Œconomia

5-1 (2015)
The Economics of Attention

Andreas Hefti et Steve Heinke

On the economics of superabundant information and scarce attention

On the economics of superabundant information and scarce attention

Andreas Hefti and Steve Heinke*

This article provides an introduction to the literature addressing the causes and consequences of limited attention in economics. We present a simple set-theoretic micro-structure describing the allocation of attention, and use this framework to explain the central notions of goal—and stimulus—driven attention mechanisms. After presenting empirical evidence on limited attention from psychology, marketing and internet research, we use our baseline setting to discuss a number of recent contributions featuring some form of limited attention or related phenomena, such as obfuscation.

Keywords: limited attention; information overload; attention allocation; salience; perception; informative advertising; rational inattention; bounded rationality

L'économie de l'information surabondante et de l'attention rare

Cet article propose une introduction à la littérature traitant, en économie, des causes et des conséquences de l'attention limitée. Nous présentons une microstucture simple de théorie des ensembles qui décrit l'allocation de l'attention, et nous utilisons ce cadre pour expliquer les deux notions centrales en jeux dans les mécanismes de l'attention : l'attention dirigée par les objectifs et l'attention dirigée par les stimuli. Après avoir exposé des preuves empiriques sur l'attention limitée empruntées à la psychologie, au marketing et au travaux spécifiques sur les comportement de recherche sur internet, nous utiliserons

* Andreas Hefti, Department of Economics, University of Zurich; Zurich University of Applied Sciences, School of Management and Law (andreas.hefti@econ.uzh.ch); Steve Heinke, Department of Economics, University of Zurich.

This article was partly drafted during a research stay of Hefti at the economics department of Harvard university. Hefti gratefully acknowledges third-party funds of the Swiss National Foundation. We are grateful to two anonymous referees and the editor of this journal for many helpful and constructive suggestions. We also wish to thank participants of seminars at the University of Zurich, Harvard University, and the Zurich University of Applied Science for their insightful comments.

notre cadre théorique pour discuter plusieurs contributions récentes qui intègrent certaines formes d'attention limitée ou de phénomènes liés, comme l'obfuscation.
Mots clés : attention limitée, surcharge d'information, allocation de l'attention, perception, saillance, publicité informative, inattention rationnelle, rationalité limitée

JEL: D43, D03, L11, L13, L86, C72, C79, M3

In the modern, digitalized economy superabundance of information is an omnipresent phenomenon. For example, a typical Google search query such as "Hawaii holiday" generates more than 100m hits in less than one second, or the Sunday New York Times contains more factual information in one edition than was available to a reader in all the written material of the fifteenth century (Davenport and Beck (2001)). Companies spend enormous amounts to influence the stream of information. Sponsored search advertising – the acquisition of specially marked search-engine links in ongoing online auctions – has become "the largest source of revenues for search engines" (Ghose and Yang (2009)).
The main concern whether individuals or societies can properly handle any amount of information has been expressed already prior to the Internet age. Among the first, Simon (1955b) reconsiders the notion of the "rational economic man" by emphasizing the existence of "internal constraints", which account for the fact that an agent faces physiological and psychological limitations to which he must obey, besides economic scarcity as represented by budget constraints. Together with such limitations a "wealth of information" generates a "scarcity of attention" (Simon (1955a)), and managing scarce attention is viewed as a key issue for success in modern business (Davenport and Beck (2001)). A by now large literature in psychology, marketing and consumer behavior research documents the importance of limitations on perceiving multiple stimuli for making decisions, storing information, planning actions and other mental processes (Pashler (1998)).

This article provides a short introduction to some of the recent approaches exploring the causes and consequences of limited attention in economics, with a special focus on what is frequently referred to as *stimulus-driven* attention allocation, as opposed to the many models featuring *goal-driven* attention. In simple terms, *stimulus-driven* (or bottom-up) attention allocation means that someone's attention is directed towards certain objects by external signals (stimuli), such as a flashy link in a web browser (the "pop-up" effect), an eye-catching placement on a supermarket shelve, or salient headlines in a newspaper. The abstract essence of this approach is that human cognition is not entirely self-contained in such that intense or repetitive exposure to external messages may influence what is on a decision-makers mind at the time of a decision (and what not). In contrast, goal-driven (or top-down) attention allocation portrays the decision-maker as having full mental control over her attention process. The essence of this approach is that while attention is a scarce resource—not everything can be considered—the decision-maker is aware of her limited attention, and has means to rationally decide (optimize) which information sources to attend and which to ignore, prior to making her choice.

Here is an overview of the article. Propelled by the evolution of the Internet and its "Big Data", the last decade has seen a surge of interest by economists and choice theorists in limited attention, which was previously a recurring theme in psychology and the marketing science. If a decision-maker is overloaded with information, e.g. a shopper facing endless shelves of cornflakes or an investor reading through several magazines trying to identify the ultimate investment opportunity, an attention allocation problem exists: which items are considered, which not? While the experience of information superabundance seems familiar from every day's life, abstracting such situations in terms of a formal model that captures the essence of the attention problem is far from trivial. When is attention scarce and information abundant? If attention is scarce, what is the mechanism governing attention allocation? Is there a difference between scarce information and scarce attention that is meaningful to economics?

Perhaps the best witness of the complexity and versatility of the attention problem is the substantial heterogeneity in thoughts and corresponding formal models of attention, making it hard for casual readers and experts likewise not to get lost in the plethora of assumptions and claims made. To address this problem, we introduce a simple logical structure in section 1 that formalizes a general attention allocation problem by means of some set-theoretic primitives. This allows us to precisely draw the dividing line between goal- and stimulus-driven attention at the individual level, and at the systemic level between economies featuring scarce information versus scarce attention. Our framework also provides a useful way of organizing the different contributions on limited attention and related phenomena.

We review some of the empirical evidence on limited attention in section 2, and continue by discussing the causes of attention as a scarce resource, or equivalently information overload, in section 3. While there are substantial differences in how attention is formally introduced, it is a common understanding that limits on attention are capacity constraints of some form on the perceivable measure of information. While some contributions take such constraints as completely exogenous characteristics of individuals (Falkinger (2008); Anderson and De Palma (2012)), they are endogenously determined in other models by decision-makers that adjust their attention gates as a response to their information exposure (Anderson and De Palma (2009); Hefti (2012)). A common insight is that attention becomes a scarce resource, meaning binding capacity constraints, if more informa-

tion senders find it attractive to enter a market, e.g. because spendable consumer income has increased or broadcasting costs have decreased.
In section 4 we review some contributions that focus on the consequences of limited attention for competition and welfare in market economies. While some papers take the attention allocation mechanism (assigning mental resources to decision items) as an exogenous feature of the model (Anderson and De Palma (2009, 2012)), other articles explicitly model the quest for attention as an additional competitive stage in the information age (Falkinger (2007, 2008); Hefti (2012)).
While the duality between rich information and scarce attention is a common theme in the aforementioned articles, several other contributions abstract away from the interplay between information (overload) and attention, and concentrate more on what the consequences of marketing methods aiming at attracting attention of new customers or at retaining attention from already captive consumers are for competitive equilibria. We review some articles of this spirit (Eliaz and Spiegler (2011a,b); Ellison and Wolitzky (2012)) in section 5.
Finally, we present a simple model of goal-driven attention allocation that exemplifies the notion of entropy-based rational inattention, which is the standard concept in this part of the attention literature. We conclude with some general remarks concerning the development of attention theory in economics.

1 Limited information, scarce information, scarce attention

To our knowledge, Falkinger (2007) was the first to provide a general equilibrium model of limited attention, thereby also addressing the logical structure of an economic model featuring information senders, and attention-constrained information receivers. For a sender (e.g. an advertising firm) it is important to recognize how many and which receivers obtain its messages (its "audience"). For a receiver (e.g. a consumer) it matters which senders she actually perceives (her "membership"). While Falkinger (2007, 2008) concentrates on general equilibrium properties of economies with limited attention and zero-measure senders and receivers in a deterministic setting, other contributions have emphasized a strategic or stochastic nature of attention-seeking. In the next section we present some set-theoretic primitives that allow us to compare the different settings through one lens.[1]

1.1 Modeling limited attention: Set-theoretic primitives

There is a measure Δ of receivers, which we usually associate with consumers. Let X be a set consisting of n information items. While we gener-

1 Naturally, our setting shares some conceptual building blocks, such as some notion of a membership or an audience, with Falkinger (2007).

ally view X as the set of all existing alternatives in some abstract decision context, many of the subsequent models operate within a product market environment, which means that X is the set of different (differentiated) products in a conventional economic sense. Let $X_i \subset X$ denote receiver i's *information set*, i.e. the subset of all items that are part of i's information technology. Conceptually, X_i are those items a consumer, in principal, *could* observe, but perhaps misses by her limited attention. For example, we would say that $j \in X_i$, if j is advertised in a magazine that i regularly reads, or j is indexed and ranked by Google for some given search query of i.

Scarce attention Let $A_i \subset X_i$ denote i's *attention set*, i.e. the subset of items perceived by i.[2] A_i is i's psychologically feasible set, i.e. the set of alternatives over which i reaches some economic decision. For example, A_i could be the collection of perceived fragrances or shampoos in a supermarket, and consumer i decides how to spend his budget over these alternatives. We say that i's attention is scarce if $A_i \subsetneq X_i$.
The sets X, X_i and A_i are elementary descriptive primitives naturally present in some form in models featuring limited attention.[3] We shall see that the notion of limited attention is formalized by some form of capacity constraint on the measure of perceivable information in most of the subsequent models. Then, attention becomes scarce if the corresponding capacity constraint becomes binding, i.e. $A_i \subsetneq X_i$ happens if the information set X_i is "too voluminous" with respect to i's capacity constraints.

Limited information We say that receiver i has limited information about X if $A_i \subsetneq X$. Hence limited information can be either caused by scarce information (but abundant attention), i.e. $A_i = X_i \subsetneq X$, or by scarce attention (but abundant information), i.e. $A_i \subsetneq X_i = X$ or, of course, by both. The first type of limited consumer information, caused by insufficient available information, has been extensively studied in economics. In the context of market economies the most relevant examples of theories on scarce information are models of consumer search or informative advertising, and we review a well-known example of informative advertising (Grossman and Shapiro (1984)) in section 4.1. An important question therefore is, whether the *cause* of limited information has economically meaningful implications, so that neither type of scarcity is a special case of the other. Empirical evidence on limited attention and related phenomena such as obfuscation as well as the models covered by this article provide many arguments in favor of a clear distinction between the two scarcity regimes. Thus a separate theory for both seems warranted.

2 The attention set broadly corresponds to what has been called consideration set or evoked set in the marketing science (see section 2), and it coincides with the notion of a membership in Falkinger (2007).

3 With descriptive we mean that these sets are logically present in models of limited attention, even if they are not modeled in terms of an explicitly given definition.

Receiver choice with limited attention A model of limited attention becomes interesting by building more logical structure around X, X_i and A_i, especially by specifying the interdependence of attention sets and consumer demand. Let receiver choice be captured by a (demand) function $d^i : \mathcal{P}(X) \to \mathbb{R}^n_+$, where $\mathcal{P}(X)$ is the power set of X, and the j-th component satisfies $d^i_j(A_i) = 0$ if $j \notin A_i$. That is, d_i describes what person i chooses from a given attention set A_i, and only perceived items can be requested.[4]

Suppose that the receiver choice process can be described by a stylized composite mapping of the form $d^i \circ P_i \circ \Phi^i$, where $X_i = \Phi^i(X)$ and $A_i = P_i(X_i)$. The models on limited attention we cover in this article feature a choice procedure consistent with such a mapping. Its essence is that decisions can be made only over alternatives that survive two different types of filtering:

- A technological filter Φ_i, describing which information is available to receiver i.
- A psychological filter P_i, describing which of the available information is actually perceived.

