

# Attention Manipulation and Information Overload\*

Petra Persson<sup>†</sup>

September 2017

## Abstract

Limits on consumer attention give firms incentives to manipulate prospective buyers' allocation of attention. This paper models such attention manipulation and shows that it limits the ability of disclosure regulation to improve consumer welfare. Competitive information supply, from firms competing for attention, can reduce consumers' knowledge by causing information overload. A single firm subjected to a disclosure mandate may deliberately induce such information overload to obfuscate financially relevant information, or engage in product complexification to bound consumers' financial literacy. Thus, disclosure rules that would improve welfare for agents without attention limitations can prove ineffective for consumers with limited attention. Obfuscation suggests a role for rules that mandate not only the content but also the format of disclosure; however, even rules that mandate "easy-to-understand" formats can be ineffective against complexification, which may call for regulation of product design.

Keywords: Information Overload, Complexity, Persuasion, Disclosure regulation, Consumer Protection, Salience

JEL Codes: D11, D14, D18, D83

---

\*I am grateful to Navin Kartik, Doug Bernheim, Patrick Bolton, Yeon-Koo Che, Pierre-André Chiappori, Matt Gentzkow, Takakazu Honryo, Samuel Lee, Uliana Loginova, Florian Scheuer, Cass Sunstein, and to seminar participants at the Consumer Financial Protection Bureau and at various universities. I also thank the faculty and participants of the Russell Sage Foundation Summer Institute in Behavioral Economics.

<sup>†</sup>Department of Economics and SIEPR, Stanford University; CEPR; NBER; Research Institute for Industrial Economics. Email: perssonp@stanford.edu.

# 1 Introduction

Governments can increase welfare by addressing market failures and providing public goods. In addition, they often try to make constituents better informed, hoping to raise welfare by inducing consumers to make better decisions. For example, government agencies such as the Food and Drug Association (FDA) and the Consumer Protection Agency (CPA) disseminate recommendations and guidelines. In addition, firms are increasingly required to disclose information that is deemed relevant to consumers, so credit card companies must disclose important details about the products they offer and food producers must adhere to labeling regulations.

Yet processing this information requires consumers' attention. If attention is unlimited, more information is weakly better, so any regulation that mandates disclosure cannot do harm. There is growing evidence that consumers' attention is limited, however (see, e.g., Chetty et al. (2009); Dellavigna and Pollet (2009); Abaluck and Gruber (2011)), which makes the welfare consequences of regulation that mandates disclosure, a priori, less clear, as consumer attention devoted to one piece of information may crowd out attention to others. The literature on rational inattention analyzes how a decision-maker with limited attention allocates it across various *passive* sources of information (e.g., Sims, 2003; Wiederholt, 2010). The core idea in this paper is that consumers' limited attention gives information providers, such as firms, incentives to be *active*. To be more precise, when a consumer's attention is limited, her ultimate purchasing decisions may hinge on what she pays attention *to*; this, in turn, incentivizes firms to engage in *attention manipulation*, that is, strategic actions to influence how she allocates her attention. This is distinct from, and operates on top of, any incentive to manipulate the substance of consumer-facing communication.

The first part of the paper shows that, in the presence of attention manipulation, competitive information supply, from firms competing for attention, can reduce consumer knowledge by causing information overload. In the second part of the paper I show that a single information provider, such as a firm mandated to disclose information, may deliberately *induce* information overload to conceal information. Thus, requiring a firm to disclose all hidden, undesirable features of its product may have no impact on social welfare; the firm will simply disclose these features along with an avalanche of irrelevant information. But if full disclosure policies can backfire, intuition suggests that there is an easy fix: simply to mandate not only *what*

firms disclose, but also *how* it should be disclosed. The paper’s third result is a disconcerting one, however: Mandating that firms provide easy-to-understand information about the products they sell will induce “complexification” of the underlying products themselves. This is essential as it limits the potential welfare gains from mandated information provision; in fact, complexification can eradicate all welfare gains. Together, these findings demonstrate that taking attention manipulation into account has important implications for the design of consumer protection regulation, by suggesting a role for rules that restrict communication, mandate not only the content but also the format of disclosure, and regulate product design.

These findings arise in a theoretical framework of information exchange between one consumer (the decision-maker; henceforth the “DM”) and various suppliers of information (firms; henceforth “experts”). To capture attention manipulation, I want to permit the experts to be *active*, in the sense that they make persuasion effort choices that, in turn, affect the optimal attention allocation of the DM. I depart from the framework by Dewatripont and Tirole (2005), who model a DM interacting with a *single* expert. The DM considers taking an action with an uncertain payoff to her, but that would surely benefit the expert (e.g., buying a good from the expert). Before the DM decides, she can communicate with the expert. Communication is a moral-hazard-in-team problem: The more attention the DM pays to the expert, and the greater the expert’s effort, the more likely it is that information is exchanged successfully between them. Dewatripont and Tirole (2005) refer to this as *issue-relevant communication*, because it concerns the DM’s actual benefits from the action. Importantly, the expert does not know what conclusion the DM will draw from the information he provides; he only knows that it may affect her decision.<sup>1</sup> In addition, Dewatripont and Tirole (2005) include a pre-play stage in which *cue communication* takes place. This does not concern the actual benefits associated with the action, but rather the decision’s ex ante appeal, that is, the likelihood that (issue-relevant communication will show that) the action is beneficial.

---

<sup>1</sup>Intuitively, a firm representative selling a product can expend effort to convey more information about the product to the buyer, without knowing with certainty whether it will make the prospective buyer conclude that her payoff from the product will be positive (and choose to buy) or negative (and choose to abstain). Thus, information provided by the seller is always truthful in this setting; put differently, no vendors are behaving in an unlawful manner – their choice regarding how much persuasion effort to make simply concerns how much effort to make to clearly and truthfully convey additional information about the product to the buyer. As we shall see, *even though all communication is required to be truthful*, this does not guarantee that a seller lacks strategies for effectively hiding information that he believes would reduce the consumer’s willingness to purchase.

To analyze attention manipulation, I introduce multitasking into this framework. A single DM (she) considers several binary actions, and can communicate with one distinct expert (he) on each of them. Each of these issue-relevant information exchanges is a moral-hazard-in-teams problem, and the DM now faces a multitasking problem because she must divide her limited attention between the various experts. In the first part of Section 2, I show that, in this framework, attention substitution leads to externalities that I refer to as attention *crowding out*: Each expert ignores the effects of his chosen persuasion effort on the attention that the DM devotes to other experts. Interestingly, an expert benefits or suffers from crowding out depending on whether he believes that attention from the DM will raise or lower the likelihood that she takes the action from which he benefits. This makes an expert’s expected payoff non-monotonic in the appeal of other experts’ proposed actions.

In the second part of Section 2, I allow each action’s ex ante appeal to be unobserved by the DM, and add a stage before issue-relevant communication in which cue communication can take place. Specifically, each expert can, at a cost, send hard information about his proposed action’s appeal to the DM, and she can process this cue at a cost. Intuitively, in the presence of competition between experts for the DM’s attention, cue communication takes place first and helps the DM select which experts (actions) to devote attention to in the second stage. Cue communication thus shapes the set of actions on which the DM ends up deliberating. For this reason, I say that cue communication takes place in the “selection stage” and issue-relevant communication in the subsequent “deliberation stage.”

I analyze how the DM’s welfare changes as the cost of sending cues—or proposing actions to the DM—falls and, consequently, more experts seek attention and more choices enter the picture. Initially, she benefits from the fact that she has more actions from which to choose. But as entry becomes cheaper, it becomes profitable for experts who propose less-appealing actions to enter. As a result, the average quality of the proposed actions deteriorates, and the DM must read cues, at a cost, to find the attractive ones. Eventually, as the supply escalates, screening ceases to be worthwhile to her, and she picks proposed actions for deliberation randomly.

Thus, at a certain point, as the competition for the DM’s attention increases and she gets more information, she processes less of it—or tunes out—and fares worse. I refer to this as *information overload*. Its immediate cause is that the quality of the proposed actions decreases with the quantity; actions worthy of deliberation become the proverbial needle in a haystack. The deeper cause, though, is negative

externalities: Entry is individually rational for each expert, even as it complicates the selection problem for the DM and spoils overall communication. A DM with limited attention may hence want to limit access to her attention space, even if that reduces her choice set. She faces a trade-off between comprehensiveness and comprehensibility.

Section 3 reinterprets the framework to capture a DM who communicates with a single expert on one action that has several *aspects*. In the deliberation stage, the DM’s multi-tasking problem now stems from the fact that the DM must decide how to allocate her scarce attention across the various aspects of the action. Recall that, in the selection stage, in the “multiple experts”-setting analyzed in Section 2, the DM is unsure of each action’s ex ante appeal. The analogue in Section 3 is that the DM is unsure which aspects of the (single) action are financially relevant, and hence worth devoting attention to in the deliberation stage.

The analysis in Section 3 picks up on the information overload result from Section 2 and shows that a single expert may induce this outcome. Specifically, in Section 2, information overload resulted from various experts’ competition for attention. Section 3 shows that a single expert, who faces no competition for the DM’s attention but who can communicate with the DM about multiple aspects of the action for which he advocates, may *strategically induce* information overload in the DM. This arises when the expert wants to divert attention from relevant aspects that are “unfavorable” in the sense that, if the DM learns more about those aspects, she may *not* take the action. Of course, if it were up to the expert, he would not bring up any such unfavorable aspect – that is, he would not send any cue about it. But sometimes the expert cannot or may not withhold such information, e.g. due to laws that mandate disclosure of relevant information. I show that, in response to such a disclosure mandate, the expert chooses to inundate the DM with cues relating mostly to irrelevant aspects of the action. This induces information overload, which, in turn, effectively conceals the inconvenient aspects that the firm was mandated to disclose in the first place. In other words, the expert shares superfluous information to strategically induce information overload, which essentially *obfuscates* the DM. When consumers’ attention is limited, simple disclosure rules can thus be completely ineffective and have no impact on social welfare.

