant matrix with positive diagonal entries. Take an arbitrary j and note that for any $A \in B_j$ $\exists$ $g(A) \in A$, $g(A) \neq j$, such that

$$\sum_{g \neq j}\left|\frac{\partial^2 V^j(A,p)}{\partial p_j \partial p_g}\right| \leq (z-1)\left|\frac{\partial^2 V^j(A,p)}{\partial p_j \partial p_{g(A)}}\right| \tag{39}$$

If condition (38) holds, then critical points also satisfy

$$-\sum_{A \in B_j} P(A)\frac{\partial^2 V^j(A,p)}{\partial p_j^2} > \sum_{A \in B_j} P(A)(z-1)\left|\frac{\partial^2 V^j(A,p)}{\partial p_j \partial p_{g(A)}}\right|$$

But then (39) and the triangle inequality imply that

$$-\sum_{A\in B_j} P(A)\frac{\partial^2 V^j(A,p)}{\partial p_j^2} > \sum_{g\neq j}\left|\sum_{A\in B_j} P(A)\frac{\partial^2 V^j(A,p)}{\partial p_j \partial p_g}\right| \tag{40}$$

which shows that $(J(-D_p))$ has a dominant diagonal. ■

We now show that the revenue function $\hat{V}$ implied by quadratic utility and linear demand, i.e.

$$d_j^i(A) = \begin{cases} \max\left\{0, \frac{1}{1+(z_i-1)\gamma_i}\left(1-\frac{1+(z_i-2)\gamma_i}{1-\gamma_i}p_j + \frac{\gamma_i}{1-\gamma_i}\sum_{g\in A: g\neq j} p_g\right)\right\} & j\in A \\ 0 & j\notin A \end{cases} \tag{41}$$

satisfies assumption 5 as well as condition (38). For simplicity, we set $c=0$ and $\Delta=1$.

Lemma 6 *$\hat{V}$ as induced by (41) satisfies assumption 5 and (38).*

Proof: Because each summand satisfies $\frac{\partial^2 V^j(A,(p_j,p_{-j}))}{\partial p_j^2} = -2\frac{1+(z-2)\gamma}{(1-\gamma)(1+(z-1)\gamma)} < 0$, $\hat{V}^j$ must be strictly concave, and it is straightforward to verify that also the remaining parts of assumption 5 are satisfied. Because $\frac{\partial^2 V^j(A,p)}{\partial p_j \partial p_g} = \frac{\gamma}{(1-\gamma)(1+(z-1)\gamma)}$ inequality (38) is satisfied iff $2+\gamma(z-3)>0$, which is true. ■

Nonexistence of asymmetric price-attention equilibria The following claim applies a result about asymmetric equilibria in symmetric games developed in Hefti (2013). Without loss of generality we set the indicative player as $j=1$.

Theorem 7 *Suppose that $n>R$ and $P(A,(f_1,f_{-1}))\in C^2((\cdot,(0,\infty)^n),[0,1])$, $A\in B_1$, is strictly concave in f_1, $\lim_{f_1\to 0}\frac{\partial P(A,\cdot)}{\partial f_1}>0$ if $f_2>0$, $P(A,\cdot)>0$ iff $f_1>0$. Moreover, take assumptions 2 - 5 as well as (38) to be satisfied for any non-degenerate attention allocation P. Then the symmetric equilibrium (p,f) as determined by (17) is the unique equilibrium of the symmetric price-attention game if for all $f_1,f_2\in(0,\infty)$:*

$$\sum_{A\in B_1} -\frac{\partial^2 P(A,\cdot)}{\partial f_1^2} + \frac{\partial^2 P(A,\cdot)}{\partial f_1 \partial f_2} \geq 0 \qquad f_3,\dots,f_n \geq 0 \tag{42}$$

Proof: Because of lemma 4 we only need to prove that there are no asymmetric equilibria. It follows from assumption 5 and the presuppositions of theorem 7 that $f_j=0$ never is part of j's best-reply. Therefore $f_j>0$ in any equilibrium, hence also $P(A)>0$ for any $A\in B_j$.