The attention models covered by this article differ with respect to their specific assumptions on these filters, but they all have in common that limited attention affects the choice process conditional on X_i, the set of items that receiver i can, in principal, gather information about.

1.2 Attention allocation

If attention is a scarce resource, then an attention allocation problem exists: Which items in X_i are perceived and which not? In the following we describe the individual attention allocation by means of a probability measure P_i on the measure space $(X_i, \mathcal{P}(X_i))$.[5] For any $A \in \mathcal{P}(X_i)$, the probability to perceive attention set A is given by $P_i(A)$. Intuitively, P_i specifies the frequency with which i observes A_i. In this vein, we interpret the function P_i as the allocation of i's attention, conditional on X_i.[6]

Once (X_i, P_i, d_i) is specified receiver-wise, one can derive a corresponding attention allocation at the population level. To see this, consider a finite population consisting of Δ receivers. Note that we can extend the probability measure P_i from the measure space $(X_i, \mathcal{P}(X_i))$ to $(X, \mathcal{P}(X))$ by setting $P_i(j) = 0 \;\forall\; j \notin X_i$. Then the population level information set is

4 The assumption that perception precedes choice is of great importance in modern choice theory when attempting to uncover preferences from choice data produced by consumers with limited attention (Masatlioglu et al. (2012)).

5 We could e.g. specify a probability measure $\hat{P}_i$ on $(X, \mathcal{P}(X))$, and for $(X_i, \mathcal{P}(X_i))$ set $P_i(A) \equiv \hat{P}_i(A|X_i)$.

6 This differs from Falkinger (2007), where an attention allocation is an economy-wide deterministic assignment of receivers (consumers) to senders (firms).

$\bigcup X_i \equiv \tilde{X} \subset X$, and it is easy to see that $P(A) = \sum_i \frac{P_i(A)}{\Delta}$ defines a probability measure on $\left(\tilde{X}, \mathcal{P}(\tilde{X})\right)$, representing the average frequency with which attention set $A \subset \tilde{X}$ is observed in the economy.
It is important to bear in mind that the objects (X, X_i, P_i, d_i) are mainly descriptive vessels of the choice process. While we view (X, X_i, P_i, d_i) as a natural micro-structure of any limited attention model, there is a plethora of possibilities how these objects could be (and have been) specified by particular models. For the objective of this article, they provide a useful and transparent way of organizing the different contributions.

1.3 An organization of the literature covered

Regarding the individual allocation of attention an important separation can be drawn between models featuring a goal-driven or a stimulus-driven attention process.[7] In terms of our simple primitives, this boils down to two fundamentally different ways of how the attention allocation function P_i could be determined.
"Goal-driven" means that one's attention allocation is essentially under one's own control, and as such a part of a decision-maker's optimization process. In terms of our simple primitives this means that P_i is endogenously determined by i's choice. Loosely spoken, P_i answers questions such as "Which part of information data should I analyze" (Sims (2003)), or "How frequently should I update my information about current economic conditions in order to optimally adjust my consumption plans" (Reis (2006)).
In contrast, "stimulus-driven" means that the individual attention process depends on the strength or conspicuousness of the arriving signals (stimuli). For example, such stimuli could be various advertising messages broadcasted over different media, or visual features increasing the salience of certain information elements such as weblinks (placement, highlighting). In terms of our primitives this means that P_i depends on the attention-seeking activities of the various senders.
In reality, attention features both goal- and stimulus-driven characteristics. It is a natural path of science to consider these aspects separately in order to distill out the essential substrata of each process. One interesting intermediate case, covered in sections 4.2 - 4.4 of this article, is to maintain P_i as determined by external stimuli, but allow individuals to adapt their "attention spans", i.e. the measure of signals that trespass the perception gate. This means that while perception is prone to external signal interference, a decision-maker may respond to changes in the economic environment by adapting how penetrable his attention filter is.

Limited attention models: An overview Given the dominant importance of individual optimization (or self-control) in the notion of most equilib-

7 This dichotomy is also omnipresent in the psychological literature on attention.

rium concepts in economics, it is not surprising that economics has squarely focused on goal-driven attention, frequently referred to as models of *rational inattention*. While we discuss rational inattention in section 6 in the context of a beauty contest, we chose to concentrate on contributions dealing more with limited attention and sender-driven attention allocations in the main part of this article, as we see this strand as important and still underrepresented in economics.

A common feature of all models covered by this article are stable preferences, meaning that a change of observed behavior is not attributed to a change of tastes, but rather to a change in either perception or decision-relevant economic variables, such as prices or qualities of products. The papers covered by our article differ in their assumptions about demand d_i (economic competition), about how information is quantified to make the notion of a capacity constraint meaningful and about how the attention allocation P_i is determined (see table 1). First, some papers introduce limited attention as a bound on signal volume, while others consider limited attention as a bound on the measure of perceivable signal sources. Next, some papers introduce salience competition, meaning that the signal strength of the messages (their salience) influences whether they are perceived or not, while other papers take the allocation function P_i as completely exogenous. Finally, the articles vary with respect to the underlying economic environment, especially with respect to the form of product market competition, or whether there is free entry, in which case X_i (or X) are endogenous to the model.

Table 1: Overview: Models with limited attention.

	I	II	III	IV	V	VI
Volume-based LA	+	+	−	−	−	−
Item-based LA	−	−	+	+	−	+
Endogenous LA constraints	−	−	+	−	−	+
Salience competition	+	+	−	−	+	+
Product market competition	−	General eq.	−	−	Duopoly	Oligopoly
Free entry (info sources)	+	+	+	+	−	+

Indexing: I Falkinger (2007), II Falkinger (2008), III Anderson and De Palma (2009), IV Anderson and De Palma (2012), V Eliaz and Spiegler (2011a,b), VI Hefti (2012)

2 Limited attention and stimulus-driven attention: Evidence

Several contributions on attention covered by our survey have been inspired and motivated by the strong evidence on limited attention from psychology, marketing science and research on consumer behavior.[8] In this section we summarize some of the main empirical facts of these fields.

Psychology The distinction of attention as an active or passive part of cognition is a core issue in the psychology of attention (see Pashler (1998)). In this literature it is a central question "whether attention is goal-driven, controlled in a top-down fashion, or stimulus driven, controlled in a bottom-up fashion" (Yantis (1998)). While both allocation mechanisms are important in explaining how subjects perform during visual search experiments, much evidence suggests that "stimulus-driven attentional control is both faster and more potent than goal-driven attentional control" (Yantis (1998)). Experimental findings on visual search show that response time of subjects to a stimulus is flat if few objects are displayed, but increases exponentially with the number of visualized objects (Mozer and Sitton (1998)). Such findings suggest that capacity limits impose a bottleneck to stimulus processing, meaning that in presence of multiple stimuli attention operates as a gating mechanism by restricting the amount of processed information. Hence some signals are inhibited, while others are processed on to the recognition network.[9]

Related studies emphasize that the "relative salience of a target" is much more important than "the occurrence of specific features" in explaining observed search time (Nothdurft (2000)). Thus 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. Experimental evidence on the spatial distribution of attention over the visual field suggests that attention frequently works as a *spotlight*. Once a location is selected, all features at that location are processed and moved on to the recognition network (e.g. Kahneman and Henik (1981)) or "receive enhanced processing" (Maunsell and Treue (2006)). In sum, there is substantial empirical evidence that the *relative salience* of a target importantly affects the amount of mental processing the object receives, at least in sufficiently complex environments.

Marketing studies While research on the equilibrium consequences of limited attention is a very recent development in economics, the idea that consumers reach their decisions conditional on a subset of all possible alternatives has a long-standing tradition in marketing, dating back to Howard

8 This is in particular the case for models featuring some form of salience competition, e.g. Falkinger (2007, 2008), Eliaz and Spiegler (2011a,b) and Hefti (2012).

9 See Milliken and Tipper (1998) for experiments on the importance of inhibition in selection experiments.

and Sheth (1969). Hauser and Wernerfelt (1990) provide an overview of empirical marketing studies dealing with consideration sets. The authors emphasize that these sets usually are very small compared to the complete set of possible alternatives. For example, the median shopper considers about 4 shampoos out of more than 30, or $2-5$ autos out of more than 160. Interestingly, the average consideration set *size* varies comparably little (between $2-6$) between product classes (Cars, Food, Medicine, Soaps...) and, perhaps surprisingly, seems not to hinge on the size of the grand set of alternatives. While consideration set sizes appear to be stable even across product categories, the content of the consideration sets are not, and depend critically on situational or marketing factors that increase the visibility or salience of the products (Fader and McAlister (1990), Mitra (1995), Allenby and Ginter (1995)). This closely matches the tenets from psychology on the pertinence of stimulus-driven attention, and also highlights why such findings are important to economics.

Internet research Search engines data provides considerable evidence indicating that information concentrated in certain salient regions receives critically more attention. For example, it is a well-documented fact that the on-screen placement of a link on an online search page is a central determinant of the number of clicks that a site receives. Using a dataset of a price listing service Baye et al. (2009) find that "a firm receives about 17% fewer clicks for every competitor listed above it on the screen", and the relationship between clicks and screen location is far from linear.[10] In e-commerce, Drèze and Zufryden (2004) show, using data on the web traffic of Amazon and Ebay, that a site can increase its visibility to potential consumers most by increasing its linkage in the web, thereby improving its indexing by search engines. In case of computer memory modules Ellison and Ellison (2009) report that on a price search engine "moving from first to seventh on the list reduces a website's sales ... by 83%". Similar findings are documented in case of sponsored search advertising (Ghose and Yang (2009)). These authors additionally document a non-monotonic relationship between rank and profitability, which they explain by the fact that both costs and clicks decrease with a worse rank (a lower on-screen position), with clicks decreasing slower than costs. A related finding has been documented in case of the online book market: Being on the first page of an online search is vital to receiving a high number of clicks, and in such perhaps more important than being the first in the list (Smith and Brynjolfsson (2001)). First-page dominance is also found by Jansen et al. (2000). Such evidence is consistent with attention working as a spotlight: What matters is to be within the cone of light of attention and less the precise location within the cone.
Finally, Pan et al. (2007) show in an eyeball-tracking experiment that the first entries on a Google search page get by far most viewing-time, but

10 The first three positions get about 75% out of all clicks of the first ten on-screen positions, while positions with rank larger than three do not differ substantially from each other.

there is no statistical difference between the first and the second item on the list. In their online experiment these authors find that people frequently select among the top items of a list in a quality based search experiment even if these upper items systematically are of inferior quality. Hence this finding suggests that a salient position of a decision alternative is perhaps even more important for explaining choice than the actual quality of the item, and it is an important vindication of the idea that perception precedes (re)cognition.
In sum, Internet data is broadly consistent with the notion of attention-constrained users, and the evidence suggests as a stylized observation that "flashy" or well-positioned links attract attention (clicks). This finding seems to coincide with the earlier mentioned literature on visual attention, which argues that the relative salience of an array of objects determines the speed or chance of their detection (Nothdurft (2000)).

Stylized findings We have summarized evidence from different fields about limited attention and its allocation mechanism. Particularly, the above evidence suggests that

- People have limited attention: They consider only a subset of all information, if exposed to a lot of information.
- The relative salience of an item (e.g. position, lumination, color) is an important determinant of its chance of perception (pop-up effect), and in this respect may outweigh the relevance of the item to a decision-maker

Moreover, it may be of less importance to be the first on some abstract list of alternatives, but what matters is to be among the first few entries (the spotlight effect). In the next sections we will describe how the various economic models have incorporated some of these features.