But if disclosure mandates can backfire because firms can induce information overload, intuition suggests an immediate solution: to mandate not only what firms disclose, but also that it be disclosed in an easy-to-understand way. The second part

of Section 3, however, shows that requiring a firm to provide easy-to-understand information about a product it sells can induce a “complexification” of the underlying product itself. I show this by extending the model with a single expert to allow the DM’s payoff from the (single) action to be comprised of many components, and to allow the expert to manipulate that composition so long as the total payoff stays constant. Intuitively, such payoff-equivalent variations amount to changing the number of financially relevant aspects of the product. This gives the expert yet another tactic with which to thwart learning: The expert can force the DM to understand more details of the action, or product, to grasp its total payoff; in other words, he can make it more *complex*. Complexification ensures that an increasing amount of relevant information slips the DM’s attention and, by the same token, that whatever she *can* learn in the deliberation stage is so trivial that it no longer affects her decision. In a nutshell, even if she fully understands all the aspects on which she can deliberate, she will always do what she would have done anyway. Complexification thus has same welfare consequences as inducing information overload — it can eradicate all intended welfare gains from mandated information provision. But complexity is a more delicate issue for regulation: Unlike strategic information overload, it cannot be tackled at the level of communication, information, and disclosure; it may call for intervention in product design.

This paper builds on and contributes to several strands of the literature. First, I advance recent work on two-sided communication as a moral-hazard-in-teams problem, where the “softness” of information is intermediate and endogenous (Dewatripont and Tirole, 2005), by introducing multiple experts that vie for a DM’s attention.<sup>2</sup> Competition for attention leads to attention substitution, which in turn invites attention manipulation. More generally, this relates to a number of studies that examine how a DM communicates with multiple experts or with a single expert on multiple topics. Krishna and Morgan (2001), Battaglini (2002), and Ambrus and Takahashi (2008) study competing experts in a soft information setting. In contrast, Milgrom and Roberts (1986) and Gentzkow and Kamenica (2017) study multiple experts in a hard information setting, and Chakraborty and Harbaugh (2007, 2010)

---

<sup>2</sup>Soft information can be misrepresented at no cost (Crawford and Sobel, 1982); hard information can be withheld but not misrepresented (Grossman, 1981; Milgrom, 1981). When communication is a moral-hazard-in-teams problem, the softness is intermediate: Communication conveys hard information with a probability that depends on effort by both sides; otherwise, information remains soft. Introducing a lying cost represents another way to bridge soft and hard information (see, e.g., Kartik et al., 2007; Kartik, 2009). Also see Caillaud and Tirole (2007).

study soft communication between a DM and one expert on several topics. In these papers, competition or multiplicity typically increases the amount of knowledge the DM gains.<sup>3</sup> Kartik et al. (2017) endogenize information acquisition in a hard information setting with multiple experts and show that adding more experts can reduce the DM’s welfare. The mechanism driving this result – that adding experts reduces each expert’s incentive to acquire costly information – is, however, distinct from information overload. In my setting, there is no information acquisition; instead, central to my finding that more information can reduce a DM’s knowledge is that experts manipulate not only the substance of communication but also the DM’s attention allocation. This suggests that limits to attention are important in determining whether individuals stand to benefit from a more competitive, or greater, information supply.<sup>4</sup>

Second, my focus on attention limitations relates to recent work in the industrial organization literature with attention-constrained consumers. Spiegel and Eliaz (2011a,b) and de Clippel et al. (2014) analyze a setting where consumers pay attention only to a subset of available options and firms can manipulate this “consideration set” through marketing or price setting. Hefti (2017) studies a similar but more general setting with imperfect price competition and derives a mechanism akin to the information overload result presented in Section 2 (and in Persson (2012)), suggesting that this result is robust to various modeling assumptions – so long as consumers’ attention is limited. The two key results presented in Section 3 relate to Carlin (2009), Wilson (2010), and Ellison and Wolitzky (2012), whose papers model obfuscation or complexification as a strategic choice firms make to raise search costs in settings with optimal consumer search.<sup>5</sup> In these papers, firms’ incentives to raise

---

<sup>3</sup>Gentzkow and Kamenica (2017) provide a general framework of costless and publicly observed information acquisition by multiple agents with flexible information structures. They show that competition may increase or decrease information, but provide conditions under which a higher number of experts cannot provide less aggregate information in the sense of Blackwell (1951).

<sup>4</sup>In these models, all the experts have information relevant to the same action. In contrast, I analyze a case in which each expert has information about a different action. This relates to studies on organizational design in which multiple division managers communicate local information to the central management (Dessein and Santos, 2006; Alonso et al., 2008). These studies, however, address neither limited attention on the part of the DM nor competition for that attention.

<sup>5</sup>In settings where multiple firms sell a homogenous good and compete on price, Wilson (2010) and Ellison and Wolitzky (2012) ask why it is individually rational for firms to raise consumer search costs, and Carlin (2009) focuses on how this affects market prices. Other papers that discuss obfuscation or related mechanisms in the context of competitive price discrimination models include Ellison (2005), Gabaix and Laibson (2006), Spiegel (2006), Inderst and Obradovits (2015), and BJORKEGREN (2016).

these search costs stem from competition between firms. The present paper shows that obfuscation and complexification can be individually rational even when a single firm operates in the absence of competition, if it is subjected to disclosure regulation. In addition, the paper makes a distinction between obfuscation – which alters communication about a product without altering the product itself – and complexification – which affects the product characteristics per se. While the consumer welfare consequences of these phenomena are similar, the distinction has crucial implications for optimal disclosure regulation.

Third, Bordalo et al. (2012, 2013) model the attributes *to which* an individual’s attention is drawn when it is limited: attention is unproportionally allocated to salient issues. I instead focus on how, when individuals have limited attention, market participants’ strategically act *to make* a certain attribute of a good salient or invisible, depending on whether the market participant wants to conceal or emphasize the attribute. Put differently, I allow interested parties to influence the relative salience of a product’s attributes. This relates to Bordalo et al. (2016) and Manzini et al (forthcoming) who study competitive markets where the salience of goods’ characteristics is endogenously determined.

## 2 Many Experts Competing for Limited Attention

In this section, I introduce the multiple expert setting and analyze, in turn, the deliberation stage and the selection stage.

### 2.1 The deliberation stage

**Set-up** I introduce multiple senders, or experts, into the framework proposed by Dewatripont and Tirole (2005). A DM faces two simultaneous decisions,  $i = 1, 2$ .<sup>6</sup> Each decision  $i$  concerns whether to take a distinct action,  $A_i$ . For each decision, there is a distinct expert who gets a deterministic payoff  $d > 0$  if the DM takes the action, and zero otherwise. The DM’s payoff  $\tilde{x}_i$  from  $A_i$  takes the value  $\bar{x} > 0$  with probability  $\alpha_i$  and otherwise the value  $\underline{x} < 0$ . The probability  $\alpha_i$  is common knowledge. The larger the  $\alpha_i$ , the more attractive  $A_i$  seems to the DM, and the more aligned are her interests with those of the expert vested in  $A_i$ . Everyone is

---

<sup>6</sup>It is not essential that the decisions be simultaneous, only that communication about both decisions is simultaneous.



risk-neutral. In the absence of additional information, the DM takes  $A_i$  if and only if its expected payoff is positive:  $\alpha_i > \alpha^* \equiv \frac{-\underline{x}}{\bar{x}-\underline{x}}$ .<sup>7</sup>

Before making any decision, the DM can learn more about the actions. For each action, the vested expert can provide information, and the DM can devote attention to processing this information. Through such communication, the DM can learn the realization of  $\tilde{x}_i$ . The expert himself knows neither whether  $\tilde{x}_i = \bar{x}$  or  $\tilde{x}_i = \underline{x}$  nor whether the (truthful) information he provides will help the DM find out. Nevertheless, his information may persuade the DM to take  $A_i$  even though  $\alpha_i < \alpha^*$ , since the DM may find out that  $\tilde{x}_i = \bar{x}$ . The probability that the DM learns  $\tilde{x}_i$  is given by  $p(s_i, r_i)$ , where  $s_i$  and  $r_i$  are, respectively, the expert's effort to communicate about  $A_i$  and the attention that the DM devotes to learning about  $A_i$ .