But then proposition 5 tells us that a unique and symmetric price equilibrium $p = (p_1, ..., p_1)$ exists for any non-degenerate attention allocation. Hence any equilibrium of the price-attention game is of the form $\{(p_1, f_1), ..., (p_1, f_n)\}$. Therefore, we need only show that the n equations $\sum_{A \in B_j} \frac{\partial P(A, f_j, f_{-j})}{\partial f_j} V(p_1, p_1, R)\Delta = C'(f_j)$ have no asymmetric solution $(f_1, ..., f_n)$. The slope condition developed in Hefti (2013) holds that if the best-reply function $f_1(f_2; f_3, ..., f_n)$ satisfies $f_1'(f_2; \cdot) > -1$, then such an asymmetric solution cannot exist. But as $\frac{\partial f_1(f_{-1})}{\partial f_2} = -\frac{\Pi_{f_1 f_2}}{\Pi_{f_1 f_1}}$, where $\Pi_{f_1 f_1} = \sum_{A \in B_1} \frac{\partial^2 P(A, \cdot)}{\partial f_1^2} V - C''(f_1) < 0$ and $\Pi_{f_1 f_2} = \sum_{A \in B_1} \frac{\partial^2 P(A, \cdot)}{\partial f_1 f_2} V$, condition (42) implies that $f_1'(f_2) > -1$. ■

We now illustrate for $R = 2$ that if P follows the ACF or (7), then (42) is satisfied in both cases. The claim also holds for $R = 3, 4, ...$ but the algebra gets increasingly messy. Note that in the special case, where $V > 0$ is a fixed prize (i.e. there is no price competition), the ACF or (7) generate a unique (attention) equilibrium. Hence the above result shows that the symmetric Tullock-Contests for R equally sized prizes implies a unique pure-strategy equilibrium, which generalizes the uniqueness result of single-prize Tullock contests (Konrad (2009) or Hefti (2013).

Corollary 3 *Suppose that $R = 2$. If P either follows the ACF or* (7) *and $V^j(A, p)$ is derived from* (41)*, then the symmetric equilibrium (p, f) is in fact the unique equilibrium.*

Proof: We use the following decomposition of B_1: $\hat{B}_2 \equiv \{A \in B_1 : 1 \in A \wedge 2 \in A\}$ and $\hat{B}_{-2} \equiv \{A \in B_1 : 1 \in A \wedge 2 \notin A\}$. Note that, as $R = 2$, we have $\hat{B}_2 = \{1, 2\}$. Let $A \in B_1$. In case of the ACF $P(A)$ takes on the form $P(A) = \frac{f_1 f_h}{\sum f_j} \left(\frac{1}{\sum f_j - f_1} + \frac{1}{\sum f_j - f_h} \right)$. Let $A = \{1, 2\}$ and $A' \in \hat{B}_{-2}$. Differentiation and some algebra yields:

$$sign \left(-\frac{\partial^2 P(A)}{\partial f_1^2} + \frac{\partial^2 P(A)}{\partial f_1 \partial f_2} \right) = sign \left(\frac{1}{(\sum f_j - f_h)^2} - \frac{1}{(\sum f_j)^2} \right) > 0$$

and

$$sign \left(-\frac{\partial^2}{\partial f_1^2} P(A') + \frac{\partial^2}{\partial f_1 \partial f_2} P(A') \right) = sign \left(\frac{(\sum f_j - f_1 - f_2)}{(\sum f_j - f_2)^3} \right) > 0$$

Hence (42) is satisfied for the ACF. If P follows (7), we obtain for $A \in B_1$: $P(A) = \frac{f_1 \sum_{g=2}^{n} f_g}{\sum_{j=1}^{n} \sum_{h=1+j}^{n} f_j f_h}$.

Some algebra reveals that (42) holds if $\left(\sum_{h=3}^{n} f_h \right)^2 \geq \sum_{h=3}^{n} f_h^2 + \sum_{j=3}^{n} \sum_{h=1+j}^{n} f_j f_h$ which is true. As both the ACF or (7) satisfy the other presumptions of theorem 7, the claim follows from theorem 7 and lemma 6. ■