3 The economics of scarce attention: Causes

As Simon put it: Poverty of attention requires an abundance of information. Thus besides the existence of constraints on how much information can be handled, an economy must produce such an amount of information that these constraints become binding. This can be complicated by the fact that these constraints themselves may be endogenous, e.g. an outcome of some decision process depending e.g. on the presumed quality of the information available.
In this section we review several models that treat the abundance of information, and scarcity of attention, as an endogenous outcome. While limited attention is introduced differently in these models, the underlying causes of information overload are similar, and roughly summarized by the following. Attention is more likely to become a scarce resource if the measure

of active senders (firms) increases, or consumers reduce their willingness to pay attention. The former occurs if market prospects increase, e.g. because consumer wealth, willingness to pay or economic market power from attracting attention increases or information transmission costs decrease. The latter may happen, for example, if the size of the attention set is a part of the (consumer) decision process and attention costs increase or attention revenues to consumers decrease (e.g. because of lower expected product quality).

The main logical reason for why the different models have a very similar prediction of what causes scarce attention is that they all are set up as regime switch models, which essentially reflects the incorporation of limited attention as a capacity constraint. As long as attention is not scarce the main equilibrium information variables (the measure of active senders or transmitted information volume) essentially follow standard theory. Only if overall information exceeds some threshold value the attention economy becomes important. It is beyond this point where the various models begin to differ with respect to their predictions.

Capacity constraints A meaningful notion of scarce attention or, equivalently, of superabundant information, requires a quantification of information items by some measure-theoretic concept. As soon as the information load is quantifiable, limited attention can be introduced formally as a restriction or a threshold on the amount of information that is processed. Regarding capacity measures, the literature can be divided into item-based and volume-based capacity constraints, and we consider both in turn. In sloppy terms the main difference between the two is that in the volume-based approach to limited attention it is less the numbers of market-shouters but rather their aggregate shouting volume that depletes a consumer's mental resources.

3.1 Volume-based capacity constraints

Falkinger (2008) introduces limited attention as a capacity constraint on the perceivable information volume by consumers into a model of non-strategic, symmetric monopolistic competition.[11]

The model Regarding information technology, a firm j is described by the size of its potential audience, i.e. by how many consumers a firm can reach (the firm range), and by their signal strength σ_j. While firm range is an exogenous characteristic, firms must choose their signal strength such to make their messages salient in their attempt to survive the competition for attention. All firms have the same attention technology. The exogeneity of the firm range means that a consumer's information set X_i cannot be

11 The paper considers the case of multi-product firms, but the essential results can be demonstrated already with single-product firms.

influenced by firm activities other than by their decisions to enter or exit the market. Each X_i consists of the same measure of firms (i.e. differentiated products).[12]

Limited attention is introduced as an exogenous threshold level τ_0 on the aggregate volume $\tau = \int_{j \in X_i} \sigma_j$ of all messages received by a consumer. If $j \in X_i$, then j is perceived if and only if j's messages are sufficiently loud, i.e. if and only if $\sigma_j \geq \sigma_{min} \geq 1$. Expressed in terms of attention profiles $(\mathcal{P}(X_i), P)$ the allocation function P can be written as

$$P(A) = \begin{cases} 1 & \text{if } A = \bigcup \{j \in X_i : \sigma_j \geq \sigma_{\min}\} \\ 0 & \text{else} \end{cases} \tag{1}$$

which implies for firm j's chances of perception

$$P_i(j \in A_i) = \begin{cases} 1 & \sigma_j \geq \sigma_{\min} \\ 0 & \text{else} \end{cases} \qquad , j \in X_i$$

Hence if all firms in X_i broadcast sufficiently loudly (with strength $\sigma_{\min}$), then $P(X_i) = 1$, and $P(A) = 0$ for any $A \subsetneq X_i$. This is a difference to models with item-based capacity constraints, where perception naturally is of a stochastic nature. To close the model, the following gating mechanism is specified:

$$\sigma_{\min} = \begin{cases} 1 & \tau \leq \tau_0 \\ \frac{\tau}{\tau_0} & \tau > \tau_0 \end{cases} \tag{2}$$

If aggregate volume τ is below the threshold value τ_0, sending at the minimal signal strength 1 asserts perception, whereas signal strength must exceed 1 if aggregate volume exceeds τ_0. Intuitively, this means that louder shouting is required to be heard as aggregate noise increases. Producing a signal of strength σ costs the firm an amount of $C(\sigma)$. Hence in this model attention costs technically work as an endogenous fixed cost: The height of these costs depends on the current noise level τ, and any firm that can afford these costs will be heard by all consumers in its range. With a measure of T firms, a consumer receives a measure of $M = \phi T$ messages, where $\phi \in (0, 1]$ denotes firm range. As all active firms send at strength σ_{min}, consumer exposure to information is $\tau = \phi T \sigma_{min}$.

Regarding the product market, the paper adopts the symmetric, non-atomistic monopolistic competition model with exogenous consumer wealth $y > 0$ and elasticity parameter ε. Equilibrium profits are given by $\Pi = \frac{\phi \Delta y}{M \varepsilon} - C(\sigma_{min})$.

Causes of an attention economy An economy is *information-poor* if signal exposure satisfies $\tau \leq \tau_0$, otherwise the economy is *information-rich*. The zero-profit condition together with the gating mechanism (2) imply that

12 As utility is symmetric with respect to all products in X, and competition is symmetric and non-strategic it does not matter which products exactly populate the information sets.

an information-rich economy emerges if and only if $\frac{\phi\Delta y}{C(1)\varepsilon} \geq \tau_0$. This inequality reveals that the economy is more likely to be information-rich if i) information technology improves (larger potential market $\phi\Delta$ or lower information (fix)costs $C(1)$) or ii) product market prospects increase (higher per-consumer wealth y or lower demand elasticity ε).

3.2 Item-based limited capacity constraints

In this section we review some contributions, where limited attention is quantified over the measure of information items rather than signal volume. Besides the conceptual difference in the way information (and attention) is measured, the zero-mass assumption in Falkinger (2008) is not completely innocent in a setting with limited attention. In the model of the last section the zero-mass assumption of senders means that important interactions of limited attention with the ability of a firm to strategically set its price, signal strength or other relevant economic attributes, such as quality, are excluded. While zero-mass firms are perhaps reasonable in a macroeconomic world without any explicit attention concerns, this seems more problematic *especially* by the main message of limited attention, namely that perception windows are potentially very small compared to the overall size of alternatives.[13] But this means that perceived firms are not small on a receiver's mind, which makes firm interaction in the perceived market essentially strategic. In fact, there is strong evidence that firms try to exploit consumer browsing behavior. For example, it is known that firms sometimes increase online prices if a particular consumer returns to their website, to give him the impression of economic scarcity. Such behavior can be understood as part of firm obfuscation tactics (see section 5.2), and generally suggests that firms try to price to the "subjective" market as represented in the consumers' heads, and less to the "objective" market. Such arguments emphasize the need to analyze the consequences of limited attention in an oligopolistic (game-theoretic) setting.

The model The following exposition builds on Hefti (2012). For simplicity we set $X_i = X$ throughout this section. A simple way of quantifying the "weight" of a collection of information items, e.g. in terms of mental resources required to recognize ("decode") these items, is to define a measure $r_i : \mathcal{P}(X) \to [0, \infty]$. For $j \in X_i$, $r_i(j) > 0$ is item j's information weight to receiver i. Limited attention means the existence of an upper bound $1 \leq R_i < \infty$ on the quantity of perceivable information. For simplicity we set $r_i = |\cdot|$ (the count measure), and capture heterogeneity in attention capacities by the distribution of $R_i \in \mathbb{N}$ over the population. Note that the threshold level R_i can be introduced as an exogenous constraint (Van Zandt (2004), Anderson and De Palma (2012)) or as part of the consumer decision process (Anderson and De Palma (2009), Hefti (2012)). Given R_i we can de-

13 This is a central emphasis of marketing studies such as Hauser and Wernerfelt (1990).

fine the psychologically feasible set as the collection of all possible attention sets that i can perceive, i.e. $\mathcal{A}_i \equiv \{A \subset X : |A| = R_i\} \subset \mathcal{P}(X)$. We call the pair $(\mathcal{A}_i, P_i)$ an *attention profile*, where P_i is a (possibly degenerate) probability distribution over the attention sets in $\mathcal{A}_i$. Note that if $|X| \leq R_i$ then $\mathcal{A}_i = X$, and the only possible probability function P_i assigns the value 1 to X. This depicts the case where the attention constraint is not strictly binding, and there is no attention allocation problem.[14]

For a given attention profile, chances of perceiving item j are given by $P(j) = \sum_{A \in B_{ij}} P_i(A)$, where $B_{ij} = \{A \in \mathcal{A}_i : j \in A\}$. Together with the demand function d_i, consumer behavior is described by $(\mathcal{A}_i, P_i, d_i)$.[15] Turning to firms we assume that attracting i's attention (joint with $|A| - 1$ competitors) generates a surplus of $V_i^j(A)$. Given the attention profiles $(\mathcal{A}_i, P_i)$ and the value function $V_i^j(\cdot)$ we obtain the following simple structure of firm j's expected earnings:

$$\pi^j = \sum\nolimits_{A \in B_{ij}} P_i(A) V_i^j(A) \tag{3}$$

In case of spare attention capacities ($|X| < R_i$), expression (3) reduces to $\pi^j = V_i^j(X)$.

Anderson and De Palma (2009) and Hefti (2012) consider firm-consumer models, where R_i is endogenously determined by consumers anticipating the expected value of the incoming messages.[16] In these models it depends on the fundamentals underlying the choice process whether or not scarce attention results in equilibrium. Anderson and De Palma (2009) consider the case of exogenously ranked firms (in terms of their profitability) with fixed and equal perception chances, whereas in Hefti (2012) firms must compete in prices on the product market as well as for consumer attention. Regarding the questions, if and when attention constraints are binding, both models broadly offer the same insights, and we concentrate here on Anderson and De Palma (2009) because of its relative simplicity.

Let the measure of firms and consumers be normalized to unity. Firms are distributed over $[0, 1]$, where firm $j \in [0, 1]$ receives an exogenous value $V(j)$ if its products are consumed. The value function $V(j)$ is strictly decreasing in firm type j, which reflects exogenous differences in firm profitability. As this model features no product market competition the transaction value $V(j)$ simply measures the value of attention to that firm. Each

14 Note that standard economic models can be placed in this framework by setting $R_i = \infty$.

15 From $(\mathcal{A}_i, P_i, d_i)$ all objects of interest can be deduced. For example, i's expected demand is $D_i = E[d_i(A) : A \in \mathcal{A}_i] = \sum_{A \in \mathcal{A}_i} d_i(A) P_i(A)$, and $i's$ demand for $j \in X$ corresponds to the j-th projection of D_i.

16 A related spirit is present e.g. in Reis (2006), where the attention decision amounts to choosing the points of time, where a consumer learns the instantaneous state of the economy by paying attention at that time, and thereafter adjust his consumption- or savings plan. As receiving such an update (i.e. being attentive) is costly, a consumer cannot be attentive all the time, but tends to be attentive more frequently if either the explicit cost of receiving an update or the implicit cost of not being attentive (e.g. failing to adjust the consumption plan in a volatile economy) increase.

consumer randomly and independently samples R firms, and acquires a product of each firm in his sample.
In terms of (3), this model satisfies $V^j(A) = V^j \equiv V(j)$ for all $A \in B_j$ (no product market competition), and $P(j) = min\{\frac{R}{n}, 1\}$ for active senders ($P(j) = 0$ otherwise), where n is the measure of active senders[17], and $R \in [0, 1]$ is each consumer's attention span. Hence $P(j)$ simply corresponds to the fraction of (randomly) examined items.
Sending a message costs γ, and profits of active senders are $\Pi^j = P(j)V(j) - \gamma$. Every firm decides whether or not to send a message (e.g. to advertise), and sending more than one message is assumed to be non-profitable.[18] The measure of active senders $n(R)$ is determined by the zero-profit condition $P(n)V(n) = \gamma$. Depending on whether or not attention constraints are binding, n either satisfies $n = n^{max} = V^{-1}(\gamma)$ (if $R \geq n^{max}$) or n solves $\frac{RV(n)}{n} = \gamma$.
Consumers rationally choose their attention span R by weighting marginal benefits against an exogenous, increasing and strictly convex attention cost $C(R)$. Let $s(j)$ quantify the benefits of item j to a consumer. Then $S(n) = \frac{1}{n}\int_0^n s(j)dj$ is the average benefit, given that n firms are active. Assuming independent draws with uniform sampling probabilities gives the net expected benefit of sampling R firms as $EU(R) = RS(n) - C(R)$. For given n the consumers choose to be fully attentive if and only if $S(n) - C'(n) \geq 0$, in which case they set $R(n) = n$. If $S(n) - C'(n) < 0$, then they rationally ignore some available information. Overall, the optimal attention span is determined by $R(n) = min\{n, C'^{-1}(S(n))\}$.