**Assumption** The function  $p(s_i, r_i)$  is twice continuously differentiable on  $[0, 1]^2$ , with  $p(0, 0) = 0$  and  $p(1, 1) = 1$ . It is strictly concave and satisfies  $p_1(\cdot) > 0$ ,  $p_2(\cdot) > 0$ ,  $p_{12}(\cdot) > 0$ , and the Inada condition  $\forall s_i \in [0, 1]$ ,  $p_2(\cdot) \rightarrow 0$  as  $r_i \rightarrow 1$  and  $p_2(\cdot) \rightarrow \infty$  as  $r_i \rightarrow 0$ .<sup>8</sup>

Successful communication is more likely the more effort the expert devotes to persuasion ( $p_1(\cdot) > 0$ ) and the more attention the DM devotes to his message ( $p_2(\cdot) > 0$ ). In short, communication is a team effort. Because communication efforts are complements ( $p_{12}(\cdot) > 0$ ), an expert's return from expending effort is higher when the DM listens more attentively, and the DM's return from paying attention is higher when the expert makes a greater effort to explain. The formulation encompasses communication technologies with the property  $p(0, r_i) \neq 0$ : Even if an expert makes no effort to transmit information to the DM, it is possible for her to find the relevant information by herself.<sup>9</sup>

Communication is costly to both parties. The DM's attention is scarce,  $\sum_i r_i \leq 1$ , so the cost is attention substitution: Paying more attention to one action necessarily comes at the expense of others. An expert's cost of persuasion effort is given

---

<sup>7</sup>W.l.o.g., I assume that she does not choose  $A_i$  when  $\alpha_i = \alpha^*$ . Dewatripont and Tirole (2005) refer to this as *supervisory* decision-making, which they distinguish from *executive* decision-making, whereby the DM chooses action  $A_i$  only if she is certain that  $\tilde{x}_i = \bar{x}$ . Intuitively, executive decision-making may capture the DM's behavior when the stakes are so high that it is prohibitively costly for her to make "the wrong" decision ( $\underline{x} = -\infty$ ). Because executive decision-making corresponds to the limiting case when  $\alpha^* \rightarrow 1$ , my analysis of supervisory decision-making when  $\alpha_i \leq \alpha^*$  characterizes the results under executive decision-making.

<sup>8</sup>The subscripts refer to the derivative of a function with respect to the  $i$ th argument.

<sup>9</sup>Dewatripont and Tirole (2005) study the particular complementary technology  $p(s_i, r_i) = s_i r_i$ , and thus do not allow for the DM to find the relevant information by herself.

by  $c(s_i)$ .

**Assumption** The function  $c(s_i)$  is twice continuously differentiable on  $(0, 1)$  and satisfies  $c'(\cdot) > 0$  and  $c''(\cdot) > 0$  as well as the Inada conditions  $c'(\cdot) \rightarrow 0$  as  $s_i \rightarrow 0$ , and  $c'(\cdot) \rightarrow \infty$  as  $s_i \rightarrow 1$ .

The persuasion efforts and the attention allocation are chosen simultaneously and non-cooperatively. I refer to the above game as the deliberation stage.

I determine the Nash equilibrium for this game, and then analyze how the experts affect each other in equilibrium.

**Lemma 1.** *If  $\alpha_1, \alpha_2 \leq \alpha^*$ , there is a unique equilibrium  $(r_1^*, s_1^*, s_2^*)$ , which is interior. If  $\alpha_1 \leq \alpha^* < \alpha_2$ , there is a unique equilibrium  $(r_1^*, s_1^*, 0)$ . If  $\alpha_1, \alpha_2 > \alpha^*$ , there is a unique equilibrium  $(r_1^*, 0, 0)$ .*

An expert's behavior hinges on whether the DM uses an *opt-in rule* or an *opt-out rule* for the decision in which he is vested. When  $\alpha_i \leq \alpha^*$ , the DM uses an opt-in rule with respect to  $A_i$ . Her default is not to take  $A_i$ , but she departs from this default—opts in—if she learns that  $\tilde{x}_i = \bar{x}$ . Hence, expert  $i$  has an incentive to communicate with her. Thus, when  $\alpha_1, \alpha_2 \leq \alpha^*$ , each expert solves

$$\max_{s_i} \{d\alpha_i p(s_i, r_i) - c(s_i)\},$$

and the DM's problem is

$$\max_{r_1 \in [0,1]} \{\bar{x} (\alpha_1 p(s_1, r_1) + \alpha_2 p(s_2, 1 - r_1))\}. \quad (1)$$

In the unique equilibrium, both experts communicate, and the DM pays attention to both. The Inada condition rules out corners; global concavity of  $p(s_i, r_i)$  guarantees uniqueness.

In contrast, when  $\alpha_i > \alpha^*$ , the receiver uses an opt-out rule with respect to  $A_i$ . Her default is to take  $A_i$ , but she departs from this default—opts out—if she learns that  $\tilde{x}_i = \underline{x}$ . Because communication can only persuade the DM *not* to take  $A_i$ , expert  $i$  makes no effort. The DM can nevertheless devote attention to  $A_i$ ; that is, she can engage in (one-sided) information acquisition. If  $\alpha_1 \leq \alpha^*$ ,  $\alpha_2 > \alpha^*$ , her problem is

$$\max_{r_1 \in [0,1]} \{\bar{x} \alpha_1 p(s_1, r_1) + \alpha_2 \bar{x} + (1 - \alpha_2) \underline{x} - p(s_2, 1 - r_2) (1 - \alpha_2) \underline{x}\}.$$

In the unique equilibrium, expert 1 exerts effort, expert 2 is passive, and the DM communicates with expert 1 about  $A_1$  and devotes some attention to acquiring information about  $A_2$ . The distinction between one-sided and two-sided communication arises endogenously.

**Crowding out** How well an expert fares in the deliberation stage depends not only on how attractive his *own* action seems to the DM, but also on the *other* expert's attractiveness.

**Proposition 1** (Crowding out). *Fix expert 2's attractiveness,  $\alpha_2$ . If expert 2 wants the DM's attention ( $\alpha_2 \leq \alpha^*$ ), his expected utility is a strictly decreasing function of the attention given to expert 1,  $r_1^*(\alpha_1)$ . If expert 2 does not want the DM's attention ( $\alpha_2 > \alpha^*$ ), his expected utility is a strictly increasing function of the attention given to expert 1,  $r_1^*(\alpha_1)$ .*

As  $r_2^*(\alpha_1) = 1 - r_1^*(\alpha_1)$ , a change in  $\alpha_1$  that causes the DM to pay more attention to expert 1 in equilibrium crowds out attention to expert 2. When expert 2 wants the DM's attention, this crowding out harms him; otherwise, it benefits him. Thus, the presence of expert 1 imposes a negative or positive externality on expert 2, and the size of this externality is captured by  $r_1^*(\alpha_1)$ .

**Corollary 1.** *Fix expert 2's attractiveness,  $\alpha_2$ . Expert 2's expected utility is non-monotonic in the attractiveness of expert 1,  $\alpha_1$ .*

This follows from the fact that the attention the DM devotes to expert 1 in equilibrium,  $r_1^*(\alpha_1)$ , is non-monotonic in  $\alpha_1$ :  $r_1^*(\alpha_1)$  increases for  $\alpha_1 \in (0, \alpha^*)$ , falls at  $\alpha^*$ , and decreases for  $\alpha_1 \in (\alpha^*, 1)$ . If expert 2 wants the DM's attention ( $\alpha_2 \leq \alpha^*$ ),  $\mathbf{EU}_{Exp_2}(\alpha_1)$  is negatively related to  $r_1^*(\alpha_j)$ ; otherwise, the reverse holds. This is illustrated in Figure 1.

## 2.2 The selection stage

We now add a pre-play stage by introducing *cue communication*, as in Dewatripont and Tirole (2005).

**Set-up** The DM can pay attention to two distinct actions in the deliberation stage.<sup>10</sup> When more than two experts seek the DM's attention, selection becomes

<sup>10</sup>The assumption that the DM can devote attention to only  $t$  topics can, in this context, be thought of as a lower bound  $\underline{r}$  on the amount of (nonzero) attention that the DM can devote to any

an important issue. To capture this, I add a pre-play stage in which the DM must select at most two experts for the deliberation stage. I refer to the pre-play stage as the “selection stage.”<sup>11</sup>  $N = N_{\bar{\alpha}} + N_{\underline{\alpha}}$  experts can enter the competition to be selected. Of these,  $N_{\bar{\alpha}}$  propose actions of high quality ( $\alpha = \bar{\alpha}$ ) and  $N_{\underline{\alpha}}$  of low quality ( $\alpha = \underline{\alpha} < \bar{\alpha}$ ).  $N_{\bar{\alpha}} > t$  is finite;  $N_{\underline{\alpha}}$  is infinite. I set parameters such that all experts want attention,  $\underline{\alpha} < \bar{\alpha} < \alpha^*$ .

As in Dewatripont and Tirole (2005), each expert’s quality is his private information. Hence, an expert may be willing to signal the quality of his action if it helps him get selected, and the DM may be willing to read such signals before choosing on which actions to deliberate, that is, with which experts to communicate. Specifically, at cost  $q_S > 0$ , an expert can send a cue (signal) that contains hard information about the quality of his action. Upon receiving a cue, the DM decides whether to process it, at cost  $q_R > 0$ , to learn the action’s quality. No expert can be selected without having sent a cue; hence,  $q_S$  can be thought of as an entry cost. As I explain below, this last assumption is not crucial.

I solve this game, deliberation stage plus selection stage, for perfect Bayesian equilibrium. Furthermore, I focus on the equilibria favored by the DM, that is, those with the maximum number of high-quality entrants.

**Information overload** As would be expected, when the cost of entry decreases, the supply of experts—and hence the number of actions from which the DM can choose—increases. However, as the DM’s choice set grows, her expected utility first increases, but then decreases. Proposition 2 states this key result:

---

one topic,  $r_i \in \{0\} \cup [\underline{r}, 1]$  for all  $i$ . This limits the number of topics on which she can deliberate  $t \equiv t(\underline{r}) \in \mathbb{N}$ . This assumption is appealing in the presence of a large number of topics; in practice, it is not possible to devote only a split second to each of (infinitely) many sources. The choice of  $t = 2$  is merely one of convenience; I show in the proof of Proposition 2 that all results go through for any finite  $t$ .