Causes of an attention economy Whether or not there is information congestion (i.e. attention constraints are binding) depends on the firms entry decision $n(R)$ and consumers' choice of attention span $R(n)$, and both decisions are interlinked. As a result attention becomes scarce if $n \geq \tilde{n}$, where $S(\tilde{n}) = C'^{-1}(\tilde{n})$. Hence congestion is more likely to occur, the higher the measure of active senders n, and n is higher if either information costs decline or the revenue distribution $V(j)$ shifts up. On the consumer side we see that higher sampling costs or lower average consumer surplus increases chances of information congestion, because consumers reduce their attention spans.

4 The economics of scarce attention: Consequences

In this section, we track down some of the essential findings regarding the consequences of limited attention for economics. We begin by con-

17 At the same time n is the marginal active type in this model.
18 This is a fundamental difference to Hefti (2012), where firms send multiple messages *because* they must compete for attention.

trasting results from theories of scarce information, particularly informative advertising (section 4.1), but unbounded attention capacities, to those obtained from models with scarce attention. This comparison exemplifies that the two types of scarcities may have fundamentally different implications on how a market or an economy works, which justifies an independent and self-contained theory of attention in economic research. For example, whereas increasing product diversity or cheaper information technology are associated with pro-competitive effects, and improve consumer welfare in case of informative advertising, such conclusions may be exactly reversed if attention becomes the scarce resource (section 4.2). Further, the channel of decreasing cultural distinction in an increasingly globalized world depends on the prevailing type of scarcity in an economy (section 4.3). If attention is a scarce resource that has to be shared among several sectors in an economy, then the sector size distribution depends on how attentive consumers are (section 4.5). We also present a model addressing how market inefficiencies caused by attention externalities could be corrected, and how such correctives might be implemented (section 4.4).

4.1 Scarce information: The case of informative advertising

In models of informative advertising, such as Butters (1977), Shapiro (1980) or Grossman and Shapiro (1984), consumers (information receivers) are ex-ante unaware of the products, and can only learn about a product by receiving ads. Advertising is informative[19] in the sense that it truthfully conveys price and product information, but does not otherwise affect a consumer's evaluation of the products. It is an (implicit) assumption in these models that $A_i = X_i$, attention sets and information sets always coincide or, equivalently, that attention capacities are unbounded ($R_i = \infty$).

If ads are randomly sent to consumers, as is the case in the aforementioned papers, information sets are stochastic and thus usually different among different consumers. In terms of our primitives, this means that while P_i just is the unit function, the information sets X_i are random[20] but depend on the firms' choices of how strong to advertise. Moreover, $X_i \subsetneq X$ can occur if firms do no fully exhaust their advertising capacities, e.g. because of strategic considerations (Grossman and Shapiro (1984)), or due to technological limitations. Hence perception is stochastic in such models, and depends on the information efforts of the competing firms - so is this setting not just a further variant of a stimulus-driven attention model? This clearly is a question of fundamental importance, as it asks if the cause of limited consumer consideration—because there is too little or too much information—matters.

With informative advertising firms have an incentive to increase their infor-

19 If advertising is of a *persuasive* nature, then a consumer's choice function d^i depends directly on actions taken by the firm (e.g. advertising-dependent preferences). See Bagwell (2007) for a comprehensive survey on the different forms of advertising.

20 More formally, informative advertising works through the Φ^i-function from section 1.1.

mation efforts, because they then possibly reach previously bad or even uninformed consumers, from which high information rents can be extracted. In the modern, digitized society consumers are far less likely to have too little information,[21] but limited attention capacities imply that they fail to consider all received messages. Hence firms have a clear incentive to increase their information efforts, not to reach new, unaware consumers, but rather to capture the attention of an overloaded consumer, and thereby possibly inhibit perception by other competitors.

Looking at the main insights from models of informative advertising, a standard comparative-static conclusion is, that cheaper setup or information costs, or larger potential markets, e.g. due to increased information accessibility by consumers, work in a pro-competitive way, and have positive consumer welfare effects by putting downward pressure on prices or increasing equilibrium product diversity.[22] We will see in the next sections that such a conclusion is greatly challenged if attention is the scarce resource. Moreover, if information costs approach zero, the equilibrium approaches the standard case of perfect information. This explains why many have celebrated the introduction and expansion of the Internet as the new supermedium, leading to frictionless commerce. Yet, roughly ten years after the Internet revolution began, many researchers were surprised about the results it produced, as evidence indicated far more price dispersion for apparently homogeneous goods, higher profit margins and generally pricing patterns that seem inconsistent with standard models (see Ellison and Ellison (2005) for a good overview). Intuitively, limited attention and related phenomena such as obfuscation, are important candidates for explaining what seem to be anomalies from the perspective of standard theory. To gauge the potential of an attention theory for filling this gap, it is vital to have a clear understanding of how the patterns of an economy with scarce attention differ from those produced by an economy with scarce information, which is the central topic of the next section.

4.2 Limited attention and oligopolistic competition

Hefti (2012) analyzes the case, where firms strategically compete for attention *and* prices both assuming an abstract model of horizontal product differentiation as well as particular examples, such as the CES or the Ideal-Variety model. Competition for attention is modeled as a contest, where the allocation rule P_i is determined by the efforts to attract attention in a relative way, and attention revenues (i.e. conditional demand) depends on the pricing strategies.

With identical consumers and symmetrically behaving opponents, a firms problem is to choose its prices and attention efforts (p, f) in order to maxi-

21 In fact the case of a completely interlinked digital information society would suggest to set $X_i = X$, as physical locations are of no importance.

22 A similar conclusion holds for conventional consumer search models if search costs decline, see section 5.2.

mize

$$\Pi = \pi \left(f, \bar{f}, n, R\right) V^j(p, \bar{p}, z) - C(f) - F \qquad z = \min\{R, n\} \tag{4}$$

where, $(\bar{p}, \bar{f})$ is an opponent's strategy, $C(\cdot)$ are attention costs, $F > 0$ is an exogenous setup cost and n is the number of active senders. Finally, $\pi \in [0,1]$ is the firm's chance of perception, where $\pi = 1$ if $R \geq n$ (unconstrained attention), and π is zero-homogeneous in $(f, \bar{f})$ (relative salience) if $n > R$. Expression (4) summarizes the interdependence of competition for attention with economic competition. Firms choose their attention efforts, i.e. their perception probabilities in dependence of their economic prospects. These depend on (strategic) pricing, which in turn depends, inter alia, on consumer attentiveness R. As a change in attentiveness R affects both perception chances and pricing if attention is constrained ($R < n$), the relationship between attention-seeking, pricing and consumer attentiveness clearly is non-trivial, thereby highlighting a central difference to the zero-mass setting of Falkinger (2008). Moreover, representing the struggle for attention with a multi-prize contest, where prizes reflect the value of attention, nicely matches the spotlight-feature of attention from section 2, according to which being detected among the first items matters for cognition, but not so much being the first among the detected items.

An important insight of the paper is that if attention rather than information is the scarce resource, the resulting positive and normative equilibrium patterns of the economy can be fundamentally different. In other words the *cause* of why consumers have limited information has first-order consequences for economic outcome.

In a symmetric equilibrium with free entry and endogenous limited attention (i.e. $n > R$) firms choose prices p and attention efforts f such that (4) is maximized. The equilibrium vector (p, f, n) is determined by

$$\begin{aligned} & V_1(p,p,R) = 0 \\ & \pi_1(f,f,n,R)V(p,p,R) = C'(f) \\ & \tfrac{R}{n}V(p,p,R) = F + C(f) \end{aligned} \tag{5}$$

In this setting, attention has two important dimensions: attentiveness (R) and responsiveness. Attentiveness, i.e. the threshold level R, affects i) the (marginal) chance of perception and ii) the conditional rents of (in)attention. Responsiveness pertains to how costly it is for firms to shift attention in their favor. An appealing property of the contest-formulation (4) is that the elasticity of the cost function $C(f)$ captures the responsiveness effect. If e.g. $C(f) = \theta f^{\eta}$, $\eta > 1$, then a larger value of η means that consumers are more resilient against marketing methods. Whether or not inattention rents are competed away by equilibrium forces (whether or not $n'(R) < 0$ or $n'(R) > 0$) depends crucially on attentiveness and responsiveness.[23] The paper shows that inattention rents rather prevail (i.e. $n'(R) < 0$) i) the more

23 This is a major difference to Eliaz and Spiegler (2011a,b), where competition engulfs all attention rents.

competitive a market is (i.e. the more products are perceived as substitutes) or ii) the more consumers are resilient against marketing methods (lower responsiveness).

Limited attention and ideal varieties Important insights on consumer welfare can be obtained by applying the above setting to the well-known circular model of ideal variety. In this model, consumers and firms are uniformly distributed over the unit circle, where a consumer's location can be interpreted as his ideal variety. Consuming at a different location is associated with some inconvenience. Consumer i's utility from consuming j-th product is $U_i = V - p_j - tw_{ij}$, where w_j is the smallest arc distance between i's and j's location, and $t > 0$ quantifies unit transportation costs. Consumer i purchases one unit at the best *perceived* location. The central measure of consumer welfare in this model are average consumer transportation costs $T(R, n)$. The paper shows that equilibrium transportation costs are

$$T(R, n) = t\bar{w} = t \cdot \begin{cases} \frac{n+1}{2n(R+1)} - \frac{1}{4n} & R < n \\ \frac{1}{4n} & R \geq n \end{cases} \tag{6}$$

Hence for $R \geq n$ more aggregate diversity always reduces the average experienced mismatch, because i) perceived and effective markets coincide and ii) firms are located closer together. Grossman and Shapiro (1984) show that this result extends to the case of scarce information (but unbounded attention capacities). In contrast, (6) shows that if and only if attention is scarce ($n > R$) aggregate transportation costs *increase* in the number of active senders n.

The fundamental difference between scarce information and scarce attention, responsible for the above result, is that perceiving some alternatives under capacity limitations necessarily reduces the spare mental capacities for other alternatives. This in turn gives an additional welfare effect, the inferiority effect, which says that if diversity increases, then consumers, on average, pick a worse item compared to their prior choice as long as all items have the same perception chance (which is the case in the symmetric equilibrium). The paper shows that in the ideal variety model this inferiority effect dominates the matching effect, capturing that with increased diversity better products were, in principal, available.[24] In sum, this model shows that a central conclusion of a model with scarce information is exactly reversed with scarce attention.