<sup>11</sup>The distinction between the selection stage and the deliberation stage is founded in cognitive science. As Cohen (2011, p.1) writes, the distinction between “attentive processing” (the deliberation stage) and “pre-attentive processing” (the selection stage) is logically inherent in the notion of *selective attention*: “A fundamental empirical phenomenon in human cognition is its limitation . . . One trademark of a limited system is its need for selection . . . Any type of selection presupposes the availability of some information in order to perform the very selection. Thus, some ‘pre-attentive’ processing must be performed prior to the operation of selective attention, and its output is used for the selection. The distinction between pre-attentive and attentive processing is essential in the study of selective attention.” In this model, the cues represent the information upon which pre-attentive processing is performed.

**Proposition 2** (Information overload). *As  $q_S \rightarrow 0$ , the DM receives more cues but eventually processes fewer. Her expected utility first increases and then decreases.*

It is instructive to describe how equilibrium behavior in the selection stage changes as the cost of sending cues,  $q_S$ , falls from prohibitively large to negligibly small. The impact on the DM’s expected utility is illustrated in Figure 2, where a decrease in  $q_S$  represents a movement from right to left on the  $x$ -axis.

- For high enough  $q_S$ , cues are so expensive that no expert enters.
- As  $q_S$  falls, it at some point becomes sufficiently attractive for some high-quality experts to enter. Here, the cues in themselves are a signal of high quality, so the DM need not process them but can select her communication partner(s) for the deliberation stage at random from the pool of entrants. (This relies on the assumption that the DM can observe that a cue was sent even if she does not assimilate it. If we relax this, the economic insights remain valid, as I explain below.) In this *signaling* outcome, the DM’s welfare increases as  $q_S$  falls so long as the number of entrants is smaller than two—or more generally, smaller than the number of experts with whom she can communicate in the deliberation stage; otherwise, the DM’s welfare remains constant. This is captured in Figure 2: As  $q_S$  decreases, the DM benefits from an expansion in information supply so long as  $q_S > \bar{q}_S$ . Then, as  $q_S$  falls further (but remains above  $\underline{q}_S$ ), the DM has access to (at least) two high-quality experts, but no low-quality experts, so her expected utility remains flat as  $q_S$  falls further.
- As  $q_S$  falls below  $\underline{q}_S$ , some low-quality experts find it attractive to enter as well. A signaling equilibrium, in which a random pick from among the entrants ensures a high-quality expert for the deliberation stage, no longer exists. The DM reacts in either of two ways: Either she continues to randomize and simply accepts the lower (average) expert quality or, if  $q_R$  is not too high, she reads cues with positive probability to screen out low-quality experts. So as not to make the selection stage trivial, I focus on  $q_R$  that is low enough for the DM to engage in active *screening*. Clearly, her welfare decreases as  $q_S$  falls, as it becomes harder to spot high quality. Already, the arrival of more cues—essentially, access to more information—makes the DM worse off. The next stage is merely the copestone.

- As  $q_S$  vanishes, the avalanche of low-quality cues reduces the average quality in the entrant pool so much that screening becomes futile—high quality becomes the proverbial needle in the haystack. As a result, there is neither signaling nor screening, just *pooling*: the DM gives up on active selection and accepts that she is all but bound to encounter low quality in the deliberation stage. Her expected utility thus approaches  $\underline{U}^*$ , her expected utility from communicating with two low-quality experts (on two low-quality actions) in the deliberation stage.

Thus, at a certain point, as the competition for the DM’s attention increases and she gets more information, she processes less of it—or tunes out—and fares worse. I refer to this phenomenon—the more cues the DM gets, the fewer she processes, and the worse she fares—as *information overload*.<sup>12</sup>

It is instructive to make precise how (the idea of) information overload is related to (the idea of) limited attention. To this end, consider this quote by Simon (1971):

What information consumes is rather obvious: It consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it. [p. 40-41]

Thus, attention, in limited supply, becomes a scarcer resource in relative terms when confronted with more information. But this does not imply information overload or that more information provided can *decrease* knowledge acquired. The idea of information overload is that a wealth of information not only “*creates... a need to allocate that attention*” (emphasis added) but actually *impairs the ability* to do so efficiently.

**Alternative assumptions** If I instead assume that the DM cannot observe that a cue was sent unless she incurs a cost to read it, the economic insights remain: She must open exactly two cues so long as only high-quality experts enter; then, she must either open exactly two cues but rely on information of lower quality, or open more than two cues on average to identify two high-quality experts. In either case, her expected utility remains constant when only high-quality types enter and decreases with the number of low-quality types.

---

<sup>12</sup>The signaling and screening outcomes arise independently of our assumption that there are infinitely many experts of low quality. The pooling outcome requires that  $N_{\underline{\alpha}}$ , the number of low-quality experts, is sufficiently large relative to  $N_{\bar{\alpha}}$ , the number of high-quality experts.

Further, the equilibria described above exist even if we relax the assumption that an expert must send a cue to enter. However, in that case, there is a further equilibrium for  $q_S \rightarrow 0$  in which the experts cease to send cues, aware that they are no longer processed, and the DM picks randomly from the entire pool of experts. Still, the DM favors the equilibrium in which she picks randomly from a subset of experts—which includes all high-quality experts—who send a cue, because it offers better odds of picking a high-quality expert.

Also, we need not *assume* differences in quality. Instead, suppose experts *invest* in quality. Specifically, suppose all  $N$  experts begin with low quality ( $\alpha = \underline{\alpha}$ ) but can invest in high quality ( $\alpha = \bar{\alpha}$ ) at some cost  $c > 0$  before entering. Proposition 2 implies that information overload frustrates investment in high quality. Intuitively, the value of quality is reflected in the expected utility difference between a high-quality and a low-quality type. For prohibitive  $q_S$ , both types expect to earn zero, so there is no incentive to invest in quality. As  $q_S$  falls, if  $c$  is not too high, some invest in quality and send cues. But as  $q_S \rightarrow 0$ , information overload erodes the premium on quality, so again no one invests. The supply of high quality collapses when it becomes too cheap to approach the DM. This is *not* because the DM ceases to value quality. On the contrary, she would like to treat high-quality experts preferentially; however, she is unable to do so when finding them amounts to looking for a needle in a haystack.

**Externalities as the driving force of information overload** The immediate cause of information overload is that the quality of proposed actions decreases with the quantity; actions worthy of deliberation become the proverbial needle in a haystack. The deeper cause, though, is negative externalities: Entry is individually rational for each expert, even as it complicates the selection problem for the DM and spoils overall communication. This is because each expert ignores how his own entry affects the communication environment as a whole. If the expert were identical for all actions, he would send cues only for two high-quality actions. In the decentralized setting, however, sending cues remains individually rational even as each cue sent aggravates the complexity of the DM’s selection problem up to a point where active selection breaks down. This, in turn, frustrates the incentives to *produce* quality. Intuitively, it is as if the low-quality experts, each seeking to be noticed, pollute the DM’s attention field. Indeed, information overload is similar to pollution or congestion and, like them, may be amenable to efficiency-improving intervention. The next

subsection addresses practical expressions of information overload, and discusses the fact that a DM with limited attention may want to limit access to her attention space, even if that reduces her choice set. In a nutshell, she faces a trade-off between comprehensiveness and comprehensibility.

### 2.3 Information overload in practice

With recent advances in information technology, individuals face massive data via more channels (phone, Internet, email, instant messages, etc.) and on more platforms (Facebook, Twitter, blogs, etc.). In the presence of information overload, such an abundance of information can be counterproductive. Indeed, a business research firm nominated information overload as the “problem of the year” in 2008, predicting that it would cost firms \$650 billion in lost productivity and innovation due to “unnecessary interruptions.” The main concern is that an escalating quantity of information comes with a decline in average quality and that this inverse relationship between amount and relevance makes it harder to find “good” information.<sup>13</sup> This makes selection, as in my model, a daunting issue, and suggests that consumers should value products that help reduce their choice sets and improve selection.

In this vein several technology firms, including Microsoft, Intel, Google, and IBM, recently formed a nonprofit organization, the Information Overload Research Group, to develop solutions to information overload.<sup>14</sup> Google’s success formula, its *ranking* algorithm, implements pre-selection. And its ubiquity on the Internet, as gateway and gatekeeper, betrays the import of information overload. The logic of pre-selection also underlies solutions such as email filters, ranking inbox messages by imputed importance, compiling communication histories for every sender, displaying email portions to allow for fast screening, and sophisticated filing and search functions. Such ranking of electronic messages minimizes information overload; that is, it reduces the “economic loss associated with the examination of a number of non- or less-relevant messages” and distinguishes “communications that are probably of interest from those that probably aren’t” (Losee, 1998).

Information overload is not just a matter of Internet and emails. In a seminal study, Jacoby et al. (1974) explore how the quality of consumption decisions depends on “information load,” measured as number of brands as well as amount of information per brand provided. Their experiment shows that the ability to pick the best

---

<sup>13</sup>The quotes in this paragraph are from Lohr (2007).