4.3 Limited attention and the channel of cultural distinction

We now return to the monopolistic competition model by Falkinger (2008), introduced in section 3.1, and present some central consequences of limited

24 Moreover, the paper argues that this negative welfare effect of increased diversity is also present, and potentially even reinforced, if attentiveness R_i is also a part of the consumer's choice problem.

attention in that setting. In this model aggregate diversity of products in the economy is given by the number of active senders, which corresponds to $n^{IP} = \frac{\Delta y}{\varepsilon C(1)}$ in an information-poor (IP) economy, and to $n^{IR} = \frac{1}{\phi}\tau_0$ in an information-rich (IR) economy. The parameters are income y, demand elasticity ε, firm range ϕ and volume threshold τ_0. Hence while traditional fundamentals such as money in the market relative to fixcosts ($\frac{\Delta y}{C(1)}$), or profitability of transactions to firms (quantified by demand elasticity ε) determine equilibrium profits and thus diversity in an IP economy, they matter for the IR case only insofar they are decisive for whether or not an IR economy emerges. Aggregate diversity in an IR economy depends only on (exogenous) information technology (ϕ) and attention threshold τ_0. Looking more at the details of the model shows that this result originates from i) the proportionality of the gating mechanism (2) and, more importantly, ii) the absence of the competitive aspects of limited attention e.g. for pricing decisions, which is a consequence of the zero-mass setting.[25]

One consequence of limited attention, emphasized by the paper, is a change in the channel of cultural distinction as an economy moves from IP to IR. In case of single-product firms cultural distinction is measured by the ratio between the measure of active senders (aggregate diversity) and the measure of individual attention sets. Hence this ratio is ≥ 1, where equality means that every consumer perceives exactly the same set of items, and there is no cultural distinction. In equilibrium, it turns out that in both regimes cultural distinction is given by $\frac{1}{\phi}$, i.e. cultural distinction depends only on the information transmission technology, and decreases if firm range $\phi \in (0, 1]$ increases. Yet there is a difference between IP and IR in what causes this comparative-static result. In an IP economy cultural distinction decreases in firm range because perceived diversity increases (more consumers hear about the firm) but aggregate diversity remains fixed in equilibrium.[26] In an IR economy cultural distinction decreases in ϕ because aggregate diversity declines but, by limited attention, perceived diversity is constant. The reason why aggregate diversity declines is that an increased firm range implies higher consumer signal exposure and thus declining attention. Therefore firms must spend more resources on maintaining their salience levels[27], which means that fewer firms survive the competition for attention.

4.4 Information congestion: Correctives

One commonality of all attention models discussed so far is that the defining property of a modern (information-rich) economy is information overflow, or information congestion. In their paper Anderson and De Palma (2009) (see section 3.2) compare the market outcome (i.e. the number of

25 For example, in the oligopolistic setting of Hefti (2012) variables such as income or demand elasticity also matter for equilibrium diversity if attention is scarce.

26 This occurs because competition is independent of the measure of received messages.

27 The result is a direct consequence of (2).

active firms) with endogenous information congestion to the optimal planner solution.[28] Information congestion (i.e. endogenous limited attention) obviously never is a socially optimal outcome. If the equilibrium is uncongested, then information is underprovided (i.e. n is too low), simply because senders do not internalize receiver benefits. Hence a subsidy on senders solves the problem.[29] What is different if attention is a scarce resource? If information is congested (i.e. $R < n$) especially because sending information is cheap (low γ), then the social optimum involves *reducing* the measure of active senders (by a sender tax) and *increasing* individual attention spans by a receiver subsidy. Intuitively, this occurs because i) the receivers do not internalize the senders' benefits of being heard (hence they are too inattentive), and ii) there are too many senders in the market because individual senders do not take into account their marginal social value. However, the authors note that, with congested equilibria, corrective taxes could also involve receiving *and* sending more (especially if γ is high). Noting that receiver subsidies might be problematic to implement as receivers might falsely claim to have examined messages, the authors investigate the welfare consequences of personal contact prices (want to call me - pay me) and a monopoly access platform. In both cases there is a nuisance cost $\omega > 0$ for any message arriving at the receiver's home (e.g. the ringing of the phone) that is borne by the consumer independent of whether or not he considers the message content. While the monopoly platform ignores ω, both the platform and personal pricing will price out congestion. To see why, note that deciding about a (personal) price means selecting the marginal firm. Suppose that for given R we have $n > R$. In the platform case the platform bears the per-message cost γ and sets a (non-discriminatory) access price equal to $p(n) = \frac{R}{n}V(n)$, i.e. all firms with $\tilde{n} < n$ have positive expected profits and acquire access to the platform. Hence the platform earns $(p(n) - \gamma)n = RV(n) - \gamma n$, and setting $n = R$ leaves the platform with higher profits because of a higher price ($\pi(R) > \pi(n)$) and lower transmission costs. The same reasoning applies to personal prices. While under both mechanisms congestion is priced out, it depends on the parameters, especially γ and nuisance ω, whether or not the platform socially outperforms personal pricing. As is intuitive personal pricing tends to work better than a platform if nuisance costs ω are small, as the platform does not take into account consumer surplus from messages.

4.5 Attentional spillovers between sectors

Anderson and De Palma (2012) consider the case, where several firms from several sectors compete for limited consumer attention. In a nutshell the authors show that attentiveness (R) and the size distribution of the different sectors are correlated, in the direction that less attentiveness tends to-

28 Remember that in their model there is no actual competition for attention as firms cannot influence their chance of perception.

29 Their model by construction excludes a business-stealing effect.

wards magnifying pre-existing differences in sector market-shares. Moreover, limited attention may imply spillover-effects between sectors, as e.g. single-sector profitability shocks affect equilibrium prices in other sectors only by a shift in the attention probabilities.

There is a number $\hat{\Theta} > 1$ of active sectors, each accommodating a measure $n_\theta > 0$ of firms. Within each sector, all active firms produce a homogeneous good at constant unit costs c_θ and, as in Anderson and De Palma (2009) can send a (untargeted) message to a consumer at cost $\gamma_\theta > 0$. Consumers in turn are identical, and have an exogenously fixed attention span $R > 1$, measuring the number of messages examined. The important attentional assumptions of the model are i) firms cannot influence their chance of perception beyond sending a single message (e.g. no salience competition), and ii) each single message has the same chance of perception, independent of which firm or sector it is sent from. As a consequence the number of active firms in a sector n_θ and the number of messages from that sector always coincide, and message costs can be thought of as entry (fixed) costs in a conventional way. In particular, these entry costs do not vary with the measure of messages sent (other than Falkinger (2008)). Moreover, the market share and the information share of a sector $\frac{n_\theta}{N}$, where $N = \sum_{\theta=1}^{\hat{\Theta}} n_\theta$, are identical. Concerning economic competition, the central assumptions are that there are no competitive spillovers between sectors (products are independent), and consumers acquire q_θ units of the cheapest perceived product in each sector, provided that the sector price does not exceed the exogenously fixed reservation price $b_\theta > c_\theta$, where c_θ are the constant sector-specific unit production costs. In a sector, firms decide whether or not to enter (i.e. to send a message), and set a product price (drawn from a sector-specific common price distribution). In terms of our primitives, we have $X_i = X = \bigcup_{\theta=1}^{\hat{\Theta}} J_\theta$, where $J_\theta = [0, n_\theta]$ is the set of all active firms in sector θ and for $R < N$ a consumer's attention set $A \subset X$ satisfies $|A| = R$. Firm j_θ of sector θ sells q_θ units of his product to a consumer, if i) it is perceived and ii) it has the lowest price among perceived within-sector competitors. Formally:

$$\begin{aligned} P(j_\theta \text{ sells}) &= P\left(j_\theta \in A, p_{j_\theta} = \min\{p_{j_\theta^i} : p_{j_\theta^i} \in A\}\right) \\ &= P(j_\theta \in A) \cdot P\left(p_{j_\theta} = \min\{p_{j_\theta^i} : p_{j_\theta^i} \in A\}\right) \\ &= \tfrac{R}{N} \cdot P\left(p_{j_\theta} = \min\{p_{j_\theta^i} : p_{j_\theta^i} \in A\}\right) \end{aligned}$$

where the second equality holds because attracting attention and price competition are independent, and the last equality follows as all items have the same chance of perception. The remaining probability depends on the sector-specific equilibrium price distribution, which is endogenously determined (by sector-wise zero-profit conditions) and depends on attention probabilities (i.e. relative sector size) as well as on exogenous parameters $(b_\theta, c_\theta, \gamma_\theta, q_\theta)$ (sector profitability). To determine the equilibrium sector-specific measure of active firms n_θ in this model, it suffices by the logic of the (sector-wise) mixed-strategy price equilibrium to require the zero-profit condition to hold for the highest possible price $p_{j_\theta} = b_\theta$, as then, given the equilibrium price distribution, no further firm has an incentive to enter the sector. With $p_{j_\theta} = b_\theta$ it is only possible to make a sale if j_θ is the only perceived firm from sector θ, which occurs with probability

$$\begin{aligned} P\left(j_\theta \in A, g_\theta \notin A \,\forall\, g_\theta \neq j_\theta\right) &= P(j_\theta \in A) \cdot \left(1 - P(j_\theta \in A)\right)^{R-1} \\ &\cong \tfrac{R}{N}\left(1 - \tfrac{n_\theta}{N}\right)^{R-1} \end{aligned}$$

where the first equality holds because attention allocation is independent within sectors. To understand the final expression on the right, note that $\frac{n_\theta}{N}$ is the probability of sampling a firm of sector θ in one of the other $R-1$ draws. The final expression underestimates the true probability of being a perceived sector monopolist, because the formula is obtained under the assumption of draws with replacement, but ignores the possibility, that firm j_θ then could be drawn several times. Plugging this probability into the zero-profit condition yields

$$\frac{R}{N}\left(1 - \frac{n_\theta}{N}\right)^{R-1} = \frac{\gamma_\theta}{(b_\theta - c_\theta)q_\theta} \equiv \frac{1}{\pi_\theta} \tag{7}$$

where π_θ can be thought of as a measure of sector profitability. In this model the exogenous sector profitability ties down relative sector sizes for given N, where, ceteris paribus, more profitable sectors have a larger market (and attention) share. Assuming $\pi_1 \geq \pi_2 \geq ...$ expression (7) can be used to find how market shares depend on attentiveness R. First, sector profitability and advertising share are positively correlated, because higher sector profitability increases relative sector size. On the one hand if R increases, then the (ad) market share of the most profitable sector ($\theta = 1$) decreases, whereas the share of the weakest sector increases. Hence more consumer attention tends to equate shares. Intuitively, this occurs because more attentiveness reduces the (conditional) chances of being the only perceived firm of a sector most for the most profitable sector, rendering that sector less attractive. On the other hand if $R \to 1$, then the most profitable sector harnesses almost the entire market. A further particular property of this model is that if information costs are halved, then message volume doubles, leaving both price distributions and the number of active sectors constant. This is driven by the fact that message volume and the measure of senders are equal.[30] If however message costs (profitability) only change for *some* sectors, the equilibrium price distribution of *each* sector is affected. Specifically, an increase of profitability in one sector *decreases* prices in that sector and *increases* prices in the other sectors, in the sense of first-order stochastic dominance. The reason is that attention is a fixed resource commonly utilized by all sectors, meaning that if one sector absorbs more attention

30 If the probability of receiving attention can be smoothly influenced as in Hefti (2012), then a symmetric reduction in attention costs can increase both messages sent at the individual and the aggregate level *without* affecting the equilibrium measure of senders.

(i.e. sends more messages), less resources are available for the other sectors. But the sector sending more messages necessarily experiences more competition, as chances of being the only perceived firm from that sector decrease, which puts downwards pressure on prices.[31] The same argument says that prices tend to increase in the other sectors, because messages of these sectors are crowded out in relative terms by the more active sector. In this model such spillovers occur only because attention is scarce, as sectors otherwise are independent. We will see in section 5.1 that a somewhat related crowding-out logic can be utilized by strategically behaving firms when designing their product lines in order to grab attention and reduce comparison to other brands.

5 Market phenomena related to limited attention

In this section we focus on recent contributions related to attention, where limited attention matters less as a bound on perception in the context of abundant information, but rather plays a role for the possibilities of firms to extract (in)attention rents from boundedly rational consumers.

5.1 Competitive marketing with sticky consideration

Eliaz and Spiegler (2011a) study a model, where each consumer is ex ante allocated to one of two firms (his status quo), but marketing devices can be used to persuade a consumer into considering the alternative product. An important conceptual difference to previously discussed articles, particularly to Falkinger (2008) and Hefti (2012), is that consumer attention capacity is not limited in the sense that consumers never "forget" their status quo product: They either consider only their status quo, or both products. Firms can use marketing devices to make non status quo consumers aware of their product. If a firm succeeds in its attempt, the corresponding consumer has a complete consideration set, i.e. she considers both alternatives when making her choice. What makes the model interesting is that product attributes, e.g. quality, and "pure" marketing features (such as packaging or advertising) can influence a consumer's consideration process.

The model In their model $X_i = X$ is a finite set of a vertically differentiated product. The elements of X can be thought of as different varieties, which are quality-ordered, i.e. $(X, \succ)$ is a linearly ordered set with $\succ$-maximal and $\succ$-minimal elements x^*, x_*. Consumers are identical up to the fact that one half of the unit population has firm j as their status quo product assignment, and the other half initially considers only the product of firm g. Hence $\mathcal{A}_{i(j)} = \{\{j\}, \{j, g\}\}$ is the set of all possible attention sets of a consumer i with status quo j. Both firms simultaneously choose their strategies, where a strategy is a pair (x, m) (called the extended product)

31 Roughly, this is the same rationale for why $p'(R) < 0$ in Hefti (2012).

with $x \in X$ and $m \in M$, where M is a finite set of marketing activities. >From the perspective of a consumer (x^s, m^s) denotes his default extended product, and (x^n, m^n) is the alternative extended product. In their setting, whether or not a consumer can be persuaded to consider the alternative product is deterministic, given the default strategy (x^s, m^s), but the equilibrium qualities and marketing activities are stochastic as firms play mixed strategies. Whether or not a consumer considers the respective alternative product is captured by a consideration function $\phi : (X \times M)^2 \to \{0, 1\}$, where $\phi = 1$ indicates that the alternative product x^n is considered. With strategies $(x^s, m^s), (x^n, m^n)$ perception chances of a consumer i with default firm j are

$$P_{i(j)}(A) = \begin{cases} 1 & A = \{j\},\ \phi(x^s, m^s, x^n, m^n) = 0 \\ 1 & A = \{j, g\},\ \phi(x^s, m^s, x^n, m^n) = 1 \end{cases} \qquad A \in \mathcal{A}_{i(j)}$$

Each consumer selects the $\succ$-maximal product in his attention set. Product quality and marketing are both associated with additive-separable costs roughly satisfying monotonicity, i.e. higher quality and larger marketing campaigns (in the sense that they enable consideration for more values of the default) are more expensive. The model is set up as a competition for market shares, and the cost function satisfies $c(x, m) < \frac{1}{2}$. Under the rational consumer benchmark, i.e. where all consumers ex-ante have *both* products as status quo and $P_i(j, g)) = 1$, the unique equilibrium strategy is to set $(x, m) = (x^*, \emptyset)$, because marketing cannot make products disperceived, and any $x \neq x^*$ can be profitably overbidden. In the rational consumer case each firm earns a payoff $\frac{1}{2} - c(x^*, \emptyset)$.

Offensive marketing The authors discuss essentially two types of marketing activities. The first type takes a consideration function of the form $\phi(x^s, m^n)$. Hence for given own quality x^s, marketing activities of the opponent determine whether or not the alternative product is considered, meaning that advertising is of a purely offensive nature in this setting. In this case the symmetric equilibrium necessarily is mixed. The authors show that i) inferior products are offered with positive probability, ii) firms advertise with positive probability and iii) firms earn the rational consumer benchmark profits. The last observation is in fact a quite robust result under the main assumptions in this setting.[32] Hence market forces compete away any potential gains from (ex-ante) limited consumer considerations.[33]

Salience competition The second type of consideration function satisfies the property that $\phi(x^s, m^s, x^n, m^n) = 1$ if and only if $m^n \geq m^s$. This type of consideration function reflects the idea that advertising may also be used to

32 The authors show that the firms possibly can earn more than the rational benchmark if e.g. there is a small population of perfectly informed consumers, or the consideration function violates being "partitional".

33 This is in stark contrast to Hefti (2012), where profitability under limited attention generically is different from the case of consumers with unbounded attention capacities.

inhibit consumer consideration of the alternative product, and is closer to the idea of salience competition. Under this consideration function, firms still earn the rational consumer profit in the symmetric mixed strategy equilibrium, but the fact that inferior products are offered with positive equilibrium probabilities shows that advertising needs not signal higher quality, contrary to standard results of the advertising literature.[34] Intuitively, this occurs because advertising in this setup jams the opponent's messages, in which cases the consideration set just consists of the default product, meaning that the competitive pressure of high quality is absent.

Consumer conversion The authors further show that if advertising has the above jamming-property, this may affect consumer conversion rates. Under the purely offensive type of consideration function, consumer conversion rates are perfect, meaning that whenever a consumer is induced to consider a new product, he ends up buying it. This extreme result may break down if advertising is both offensive and defensive. Then, while consumers are persuaded into considering the alternative, they might not buy it in the end. To see why, suppose that $x^* \neq y \succ x$ and $m > n$. Because of the latter, a consumer with status quo (y, n) buys y despite considering (x, m), but the reason why (x, m) is offered is that this strategy protects demand from consumers with status quo product x from considering y.

Attention-grabbers and product lines In an extension of their baseline model, Eliaz and Spiegler (2011b) consider the consequences of inertia in consumer attention for product lines, when certain varieties have the only goal of grabbing consumers' attention. Product lines can be thought of as menus, and consumers have well-defined preferences over menus. Formally, X is the set of all possible menu items, and $M \in \mathcal{P}(X)$ is a menu. As before, there are two firms simultaneously choosing their menus M, M', and consumer attention features inertia in the sense that otherwise homogeneous consumers initially are assigned to a status quo firm. Importantly, consumers do not necessarily derive utility from all items in a menu. Some items may be irrelevant to the consumer's well-being, but nevertheless matter for consumer choice as they persuade the consumer into considering the alternative menu. For any menu $L(M) \subset M$ denotes the set of content items in M, i.e. the set of items from M that affect a consumer's utility. In this version of their model the consideration function ϕ is a mapping: $\phi : X \times \mathcal{P}(X) \to \{0, 1\}$, $(x, L(M)) \mapsto \phi(x, L(M))$, i.e. ϕ depends on items rather than on marketing strategies. It should be remarked that this type of consideration function gives a clear offensive touch to attention-grabbing, as a firm cannot use attention-grabbers to prevent a consumer from switching. Moreover, the use of attention-grabbers is assumed to be neutral for status quo consumers, meaning that pure attention-grabbers can never annoy consumers. Menus are quality-ordered in the sense that $M \succ M'$ im-

34 See Bagwell (2007).

plies that $c(L(M)) > c(L(M'))$, i.e. producing a more preferred menu involves more expensive content items. In the absence of inertial attention, i.e. if consumer always consider both menus, this means that $M^* = L(X)$, i.e. the most preferred menu is the only menu offered in equilibrium.
The authors now ask how the equilibrium changes if consumer attention features inertia and certain items may work as pure attention grabbers, i.e. their only function is to make the consumer consider the alternative menu. They show that if i) for any set of content items $L(M)$ there exists an item x such that $\phi(x, L(M)) = 1$, and ii) $\exists M$ such that $M^* \succ M$ but $f(M^*, L(M)) = \{0\}$, then pure attention grabbers are offered in equilibrium with positive probability. Assumption i) means that there always are items (potentially pure attention grabbers) to attract consumer attention to the alternative, and ii) means that even if the contender chooses the most preferred menu, the status quo firm can select a set of inferior content items preventing the consumer from considering the alternative.
The intuition why pure attention-grabbers are used in equilibrium is the following. First, M^* is not a best response to M^*, because ii) assures that the defender can select an inferior (i.e. cheaper) menu but prevent the consumer from considering the alternative, by providing sufficiently many content items. But if j plays $M \prec M^*$, g may have an incentive to play M^* together with attention-grabbers in order to induce a comparison.

It is important to keep in mind that the result on attention-grabbing with zero-utility products was derived under the assumption of a purely offensive nature of attention-seeking, where jamming or reducing the competitor's salience is not possible. Hence attention-grabbing with zero-utility as a tactic can be successful only if at the same time the own menu is superior to the competitor's choice, which explains why the model naturally features only mixed-strategy equilibria (except in the benchmark case of perfectly rational consumers). It also explains why in their model a menu of the form $M \supsetneq M^*$ always is chosen with positive probability (necessarily containing pure attention grabbers). In view of limited attention, one could expect the firms trying to reduce or even abolish comparison of their products with competing products, conditional on being perceived. This line of argument is discussed in the next section.

5.2 Obfuscation

Consumer search Besides the theory of informative advertising, models of consumer search address the problem of scarce information. Perhaps most well-known is the sequential search model by Stahl (1989). Consumers face an exogenous search cost and hold rational expectations about the (equilibrium) price distribution. A number of n identical firms offer a homogeneous product, and simultaneously and non-cooperatively choose their product prices, holding rational expectations about consumer search depth. Searching means sampling a particular firm and learning its price.

Consumers continue to search if the expected benefit of sampling another firm exceeds the cost of doing so, and purchase at the lowest perceived price. The probability to perceive an individual firm is exogenous and equal to $1/n$, but how many firms a consumer samples is potentially complex and depends on his sample history (and his beliefs).
In terms of our primitives, consumer search means that $X_i = A_i$. As with informative advertising $X_i \subset X$ is possible. The difference to informative advertising is that limited consumer information results from incomplete consumer search rather than from incomplete advertising efforts of firms. To make the model interesting, a fraction $\mu \in (0,1)$ of the unit population are costless searchers, i.e. they learn each firm's price for free, the remaining $1-\mu$ consumers must conduct a costly sequential price search.
As firms cannot discriminate between either type of consumer, this generates informational discrepancies between consumers, which - similar to the advertising model of Butters (1977) - induces firms to play mixed pricing strategies. Also, they earn positive (expected) profits, because the costly searchers do not learn all prices. A prediction of this search model (and others) is that if consumer search costs decline (or the fraction of costless searchers increases) this puts downward pressure on prices because the "threat" that consumers continue to search after sampling a high price quota increases.[35] If consumer search costs approach zero, the equilibrium converges to the Bertrand equilibrium, i.e. all firms equate prices with unit costs. Thus lower information costs have pro-competitive effects, which is similar to the conclusion of informative advertising.

Evidence of obfuscation This prediction stands in stark contrast to a recent piece of evidence on estimated e-commerce markups (Ellison and Ellison (2009)). In their empirical study, the authors consider the e-market for computer memory modules on a price comparison website. The website (Pricewatch.com) lists different modules of various companies in an ascending price order. Hence the cheapest modules get the salient top positions. One could be lead to think that such a comparison site warrants a competitive outcome, as search costs for the lowest price are virtually zero, and salience competition naturally puts downwards pressure on prices in this environment. The striking finding of the authors is that while demand for low-quality memory modules is highly price-sensitive, a supplier sells *more* (profitable) medium- and high-quality modules if its low-quality modules occupy a better ranking on the price filter. Hence low-quality memory modules work as an effective loss-leader.
The authors intuitively explain their finding by consumers in fact clicking on the salient (top-list) positions much more frequently, but then the link takes them to the suppliers own website, where the consumer frequently is confronted with several "bargain offers", persuading him into acquiring a higher-quality memory module of the same supplier. The authors esti-

35 In fact the equilibrium is such that a costly searcher searches exactly once.

mate that such an obfuscation strategy of firms implies an effective average markup of about 12%, while a naive application of demand elasticities only in case of low-memory modules (the loss leader) would suggest markups between 3% - 6%. Generally, their findings suggest that both attracting and retaining (or managing) attention are of fundamental importance to modern business - and vital for economists and regulators to understand.