<sup>14</sup>The Information Overload Research Group’s web site is <http://iorgforum.org/>.



product dropped off at high levels of information load. In Jacoby et al. (1973), a companion paper, they further show that the subjects spent less time on processing information—or in their words, tuned out—once the information load exceeded a certain threshold.<sup>15</sup> Many other experiments in organization science, accounting, marketing, and information science corroborate the notion that more information can impair cognitive processes and decisions (Edmunds and Morris, 2000; Eppler and Mengis, 2004).

Because information overload is a driving force behind innovations in communication and information management, it is connected to recent research on choice architecture, that is, how the presentation of choices affects decisions (Thaler and Sunstein, 2008). Cronqvist and Thaler (2004), for example, study the introduction of a new retirement savings plan in Sweden in 1993. Eligible Swedes were encouraged to choose five out of 456 funds, to which their savings would be allocated. One third of all eligibles made no active choices; their savings were instead allocated to a default fund (essentially a pre-selection by the government). Information overload seems a likely reason that so many Swedes rely on the default choice: Comparing hundreds of funds is a Herculean task for ordinary households, and one might expect many of them to resort to the default or make superficial active decisions. Indeed, studying the same Swedish reform, Karlsson et al. (2006) show that funds that (for exogenous reasons) were better represented in the fund catalogue—that is, have better “menu exposure”—received more active contributions.<sup>16</sup>

All of the above examples suggest that decision-makers can benefit from receiving less information, despite the associated decrease in choice set. Indeed, in the presence of information overload, there is a trade-off between variety and simplicity, or between comprehensiveness and comprehensibility. This evidence contrasts with models of decision-making under unlimited attention, where a larger choice set cannot make an individual worse off.

---

<sup>15</sup>In the same vein, Iyengar (2011) provides direct empirical evidence that a reduction in choices can benefit decision-makers.

<sup>16</sup>In the same vein, studying a retirement savings plan in the United States, Beshears et al. (2013) show that making the decision problem less complex—by collapsing a multidimensional problem into a binary choice—increases enrollment in the plan. In a similar study, Choi et al. (2012) report that sending short email cues that draw attention to selective details of the savings program significantly affects participation. Further, in the context of Denmark, Chetty et al. (2014) show that wealth accumulation is highly responsive to automatic retirement contributions, which effectively eliminates the need to process information ahead of making retirement contributions.

### 3 A Single Expert and Strategic Attention Manipulation

In the previous section, we saw that crowding out and information overload result from the strategic interaction between multiple experts, each of whom try to persuade the DM to take one distinct action. By making minor modifications to the original setting, this section reinterprets the framework to study a DM who communicates with a *single* expert on one action that has several *aspects*. In the deliberation stage, the DM’s multi-tasking problem now stems from the fact that the DM must decide how to allocate her scarce attention between the various aspects of the action. In the selection stage, recall that in the “multiple experts”-setting analyzed in Section 2, the DM is unsure of each action’s ex ante appeal. The analogue in Section 3 is that the DM is unsure which aspects of the (single) action are financially relevant, and hence worth devoting attention to in the deliberation stage. Next, we introduce the set-up in detail.

#### 3.1 Modified set-up: Deliberation and selection stages

The DM communicates with a single expert about one action,  $A$ , which has many aspects. Specifically, suppose the DM’s payoff from  $A$  can be expressed as the sum of  $N_R$  components:  $\tilde{x} = \sum_1^{N_R} \tilde{x}_i$ . Each component takes the value  $\bar{x} > 0$  with probability  $\alpha_i$  and otherwise the value  $\underline{x} < 0$ . So the expected payoff from  $A$  is  $E(\tilde{x}) = N_R [\bar{\alpha}\bar{x} + (1 - \bar{\alpha})\underline{x}]$ , where  $\bar{\alpha} = \frac{1}{N_R} \sum_1^{N_R} \alpha_i$ . Similar to before, absent more information, the DM takes  $A$  in the deliberation stage if and only if  $\bar{\alpha} > \alpha^* \equiv \frac{-\underline{x}}{\bar{x} - \underline{x}}$ . Or put differently, if the DM can obtain more information, she uses an opt-in rule if  $\bar{\alpha} \leq \alpha^*$  and an opt-out rule if  $\bar{\alpha} > \alpha^*$ .

In addition to the  $N_R$  components of the action  $A$  that are relevant to the action’s payoff, there exist an additional  $N_\emptyset$  components of  $A$  that are irrelevant to the DM’s payoff. I assume that the DM does not know which components are relevant,<sup>17</sup> and is potentially unaware of components per se. Intuitively, this captures a plausible situation: Inclined towards a particular choice, the DM might yet discover (financially relevant) aspects that change her opinion. Thus, the DM is faced with two sets of questions: What components exist, and which ones are relevant (selection stage)? And how much attention should a given component receive (deliberation stage)?

---

<sup>17</sup>This is analogous to the assumption, in the setting with multiple actions presented in Section 2, that the DM does not know which actions are of a high quality and which actions are of a low quality.

As before, the expert’s communication incentives hinge on the DM’s decision rule:

If the DM follows an opt-in rule, the expert must persuade her to take the action. To maximize the chances that the DM revises her beliefs upwards, and so opts in, the expert seeks to draw her attention to those aspects that are most likely to yield favorable information. That is, he sends her cues about, and exerts persuasion effort on, the topics with the highest  $\alpha_i$ .

By contrast, if the DM follows an opt-out rule, her default is to take action  $A$ , but she may depart from this default—and opt out—if she learns about any relevant aspects of the action that are unfavorable. Then, the expert wants to withhold the relevant information. In the model, this means that, on his own accord, he would never send a cue about any relevant aspect in the selection stage, since this could only induce the DM to devote attention to it in the deliberation stage, and subsequently to change her mind and abandon the action after all. In this situation, a mandatory disclosure law would make a difference to the expert’s communication strategy: By mandating that the expert (firm) reveals any relevant financial aspects of the product, it would be illegal to withhold this information.

To formalize what happens when the expert is subjected to such a mandate, assume for simplicity that there is one relevant topic,  $N_R = 1$ , that the number of irrelevant topics  $N_\emptyset$  is infinite, and that the DM can at most deliberate on two topics (as in Section 2). These assumptions are merely simplifying; indeed, the “strategic information overload”-result reported below holds as long as the DM can devote attention to only a limited number of topics in the deliberation stage and  $N_\emptyset$  is sufficiently large relative to  $N_R$ . Finally, as before, it is costly for the DM to process cues:  $q_R > 0$ . For convenience, I refer to cues about relevant (irrelevant) topics as relevant (irrelevant) cues.

### 3.2 Strategic information overload and the limitations of mandatory disclosure laws

When  $\bar{\alpha} > \alpha^*$ , the DM follows an opt-out rule. Thus, if the DM lacks access to—or is unaware of—the relevant topic, the expert has no reason to bring it to her attention by sending a relevant cue. In fact, the expert is best off sending no cues at all. Now suppose that, by a disclosure mandate, the relevant cue must be sent.

**Proposition 3** (Strategic information overload). *Let  $\bar{\alpha} > \alpha^*$ . Suppose that disclo-*

sure laws mandate that the relevant cue is sent. As  $q_S \rightarrow 0$ , the expert sends an increasing swarm of irrelevant cues, and the DM's expected utility decreases.

Proposition 3 is closely related to the information overload result in Section 2; yet, the underlying economic mechanism is different. In Section 2, information overload resulted from various experts' competition for attention. Proposition 3 shows that a single expert, who faces no competition for the DM's attention but who can communicate with the DM about multiple aspects of the action for which he advocates, *strategically induces* information overload in the DM when he is subjected to a disclosure provision that mandates drawing the DM's attention to the financially relevant aspect of the action (the product he tries to persuade the DM to buy).

Intuitively, the expert is afraid that the DM, by paying attention to the relevant topic, might discover unfavorable information about the action and opt out. To reduce the odds that the DM identifies — that is, selects — the relevant topic, the disclosure mandate thus gives the expert an incentive to supply irrelevant aspects in the selection stage, by sending out irrelevant cues, even though this is costly for him. A swarm of mostly irrelevant cues, in turn, thwarts the DM's chance, and hence her incentives, to pinpoint the relevant topic. In other words, the expert intentionally induces information overload, which effectively conceals the inconvenient aspect that the firm was mandated to disclose in the first place. In a nutshell, the disclosure mandate, which is intended to raise the DM's awareness of the financially relevant aspect of the product that she considers, generates a response on the part of the expert that in practice renders the DM financially illiterate, or *obfuscates* her.

As a result, the disclosure mandate has no impact on the DM's final decision; she buys the product even if an understanding of the financially relevant aspect would have pushed her to opt out, since strategic information overload makes her as uninformed in the presence of a mandate as she is in its absence. This sharply illustrates that, when consumers' attention is limited, simple disclosure rules can be completely impotent and have no impact on social welfare.

**Strategic information overload in practice** *USA Today* recently ran an internal study on the costs of maintaining a basic checking account at the ten largest US banks and credit unions. While the most basic fees were found to be disclosed on the institutions' websites, many others were listed only in the "Schedule of Fees and Charges." That, however, turned out to be difficult to find.

But even the world’s largest search engine couldn’t unearth a fee schedule for HSBC, TD Bank, Citibank and Capital One. To get their fee information, we had to e-mail or call the banks.

Determined customers can search for information about fees in banks’ official disclosure documents, but they’ll need a lot of time and a couple of cups of coffee, too. An analysis of checking accounts for the 10 largest banks by the Pew Health Group found that the median length of their disclosure statements was 111 pages. None of the banks provided key information about fees on a single page...<sup>18</sup>

Note that the issue was not only that the “inconvenient” information was at times unavailable. It was also that, even when provided, it was made available in a way that made it costly to locate the relevant information; that is, in a way that induced strategic information overload in the customer. Ordinary customers would be hard-pressed to know not only *where* to look for relevant items but also *what* items to look for. Similar conditions prevail in other countries. In 2008, the website *This Is Money* cited a warning by the British consumer and competition authority, the Office of Fair Trading, that

[credit card] providers can add to the problem knowing that consumers cannot process complex information . . . They can create “noise” by increasing the quantity and complexity of information, which makes it difficult for consumers to see the real price.<sup>19</sup>

Note that the concern here is the complexity *of information*; even if a product itself is simple, information overload effectively presents it in an overly complex fashion, by hiding it in very long disclosure statements. The financial products market seems rife with such practices.<sup>20</sup> Credit cards are perhaps the most widely debated example.

---

<sup>18</sup>Tilghman, Molly, and Sandra Block. 2011. “Finding Info on Bank Fees May Take Digging.” USA Today, October 21, <http://www.usatoday.com/money/perfi/credit/story/2011-10-20/comparing-bank-fees/50845842/1>

<sup>19</sup>Daily Mail Reporter and Sean Poulter. 2008. “Credit Card £400m Small Print Rip-Off,” *This Is Money*, October 29, <http://www.thisismoney.co.uk/money/cardsloans/article-1619869/Credit-card-400m-small-print-rip-off.html>

<sup>20</sup>Richards et al. (2016) provide an example of strategic information overload, or obfuscation, in another market: consumer retailing. They note that consumer product manufacturers tend to offer retailer-specific variants of common brands, and argue that this multiplicity of near-identical products is intended to prevent direct price comparison, and to thereby raise consumer search costs. They present empirical evidence consistent with such obfuscation using German and French retail scanner data.

As quoted in a 2009 Reuters article, President Obama said “No more fine print, no more confusing terms and conditions,” in a meeting with US credit card company executives on consumer protection regulation.<sup>21</sup>

These examples underscore an interesting aspect that eludes many communication models: mandatory disclosure is not a panacea. In the above examples, the communication problem is neither a willful misrepresentation (“cheap talk”) nor the withholding of facts (“strategic non-disclosure”). Here, the banks are mandated to provide fee information; cheap talk or non-disclosure are illegal and would have serious consequences *ex post*. Still, this does not mean that consumers become well-informed. Even when information is hard and disclosed, senders can still—through strategic attention manipulation—conceal what is relevant by manipulating *the sheer amount* of information. Thus, simple disclosure rules like the Truth in Lending Act may prove completely ineffective when consumers’ attention is limited.<sup>22</sup>

### 3.3 Complexification and the limitations of “simple labels” for consumer protection

The previous subsection illustrates that too much disclosure can be as concerning as too little; a surfeit of details can prove as uninformative as a dearth thereof. But if disclosure mandates can backfire because firms can induce information overload, intuition suggests an immediate solution: to mandate not only what firms disclose, but also that it is disclosed in an easy-to-understand way. Indeed, more recent disclosure rules often are of this flavor: Credit card companies must now disclose key details of the products they offer in a salient fashion, health insurance companies must provide a Summary of Benefits and Coverage (SBC) of each plan offered, and food producers must adhere to standardized labeling regulations to declare all ingredients and the caloric content of each serving, for example.

In this subsection, I use the theoretical framework to analyze the strategic response on the part of an expert (firm) when faced with a mandate to provide *easy-to-understand* information about the product it sells. As we will see, even detailed mandates to disclose easy-to-understand information can backfire, by inducing a

---

<sup>21</sup>Alexander, David, and John Poirier. 2009. “Obama Calls for Credit Card Reforms,” Reuters, April 23, <http://www.reuters.com/article/2009/04/23/us-obama-creditcards-idUSTRE53M10720090423>

<sup>22</sup>The Truth in Lending Act, of 1968, is a federal law intended to “safeguard the consumer in connection with the utilization of credit by requiring full disclosure of the terms and conditions of finance charges in credit transactions or in offers to extend credit.”

“complexification” of the underlying product itself.

I show this by extending the model with a single expert to allow the DM’s payoff from the (single) action to be comprised of many components, and to allow the expert to manipulate that composition so long as the total payoff stays constant. Intuitively, such payoff-equivalent variations amount to changing the number of financially relevant aspects of the product. Below, I formalize this:

To allow for complexification, I introduce an option enabling the expert to design  $A$  in a way that makes its payoff less transparent. As before, the DM’s payoff is—at least initially—the sum of  $N_R$  components,  $\tilde{x} = \sum_1^{N_R} \tilde{x}_i$ , and each component takes the value  $\bar{x} > 0$  with probability  $\alpha_i$  and otherwise the value  $\underline{x} < 0$ . However, the expert can now recompose the payoff structure into any form  $\tilde{y} = \sum_1^N \tilde{y}_j$  so long as the total (realized) payoff is invariant:  $\tilde{y} = \tilde{x}$ . I call  $\tilde{y}$  a *payoff-equivalent variation* (of  $\tilde{x}$ ). Crucially, I assume that the DM does not know the “original” composition,  $\{\tilde{x}_i\}$ .<sup>23</sup> In any case, absent communication, such variation does not matter; it affects neither the DM’s decision nor her welfare.

As an illustration, consider two simple mathematical operations the expert can use to create payoff-equivalent variations. One is to divide each  $\tilde{x}_i$  into  $m$  parts,  $\sum_{h=1}^m \tilde{x}_{ih} = \tilde{x}_i$ , so that the new payoff structure has  $N = mN_R$  components:  $\tilde{y} = \sum_{i=1}^{N_R} \sum_{h=1}^m \tilde{x}_{ih} = \sum_{j=1}^N \tilde{y}_j$ . For example, instead of incorporating total expenses into one salient item, such as the monthly rent, a landlord might disaggregate them into various fees, such as for maintenance, utilities, move-in or move-out, parking, laundry room, or other administrative services. The other operation is to add components that neutralize each other, as in  $\tilde{y} = \sum_{i=1}^{N_R} \tilde{x}_i + \sum_{i=1}^{N_R} \tilde{x}_i - \sum_{i=1}^{N_R} \tilde{x}_i = \sum_{j=1}^N \tilde{y}_j$  where  $N = 3N_R$ . A real-world example is a purchase involving a nominal price, fees, taxes, discounts, bonuses, rewards, and so forth, which partly offset each other. In both examples, seeing one component is not informative about the others.

Even if the total expected payoff remains unchanged, the payoff composition matters for communication. As before, suppose the DM cannot deliberate on more than a certain number of topics (in keeping with the previous framework, say, two). A proliferation of components then causes more relevant information to slip her attention. Furthermore, disaggregating the payoff can make each component, in and

---

<sup>23</sup>Alternatively, I could assume that she does not know how a given payoff-relevant variation is related to the original composition. Either assumption captures situations of the following kind: A consumer is inclined to buy a good based on superficial information. That said, she is not aware of all aspects, such as hidden costs, that could influence her decision. Further, discovering one aspect is not necessarily informative about undiscovered aspects.

of itself, less important. To see this point, let  $\tilde{x}_{ih} = \frac{\tilde{x}_i}{m}$  for all  $i$  in the first (landlord) example above. Increasing  $m$  leaves the total payoff,  $\tilde{y}$ , unchanged but shrinks every component,  $\frac{\tilde{x}_i}{m}$ . Crucially, this reduces how much the DM can learn from a given number of components.

Payoff-equivalent variation is thus a means of manipulating the DM’s learning process. How the expert uses such means hinges, as before, on the DM’s decision rule. If she uses an opt-in rule, the expert wants to help her learn more about  $A$ . He would set  $N \leq 2$ —such that no relevant aspect escapes deliberation—and exert communication effort on all the components. That is, he would simplify the payoff structure and strive to explain.

By contrast, if the DM follows an opt-out rule, the expert wants to do the exact opposite. He would increase the number of components, even if it were costly to do so, only to thwart learning. Suppose he must pay  $v > 0$  to raise  $N$  by one.

**Proposition 4** (Complexification). *Let  $\bar{\alpha} > \alpha^*$ . Suppose all relevant cues are sent out. As  $v \rightarrow 0$ , the expert sets  $N \rightarrow +\infty$ , and the DM’s expected utility falls to  $E(\tilde{x})$ .*

As  $v \rightarrow 0$ , the expert increases  $N$  such that more relevant information escapes the DM’s attention, given that she can deliberate on only a limited number of components. At the same time, he disaggregates the payoff to reduce the amount of information she can possibly wrest from any given component. In the limit, even what she can learn from the components she is capable of studying becomes so trivial that it no longer affects her decision: She chooses what she would have chosen without the information. Intuitively, the expert makes the action unnecessarily complex, and thereby successfully prevents the DM from getting the full picture.

Complexification thus has the same welfare consequences as inducing information overload — it can eradicate all intended welfare gains from mandated information provision, even if the mandate prohibits unclear communication.

**Complexification in practice** According to Edward L. Yingling, president and chief executive of the American Bankers Association (ABA), during his first term Obama urged credit card companies “to issue a *simple* credit card product” (emphasis added). A year earlier, after similar comments by Federal Reserve Chairman Ben S. Bernanke “that improved disclosures alone cannot solve all of the problems consumers face in trying to manage their credit card accounts,” the ABA and other industry



representatives had signaled strong opposition to such interventions.<sup>24</sup>

The UK financial regulator, the Financial Services Authority (FSA), is also quoted as saying:

[P]roviders of financial products may gain from the lack of price transparency about their products. [...] It may be in the provider's interest to increase the complexity of the product charges.