A search-cost model of obfuscation Ellison and Wolitzky (2012) provide a formalization of obfuscation by modifying the sequential search model in Stahl (1989). Obfuscation is modeled as a firm's possibility to increase consumer search costs. Search costs are of the form $g(\tau + t_j)$, where $\tau > 0$ is a firm-independent exogenously fixed time per search and t_j is j's obfuscation level, i.e. the extra time a consumer requires to learn j's price. $g(\cdot)$ itself is strictly convex. Obfuscation times are ex ante unknown to consumers and do not influence the initial sampling chances (but the propensity to search on), and consumers hold rational obfuscation expectations. The authors extensively discuss the case of costless obfuscation.[36] This intuitively matches a situation, where a firm decides not to take efforts to ease the evaluation of their products, e.g. by not attempting to make a product website as informative and transparent as it could be. In this scenario if a consumer samples an obfuscating firm, obfuscation reduces the propensity of consumers to search, because it increases his future search costs by the convexity of $g(\cdot)$. The authors emphasize an intriguing difference between the benchmark model, where obfuscation by assumption is not possible, and the case of costless obfuscation. In both models firms play mixed strategies with respect to prices, and consumers search exactly once. There are two types of price distributions in the benchmark case, depending on whether search costs are high or low. While with high search costs the monopoly price p^m belongs to the equilibrium support and the price distribution does not change if search costs increase further, the upper bound of the support is strictly smaller than p^m with low search costs, because otherwise consumers would continue to search. The possibility to obfuscate annihilates the possibility of this second equilibrium, exactly because firms can prevent this second search under a (slightly) higher price by obfuscating, i.e. increasing consumer search costs. In fact, the equilibrium price distribution with costless obfuscation corresponds exactly to the price distribution of the benchmark model with sufficiently high exogenous search costs *independent* of the general search cost parameter τ. This means that even if general information costs decline, which certainly is the case with new media such as the Internet, this does not automatically foster competition, as firms might be able to avoid competition by obfuscation.

All in all obfuscation strategies are played by firms in this setting as obfuscation increases a consumer's time cost required to learn the price of a

36 The model becomes far less tractable, though also richer in terms of possible equilibrium patterns, if obfuscation is costly for firms.

product, which diminishes his propensity to search on.[37]

6 Rational Inattention

In contrast to the previously encountered models of stimulus-driven limited attention, models of rational inattention concentrate on goal-driven attention allocation processes. As was pointed out in section 1.3, both approaches essentially differ in how the attention allocation function P_i is determined. While P_i is exogenous to the decision-maker in case of stimulus-driven attention but endogenous to the model, models of rational inattention endogenize P_i as an outcome of an attention optimization process.

A narrative example We illustrate the difference between the two approaches with a simple narrative example. Consider an investor who wishes to decide on selling or buying some assets. Every morning he receives a newspaper and must decide how much time he devotes to the entire newspaper, as well as how much of time to allocate to the economics, finance and politics subsections. Increasing overall reading time tends to reduce the likelihood of making mistakes in the investment decision due to ignored or misinterpreted information, but also reduces the time available for other tasks, such as not being late for work. Given that the investor does not read the entire newspaper, there is the problem of how to allocate his reading time over the different subsections. For an investment decision reading the finance section might be most relevant. Nevertheless, reading the politics or general economics section might also be important since topics such as economic policies, decision on warfare, strikes etc. will be discussed there, which could matter for an investment decision as well. Under the presumption that the investor has known his newspaper since long, he might have an ex-ante judgment ("beliefs") about the information potential of each section of the particular redaction in his hands. He then may decide rationally on the allocation of a given amount of time over the subsections, as well as on total reading time, weighting the expected benefits of the optimal reading strategy against the costs of doing something else. In contrast, the stimulus-driven approach to attention would emphasize that fancy headlines, pictures or report framing, may have some impact on what subsections the investor considers, or for how long he reads the newspaper.

37 The authors also explore a version, where obfuscation affects consumers *expectations* about future search costs, given that the search time per product is initially unknown and must be inferred from observed search time and prices. In this version the additional problem of excess obfuscation arises.

6.1 The two-stage decision problem of a totally rational agent

Sims (1998, 2003, 2005) was the first to formalize the goal-driven attention allocation problem, using the entropy from Shannon's information theory as the essential measure of informativeness of a channel.[38] The approach of De Oliviera et al. (2013) adds a decision-theoretic foundation to this literature, which we shall explain briefly.

The model In their setting the agent holds some prior beliefs $\bar{p}$ about the true state of the world $\omega \in \Omega$. An agent can attend to different channels that provide her with further information (her posterior beliefs) about the realized state. Different channels potentially imply different posteriors p. Based on the received information the agent chooses an act out of a menu of feasible acts, i.e. $f \in F$, where f is a mapping associated with the consequence $f(\omega)$ for each state of the world. In our newspaper example, we can think of the different channels as a reading-time allocation over the subsections, while f is the investment decision (e.g. buying or selling) depending on the presumed state of the world. An agent optimally chooses which channels to observe, with the main goal of obtaining the best possible information (posteriors) about the state of the world at minimal information costs.
The entire decision problem of an agent can thus be depicted as a two-stage choice problem. In stage II the agent solves a standard utility maximizing problem, which gives the optimal action f for a given posterior distribution p. The objective of stage I is to choose the optimal channel, i.e. the optimal allocation of her attention, that maximizes her information value, considering also her subjective attention costs.
Formally, the two-stage decision problem can be summarized as

$$\max_{\pi \in \Pi(\bar{p})} \left[\int_{\Delta(\Omega)} \underbrace{\max_{f \in F} \left(\int_{\Omega} u\left(f(\omega)\right) p(d\omega) \right)}_{\text{Stage II}} \pi(dp) - c(\pi) \right] \tag{8}$$

with the utility function on consequences $u : X \to \mathbb{R}$, and $\Pi(\bar{p})$ as the set of all channels for a given prior $\bar{p}$. The subjective information costs $c : \Pi(\bar{p}) \to [0, \infty]$ capture the costs (in terms of utils) associated with each channel.[39]
In this framework the agent is totally rational in the sense that she optimizes both over information acquisition and investment actions. Due to the fact that information is subjectively costly, she will be inattentive to information that is not promising to be valuable.[40]

38 A channel can be thought of as a signal that is correlated with the state of the world, which the decision maker would like to learn

39 In our narrative the costs would depend on the overall reading time and the reading difficulties of the selected subsections.

40 Most of the modeling on rational inattention focused on a variation of the subjective cost

6.2 The entropy model: An example

The rational inattention approach has been applied to diverse macroeconomic subfields such as sticky prices (Sims (2003); Woodford (2009)), differences in the price reactions due to different shocks (Maćkowiak and Wiederholt (2009); Matejka (2012)), understanding the forward discount puzzle of the uncovered interest rate parity condition (Bacchetta and van Wincoop (2006)), business cycles (Maćkowiak and Wiederholt (2009)) and consumption choice with asymmetric responses to wealth shocks (Tutino (2013)). In finance models of rational inattention have been used to study portfolio allocation decision (Peng (2005)), diversification van Nieuwerbrugh and Veldkamp (2010), understanding home bias (Mondria et al. (2010)) or sectoral instead of firm specific learning (Peng and Xiong (2006)). Other papers investigate rational inattention in the context of coordination games (Hellwig and Veldkamp (2009)), or team-production problems (Dessein and Santos (2013)).[41]
Most of the work mentioned above models the information cost function based on the concept of Shannon capacity,[42] sometimes called mutual information I. Mutual information quantifies by how much a signal reduces the uncertainty about the outcome (the entropy) of a random variable. In his seminal article, Sims (2002) introduces limited attention as an upper bound κ on mutual information. We now present a simple example of a beauty contest with rationally inattentive agents as an illustration for a model with bounded Shannon capacities.[43]

Rational inattention in the beauty contest There is a unit measure of agents, indexed by $i \in [0,1]$. Each agent chooses a price $p_i \in \mathbb{R}$ in order to minimize her loss function

$$u(p_i) = -\frac{1}{2}\left(p_i - (1-r)X - r\int p_i di\right)^2 \tag{9}$$

In sloppy terms, an agent wants to choose her price such to minimize the difference to the (random) state of the world and to the average price chosen by the other agents. The parameter $r \in [-1,1]$ quantifies the utility that arises from the average price of all other agents $\bar{p} = \int p_i di$ versus the utility from an unknown state X. Each agent has beliefs about the state of the world and about the behavior of the other agents.
Suppose that X is Gaussian with mean μ and variance σ_X^2. The agent learns about the state of the world in form of a signal $s_i = X + \epsilon_i$, where ϵ_i is a

function $c(.)$, see Veldkamp (2011) or Hellwig et al. (2012) for a discussion on information choice technologies.

41 Veldkamp (2011) is a comprehensive source of further applications.

42 Recently, some authors have used non entropy-based cost functions, as the Shannon capacity implies decreasing marginal cost of information, e.g. Pavan (2013) and Veldkamp (2011).

43 This is a simplified version of Hellwig, Kohls and Veldkamp (2012), also taking advantage of results from Wiederholt (2010).

normally-distributed zero-mean observation error with variances σ_ϵ^2. Thus X and s_i have a multivariate normal distribution with conditional variance $\sigma_{X|s}^2$. The mutual information contained in signal s_i about the random variable X is given by $I(X;s) = \frac{1}{2}log_2\left(\frac{\sigma_X^2}{\sigma_{X|s}^2}\right)$.[44] This quantifies the reduction of uncertainty about the true state of the world by observing signal s_i. Limited attention is an upper bound κ on mutual information, i.e. the capacity constraint is of the form $I(X;s) \leq \kappa$, which means that

$$\frac{\sigma_X^2}{\sigma_\epsilon^2} \leq 4^\kappa - 1. \tag{10}$$

Selecting an attention allocation essentially means deciding on the size of the noise σ_ϵ^2 or, equivalently, choosing the precision of the information about the unknown state X. By doing so, she faces a cost τ for each unit κ of bits. The complete decision problem with rational inattention can be described by:

$$\min_{\kappa\geq 0} E\left(-\frac{1}{2}\left(p_i^* - (1-r)X - r\int p_i di\right)^2 - \tau\kappa\right) \tag{11}$$

with respect to the information capacity constraint (10) and optimal pricing strategy

$$p_i^* = arg\min_{p_i} E\left(-\frac{1}{2}\left(p_i - (1-r)X - r\int p_i di\right)^2 \Big| s_i\right) \tag{12}$$

This problem can be simplified in several steps. At first, following Hellwig, Kohls and Veldkamp (2012) one can represent the expected utility (11) of an agent who acts optimally when choosing the price p_i with the indirect utility function

$$E(u) = -\frac{1}{2}r^2 Var(\bar{p}|s_i) - r(1-r)Cov(\bar{p}, X|s_i) - \frac{1}{2}(1-r)^2 Var(X|s_i) - \tau\kappa \tag{13}$$

When choosing the optimal information structure, the agent is only concerned about reducing the expected variance of his optimal action. Therefore it is enough to consider the implications of the information structure for the conditional variance of the state X, the conditional variance of the average action $\bar{p}$ and the covariance between both. Secondly, the solution to the optimal pricing problem (12) can be derived by assuming that each agents optimal price can be represented as a linear combination of the mean μ of the true state X and the individually received signal s_i

$$p_i = (1-\gamma)\mu + \gamma s_i \tag{14}$$

44 For a detailed derivation see Cover (2002).

implying an expected average action of $\bar{p} = (1-\bar{\gamma})\mu$ with $\gamma = (1-r)\frac{\sigma_X^2}{\sigma_X^2+\sigma_\epsilon^2} + \bar{\gamma}r$ being the weight on the private signal of the agent. For simplicity we consider only symmetric information choices, i.e. in equilibrium the agents will choose the same precisions of the signal and thus the same action rules, which nevertheless allows for ex-post different received signals and therefore different actions. Then, the attention allocation (11) becomes:

$$E(u) = -\frac{1}{2}\left[r\gamma + (1-r)\right]^2 \frac{\sigma_\epsilon^2\sigma_X^2}{\sigma_X^2 + \sigma_\epsilon^2} - \tau\kappa \tag{15}$$

subject to the information constraint (10). Solving this decision problem one gets as optimal bits to process κ^*:

$$\kappa^* = \begin{cases} \frac{1}{2}log_2\frac{\sigma_X^2[r\gamma+(1-r)]^2 ln2}{\tau} & \text{if } \frac{\sigma_X^2[r\gamma+(1-r)]^2 ln2}{\tau} \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

Expression (16) means the following. Suppose that an agent currently pays no attention to the state of the world ($\kappa = 0$). This is not optimal if his marginal utility of attention about X, $\sigma_X^2\left[r\gamma + (1-r)\right]^2 ln2$, is larger than the marginal cost of information processing τ. Then, the agent is going to allocate attention towards observing the real state. The marginal utility of attention depends positively on state variance σ_X^2, which is intuitive as the higher the variance σ_X^2, the more valuable reducing uncertainty about X and avoiding mistakes becomes. It also depends negatively on the strength of the coordination motive r. The more important the actions of other agents are, the less important the precision of her private about X becomes. That is, her optimal κ decreases as our agent keeps track of the real state X with more noise.[45] Allowing for higher noise also means that she will put even less weight on her private signal since $\gamma = \bar{\gamma} = \frac{\sigma_X^2}{\sigma_X^2+\sigma_\epsilon^2}$ when deciding on her optimal pricing (14).