Other examples of complexification come from the food industry, where legislation mandates that all ingredients be listed in a food declaration; however, “incidental additives” need not be listed. This has provided food producers with incentives for complexification, most prominently by removing an ingredient that consumers may want to avoid, and replacing it with one or several incidental additives that accomplish the same effect in food but that remain invisible (provided each of them is included in small enough a quantity). A recent example is Starbucks’s use of a color additive derived from the animal world classified as an incidental additive instead of an ingredient, which effectively caused a non-vegan product to be labeled vegan and caused a stir among vegan consumers.<sup>25</sup> Similarly, widespread consumer awareness of the dangers of BPA has spurred development of a new range of plastic products labelled “BPA Free”; however, these products contain BPA replacement substances that have produced health problems similar to those associated with BPA itself.<sup>26</sup> Clearly, in the absence of mandated disclosure of the ingredients in food, none of these complexifications was necessary; in the presence of this disclosure rule, however, they represented strategic responses on the part of food producers.<sup>27</sup>

As opposed to the examples presented in the previous subsection, which showed manipulation of the complexity of *information*, the examples provided here refer to manipulation of the complexity of the *product* itself. Whether the complexity of information or the complexity of the product itself is involved, the danger is that

---

<sup>24</sup>Labaton, Stephen. 2008. “U.S. Seeks New Curbs on Credit-Card Practices,” *New York Times*, May 3, <http://www.nytimes.com/2008/05/03/business/03credit.html>

<sup>25</sup>Source: Fortney D. Beware: Starbucks’ Soy Strawberries & Creme Frappuccino Is NOT vegan [weblog entry]. This Dish Is Veg (14 Mar 2012). Available: <http://goo.gl/kPpj5> [accessed 11 Mar 2016].

<sup>26</sup>Source: LaMotte S. BPA-free plastic alternatives may not be safe as you think (1 Feb 2016). Available: <http://www.cnn.com/2016/02/01/health/bpa-free-alternatives-may-not-be-safe/> [accessed 11 Mar 2016].

<sup>27</sup>Bjorkegren (2016) analyzes data from the FDA Food Labeling and Packaging Survey from 1976 through 2006. He documents rising product complexity in food over this time period, and explains how food product design reflects the introduction of various types of labeling regulations with great precision.

misguided decisions affect consumers' risk of getting into debt or of buying food products that they wish to avoid.

However, while complexity of information has a simple legislative recipe mandating that firms provide easy-to-understand information, complexification is much harder to address from a legislative perspective. In fact, Proposition 4, along with the real-world examples of complexification presented in this subsection, underscore that even the most elaborate disclosure rule – which specifies exactly how easy-to-understand-information should be provided – may be completely ineffective for the consumer. This is because *any* regulation at the communication level proves futile if the seller can modify the *object* of communication in a way that makes it intellectually challenging to grasp, even with all details correctly disclosed. While the communication is correct, the matter to be decided becomes too complicated. In such a case, to be effective, regulation may have to target the object of communication per se, that is, product design. This, as the examples illustrate, is a much thornier issue.

## 4 Conclusion

Limits on consumer attention limitations give firms incentives to manipulate prospective buyers' allocation of their attention. To the best of my knowledge, this paper is the first to model such attention manipulation. In its presence, competitive information supply can reduce consumer knowledge by causing information overload. Moreover, a single firm subjected to a disclosure mandate may deliberately induce information overload to obfuscate financially relevant information, or engage in product complexification to bound consumer financial literacy.

These findings demonstrate that attention limitations matter crucially for whether disclosure regulation improves consumer welfare: Disclosure rules that would improve welfare for agents without attention limitations can prove ineffective for consumers with limited attention. Obfuscation suggests a role for rules that mandate not only the content but also the format of disclosure; however, even rules that mandate disclosure of “easy-to-understand” information are ineffective against complexification, which may call for regulation of product design.

An interesting avenue, not pursued in the present paper, is that heterogeneity in attention constraints may provoke multiple forms or degrees of attention manipulation. Banerjee and Mullainathan (2008) posit that the poor are subject to tighter

attention constraints than the rich, who can afford better technologies to free up attention. They then show that this induces differences in productivity that amplify the differences in initial endowment; inequality breeds more inequality. Allcott and Taubinsky (2015) and Taubinsky and Rees-Jones (2015) similarly document considerable heterogeneity in attention limitations, and Taubinsky and Rees-Jones (2015) show that lower-income individuals have more severe attention limitations. The current paper's findings suggest that the problem may be even worse: The poor may not only start out with tighter attention constraints, but may also find their limited attention exploited more than the rich. In short, the tighter constraints may make them less productive *and more manipulable*. Manipulation is perhaps the more worrisome problem in that it is, as shown in this paper, prone to create externalities, and thus constrained inefficient outcomes. But such questions are left for future research.

The poor may not only start out with tighter attention constraints, but may also find their limited attention exploited more than the rich. In short, the tighter constraints may make them less productive *and more manipulable*. Manipulation is perhaps the more worrisome problem in that it is, as shown in this paper, prone to create externalities, and thus constrained inefficient outcomes. But such questions are left for future research.

## References

- Abaluck, Jason and Jonathan Gruber**, “Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program,” *American Economic Review*, 2011, *101* (4), 1180–1210.
- Allcott, Hunt and Dmitry Taubinsky**, “Evaluating Behaviorally Motivated Policy: Experimental Evidence from the Lightbulb Market,” *American Economic Review*, 2015, *105* (8), 2501–38.
- Alonso, Ricardo, Wouter Dessein, and Nico Matouschek**, “When does coordination require centralization?,” *The American Economic Review*, 2008, *98* (1), 145–179.
- Ambrus, Attila and Satoru Takahashi**, “Multi-sender cheap talk with restricted state spaces,” *Theoretical Economics*, 2008, *3* (1), 1–27.
- Banerjee, Abhijit V. and Sendhil Mullainathan**, “Limited Attention and Income Distribution,” *American Economic Review Papers and Proceedings*, 2008, *98* (2), 489–93.
- Battaglini, Marco**, “Multiple Referrals and Multidimensional Cheap Talk,” *Econometrica*, 2002, *70* (4), 1379–1401.
- Beshears, John, James Choi, David Laibson, and Brigitte C. Madrian**, “Simplification and Saving,” *Journal of Economic Behavior and Organization*, 2013, *95*, 130–145.
- Bjorkegren, Dan**, “Hidden Quality,” *Mimeo*, 2016.
- Blackwell, David**, “Comparison of Experiments,” in “Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability” University of California Press Berkeley, Calif. 1951, pp. 93–102.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer**, “Salience Theory of Choice Under Risk,” *The Quarterly Journal of Economics*, 2012, *127* (3), 1243–1285.
- , – , and – , “Salience and Consumer Choice,” *Journal of Political Economy*, 2013, *121* (5), 803–43.

- , – , and – , “Competition for Attention,” *The Review of Economic Studies*, 2016, 83 (2), 481–513.
- Caillaud, Bernard and Jean Tirole**, “Consensus Building: How to Persuade a Group,” *American Economic Review*, 2007, 97 (5), 1877–1900.
- Carlin, Bruce I.**, “Strategic price complexity in retail financial markets,” *Journal of Financial Economics*, 2009, 91 (3), 278–87.
- Chakraborty, Archishman and Rick Harbaugh**, “Comparative cheap talk,” *Journal of Economic Theory*, 2007, 132, 70–94.
- and – , “Persuasion by Cheap Talk,” *American Economic Review*, 2010, 100 (5), 2361–2382.
- Chetty, Raj, Adam Looney, and Kory Kroft**, “Salience and Taxation: Theory and Evidence,” *American Economic Review*, 2009, 99 (4), 1145–77.
- , **John N. Friedman, Sren Leth-Petersen, Torben Heien Nielsen, and Tore Olsen**, “Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark,” *The Quarterly Journal of Economics*, 2014, 129 (3), 1141–1219.
- Choi, James J., Emily Haisley, Jennifer Kurkoski, and Cade Massey**, “Small Cues Change Savings Choices,” *NBER Working Paper No. 17843*, 2012.
- Cohen, Asher**, “Selective Attention. Encyclopedia of Cognitive Science,” *Nature Publishing Group, Macmillan.*, 2011.
- Crawford, Vincent P. and Joel Sobel**, “Strategic information transmission,” *Econometrica*, 1982, 50 (6), 1431–1451.
- Cronqvist, Henrik and Richard Thaler**, “Design choices in privatized social-security systems: Learning from the Swedish experience,” *The American Economic Review*, 2004, 94 (2), 424–28.
- de Clippel, Geoffroy, Kfir Eliaz, and Kareen Rozen**, “Competing for Consumer Inattention,” *Journal of Political Economy*, 2014, 122 (6), 1203–1234.
- Dellavigna, Stefano and Joshua M. Pollet**, “Investor Inattention and Friday Earnings Announcements,” *The Journal of Finance*, 2009, 64 (2), 709–49.