In sum, an increase of the coordination motive means that the agent will rely more on the commonly known prior μ, and allocate less attention to privately learning the true state. However, this tendency towards common information is reinforced by the fact that agents take into account the attention allocation of other agents which, as Hellwig and Veldkamp (2009) show, with complementarity may destroy the standard uniqueness results for this class of games. The discussed example illustrates that the attention allocation reflects the forces of the underlying game, meaning that if an agent wants to coordinate with the action of others, she has to know what the others know. This means that private information about the state of the world is less important.

45 This means a higher σ_ϵ^2 by the capacity constraint (10), which is binding, i.e. $\kappa = \frac{1}{2}log_2\left(\frac{\sigma_X^2}{\sigma_\epsilon^2}\right)$.

7 Conclusion

The allocation of scarce resources is a main theme in economics. In an information-rich society consisting of agents with finite mental capacities, attention naturally becomes a scarce resource, but one that is far less tangible than a conventional monetary budget constraint. Therefore, it is a central task for economics as an academic discipline to explore and research the causes and consequences of this special type of scarcity for market outcome and welfare.

In this article we presented a simple framework describing the allocation of limited attention, and we used this framework to organize and discuss a number of recent economics contributions on attention.

One important separating line is whether attention follows a goal-driven or a stimulus-driven process. Given that constrained optimization has been a central tool used by economists to micro-found the allocation of scarce resources by individual decision-makers, it may be less surprising that the larger part of the literature on attention has focused on goal-driven attention allocation. As is voiced in the corresponding catchword of "rational inattention", this strand of the literature takes into account individual processing capacity limitations and the resulting attention allocation problem, but presumes that what is ignored and what considered is under complete control of the decision-maker. This requires that the agents are aware of what they miss, as only then they can consciously decide not to look at some alternatives. Hence the agents must be endowed with ex-ante information about the values and the likelihoods of the possible alternatives, and it is usually left open by these models where this prior knowledge (which is usually assumed to be "correct" in equilibrium) comes from.

While the goal-driven approach to (in)attention is important in terms of the theoretical benchmark it creates, it rules out the possibility of heteronomous or manipulative influences. Given the substantial evidence from different fields such as Psychology, Marketing or Internet research, documenting that and how external instances influence perception, this seems incomplete at best when trying to understand observed behavior. For example, a standard model of rational inattention has great difficulties to account for obfuscation, since rational agent would learn about such strategies and avoid e.g. misguided information links on the web.

One criticism against abandoning that subjects have the sovereign power over their mental resources is, that one trades a crystal-clear decision machinery against a plethora of possible attention mechanisms (entering into the wilderness of bounded rationality). Nevertheless, the empirical evidence suggests that the alternative road of sender-driven attention allocation needs to be taken, and the models presented in this article show that a rigorous theoretical assessment of the matter is possible. We believe that the existing theories and insights form a substantial starting point, but further research is required to uncover the attentional edge of the digital age.

References

Allenby, Greg M and James L Ginter. 1995. The effects of in-store displays and feature advertising on consideration sets. *International Journal of Research in Marketing*, *12*(1): 67–80.

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

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

Bacchetta, P. and Eric van Wincoop. 2006. Incomplete Information Processing: A Solution to the Forward Discount Puzzle. *American Economic Review*, *88 (4)*: 724–748.

Bagwell, Kyle. 2007. The Economic Analysis of Advertising. In M. Armstrong and R.H. Porter (eds.), *Handbook of Industrial Organization*, Amsterdan: Elsevier B.V., volume 3.

Baye, Michael R., J. Rupert J. Gatti, Paul Kattuman, and John Morgan. 2009. Clicks, Discontinuities, and Firm Demand Online. *Journal of Economics Management Strategy*, *18*: 935–975.

Butters, Gerard R. 1977. Equilibrium Distribution of Sales and Advertising Prices. *Review of Economic Studies*, *44*: 465–491.

Davenport, Thomas H. and John C. Beck. 2001. *The Attention Economy*. Boston, Massachusetts: Harvard Business School Press.

De Oliviera, Henrique, Tommaso Denti, Maximilien Mihm, and M. Kemal Ozbek. 2013. Rationally Inattentive Preferences. *Discussion paper*.

Dessein, Andrea, Wouter; Galeotti and Tanos Santos. 2013. Rational Inattention and Organizational Focus. *Discussion paper*.

Drèze, Xavier and Fred Zufryden. 2004. Measurement of online visibility and its impact on internet traffic. *Journal of interactive marketing*, *18*: 20–37.

Eliaz, Kfir and Ran Spiegler. 2011a. Consideration Sets and Competitive Marketing. *Reviews of Economic Studies*, *78*: 235–262.

Eliaz, Kfir and Ran Spiegler. 2011b. On the strategic use of attention grabbers. *Theoretical Economics*, *6*: 127–155.

Ellison, Glenn and Sara Fisher Ellison. 2005. Lessons About Markets from the Internet. *Journal of Economic Perspectives*, *19*: 139–158.

Ellison, Glenn and Sara Fisher Ellison. 2009. Search, Obfuscation and Price Elasticities on the Internet. *Econometrica*, *77*: 427–452.

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

Fader, Peter S. and Leigh McAlister. 1990. An Elimination by Aspects Model of Consumer Response to Promotion Calibrated on UPC Scanner Data. *Journal of Marketing Research*, *27*: 322–332.

Falkinger, Joseph. 2007. Attention economies. *Journal of Economic Theory*, *133*: 266–294.

Falkinger, Joseph. 2008. Limited Attention as the Scarce Resource in an Information-rich economy. *The Economic Journal*, *118*: 1596–1620.

Ghose, Anindya M. and Sha Yang. 2009. An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets. *Management Science*, *55*: 1605–1622.

Grossman, Gene M. and Carl Shapiro. 1984. Informative Advertising with Differentiated Products. *The Review of Economic Studies*, *51*: 63–81.

Hauser, John R. and Birger Wernerfelt. 1990. An Evaluation Cost Model of Consideration Sets. *The Journal of Consumer Research*, *16*: 393–408.

Hefti, Andreas M. 2012. Limited Attention, Competition and Welfare. *ECON working paper No. 28*.

Hellwig, Christian, Sebastian Kohls, and Laura Veldkamp. 2012. Information Choice Technologies. *American Economic Review Papers and Proceedings*, *102 (3)*: 35–40.

Hellwig, Christian and Laura Veldkamp. 2009. Knowing What Others Know: Coordination Motives in Information Acquisition. *The Review of Economic Studies*, *76*: 223–251.

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

Jansen, Bernard J., Amanda Spink, and Tefko Saracevic. 2000. Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing and Management*, *24*: 461–476.

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

Maćkowiak, Bastosz and Mirko Wiederholt. 2009. Optimal Sticky Prices under Rational Inattention. *American Economic Review*, *99 (3)*: 769–803.

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y. Ozbay. 2012. Revealed attention. *The American Economic Review*, *102*(5): 2183–2205.

Matejka, Filip. 2012. Rationally Inattentive Seller: Sales and Discrete Pricing. *Discussion paper*.

Maunsell, John R. and Stefan Treue. 2006. Feature-based attention in visual cortex. *TRENDS in Neurosciences*, *29*: 317–322.

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

Mondria, Jordi, Thomas Wu, and Yi Zhang. 2010. The determinants of international investment and attention allocation: Using internet search query data. *Journal of International Economics*, *82*(1): 85–95.

Mozer, Michael C. and Mark Sitton. 1998. Computational modeling of spatial attention. In H. Pashler (ed.), *Attention*, Psychology Press.

van Nieuwerbrugh, Stijn and Laura Veldkamp. 2010. Information Acquisition and Under-Diversification. *The Review of Economic Studies*, *77 (2)*: 779–805.

Nothdurft, Hans-Christoph. 2000. Salience from feature contrast: additivity across dimensions. *Vision Research*, *40*: 1183–1201.

Pan, B., H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka. 2007. In Google We Trust: Users Decisions on Rank, Position, and Relevance. *Journal of Computer-Mediated Communication*, *12*: 801–823.

Pashler, Harold E. (ed.). 1998. *Attention*. Psychology Press.

Peng, Lin. 2005. Learning with Information Capacity Constraints. *Journal of Financial and Quantitative Analysis*, *40*(02): 307–329.

Peng, Lin and Wei Xiong. 2006. Investor attention, overconfidence and category learning. *Journal of Financial Economics*, *80*: 563–602.

Reis, Ricardo. 2006. Inattentive Consumers. *Journal of Monetary Economics*, *53*: 1761–1800.

Shapiro, Carl. 1980. Advertising and welfare: Comment. *The Bell Journal of Economics*: 749–752.

Simon, Herbert A. 1955a. A Behavorial Model of Rational Choice. *The Quarterly Journal of Economics*, *69*: 99–118.

Simon, Herbert A. 1955b. Designing Organizations for an Information-rich World. In M. Greenberger (ed.), *Computers, Communications and the Public Interest*, Baltimore: John Hopkins Press.

Sims, Christopher A. 1998. Stickiness. *Carnegie-Rochester Series on Public Policy*, *49 (1)*: 317–356.

Sims, Christopher A. 2003. The Implications of rational inattention. *Journal of Monetary Economics*, *50*: 665–690.

Sims, Christopher A. 2005. Rational Inattention: A research agenda. *Discussion paper Series 1 / Volkswirtschaftliches Forschungszentrum der Deutschen Bundesbank*.

Smith, Michael D. and Erik Brynjolfsson. 2001. Consumer Decision-Making at an Internet Shopbot: Brand still matters. *Journal of Industrial Economics*, *49*: 541–558.

Stahl, Dale O. 1989. Oligopolistic pricing with Sequential Consumer Search. *American Economic Review*, *79*: 700–712.

Tutino, Antonella. 2013. Rationally inattentive consumption choices. *Review of Economic Dynamics*, *16*: 421–439.

Van Zandt, Timothy. 2004. Information overload in a network of targeted communication. *RAND Journal of Economics*: 542–560.

Veldkamp, Laura. 2011. *Information Choice in Macroeconomics and Finance*. Princeton: University Press.

Woodford, Michael. 2009. Information-Constrained State-Dependent Pricing. *Journal of Monetary Economics*, *56 (1)*: 100–124.

Yantis, Steven. 1998. Control of Visual Attention. In Harold E. Pashler (ed.), *Attention*, Psychology Press.