- Dessein, Wouter and Tano Santos**, “Adaptive organizations,” *Journal of Political Economy*, 2006, 114 (5), 956–95.
- Dewatripont, Mathias and Jean Tirole**, “Modes of Communication,” *Journal of Political Economy*, 2005, 113 (6), 1217–1238.
- Edmunds, Angela and Anne Morris**, “The Problem of Information Overload in Business Organizations: A Review of the Literature,” *International Journal of Information Management*, 2000, 20 (1), 17–18.
- Ellison, Glenn**, “A Model of Add-On Pricing,” *The Quarterly Journal of Economics*, 2005, 120 (2), 585–637.
- **and Alexander Wolitzky**, “A search cost model of obfuscation,” *The RAND Journal of Economics*, 2012, 43 (3), 417–441.
- Eppler, Martin J. and Jeanne Mengis**, “The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines,” Technical Report 2004.
- Gabaix, Xavier and David Laibson**, “Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets,” *The Quarterly Journal of Economics*, 2006, 121 (2), 505–540.
- Gentzkow, Matthew and Emir Kamenica**, “Competition in Persuasion,” *The Review of Economic Studies*, 2017, 84 (1), 300.
- Grossman, Stanford J.**, “Informational Role of Warranties and Private Disclosure about Product Quality,” *Journal of Law & Economics*, 1981, 24 (3), 461–483.
- Hefti, Andreas**, “Limited Attention, Competition and Welfare,” *Mimeo*, 2017.
- Inderst, Roman and Martin Obradovits**, “Too Much Attention on Low Prices? Loss Leading in a Model of Sales with Salient Thinkers,” *CEPR Discussion Paper No. DP10813*, 2015.
- Iyengar, Sheena**, “The Art of Choosing,” *Grand Central Publishing*, 2011.
- Jacoby, Jacob, Carol A. Kohn Berning, and Donald D. Speller**, “Time Spent Acquiring Product Information as a Function of Information Load and Organization,” *Proceedings, 81st Annual Convention, American Psychological Association*, 1973, 8 (2), 1–1.

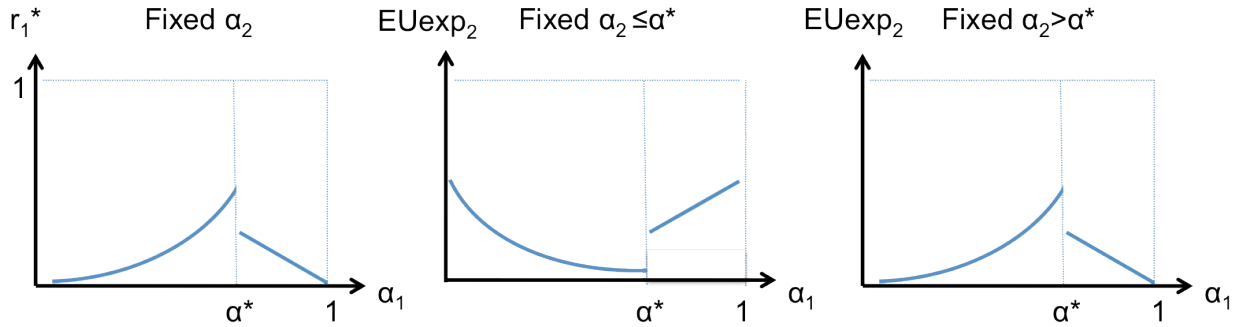
- , **Donald E Speller**, and **Carol A. Kohn Berning**, “Brand Choice Behavior as a Function of Information Load: Replication and Extension,” *Journal of Consumer Research*, 1974, 1 (1), 33–42.
- Karlsson, Anders**, **Massimo Massa**, and **Andrei Simonov**, “Portfolio Choice and Menu Exposure,” *Mimeo*, 2006.
- Kartik, Navin**, “Strategic communication with lying costs,” *Review of Economic Studies*, 2009, 76 (4), 1359–1395.
- , **Frances Xu Lee**, and **Wing Suen**, “Investment in concealable information by biased experts,” *The RAND Journal of Economics*, 2017, 48 (1), 24–43.
- , **Marco Ottaviani**, and **Francesco Squintani**, “Credulity, lies, and costly talk,” *Journal of Economic Theory*, 2007, 134 (1), 93–116.
- Krishna, Vijay** and **John Morgan**, “A model of expertise,” *The Quarterly Journal of Economics*, 2001, 116 (2), 747–775.
- Lohr, Steve**, “Is Information Overload a \$650 Billion Drag on the Economy?,” *The New York Times Bits*, December 2007.
- Losee, Robert M. Jr**, “Minimizing Information Overload: The Ranking of Electronic Messages,” *Journal of Information Science*, 1998, 15 (3), 179–189.
- Manzini, Paola** and **Marco Mariotti**, “Competing for Attention: Is the Showiest Also the Best?,” *The Economic Journal*, pp. n/a–n/a.
- Milgrom, Paul**, “Good news and bad news: Representation theorems and applications,” *The Bell Journal of Economics*, 1981, 12 (2), 380–391.
- and **John Roberts**, “Relying on the information of interested parties,” *The RAND Journal of Economics*, 1986, 17 (1), 18–32.
- Persson, Petra**, “Attention Manipulation and Information Overload,” *SSRN Working Paper no. 2102155*, 2012.
- Simon, Herbert A.**, “Designing organizations for an information rich world,” in Martin Greenberger, ed., *Computers, communications, and the public interest*, 1971, pp. 37–72.

- Sims, Christopher A.**, “Implications of rational inattention,” *Journal of Monetary Economics*, 2003, 50 (3), 665–690.
- Spiegler, R and K. Eliaz**, “Consideration Sets and Competitive Marketing,” *Review of Economic Studies*, 2011, 78, 235–262.
- and –, “On the Strategic Use of Attention Grabbers,” *Theoretical Economics*, 2011, 6, 127–155.
- Spiegler, Ran**, “Competition over agents with boundedly rational expectations,” *Theoretical Economics*, 2006, 1 (2), 207–231.
- Taubinsky, Dimitry and Alex Rees-Jones**, “Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment,” *mimeo*, 2015.
- Thaler, Richard and Cass R. Sunstein**, “Nudge: Improving Decisions about Health, Wealth, and Happiness,” *Yale University Press*, 2008.
- Wiederholt, Mirko**, “Rational Inattention,” *The New Palgrave Dictionary of Economics*, 2010.
- Wilson, Chris M.**, “Ordered search and equilibrium obfuscation,” *International Journal of Industrial Organization*, 2010, 28 (5), 496 – 506.



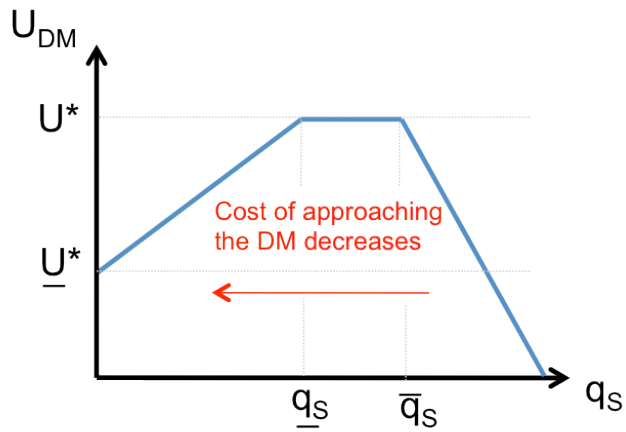
# A Figures

Figure 1: Crowding Out



*Note:* The DM's attention devoted to Expert 1 in equilibrium,  $r_1^*(\alpha_1)$ , for a given  $\alpha_2$ , is non-monotonic in Expert 1's attractiveness (left panel). This makes Expert 2's utility nonmonotonic in  $\alpha_1$ : When Expert 2 desires attention ( $\alpha_2 \leq \alpha^*$ ),  $r_1^*(\alpha_1)$  represents a negative externality on Expert 2, so  $\mathbf{EU}_{Exp_2}(\alpha_1)$  is negatively related to  $r_1^*(\alpha_1)$  (middle panel). When Expert 2 does not desire attention ( $\alpha_2 > \alpha^*$ ),  $r_1^*(\alpha_1)$ , represents a positive externality on Expert 2, so  $\mathbf{EU}_{Exp_2}(\alpha_1)$  is positively related to  $r_1^*(\alpha_1)$  (right panel).

Figure 2: **Information overload**



A movement from right to left on the x-axis represents a decrease in the cost of entry,  $q_s$ . As  $q_s$  decreases, the DM first benefits from an expansion in information supply (so long as  $q_s > \bar{q}_s$ ). Then, the DM has access to (at least) two high-quality experts, but no low-quality experts, so she obtains her maximum possible (decision) payoff,  $U^*$ . As the cost of entry falls below  $\underline{q}_s$ , low-quality experts join the battle for access to the DM's attention, and her expected utility falls below  $U^*$ . When  $q_s \rightarrow 0$ , the number of low-quality experts who enter tends to infinity, so the DM ceases to screen experts. Her decision payoff falls, as she relies on information of lower quality on average. Thus, when information becomes cheap enough, the more information she gets, the less information she processes, and the worse she fares.

## B Proofs

### B.1 Proof of Lemma 1

**Claim A1.1** When  $\alpha_1, \alpha_2 \leq \alpha^*$ , there exists a unique, interior equilibrium  $(r_1^*, s_1^*, s_2^*) \in (0, 1)^3$  of this game.

**Proof of Claim A1.1** When  $\alpha_1, \alpha_2 \leq \alpha^*$ , the problem of  $S_i$ ,  $i \in \{1, 2\}$  is given by

$$\max_{s_i \geq 0} \{d\alpha_i p(s_i, r_i) - c(s_i)\}.$$

For a given (conjectured)  $\hat{r}_1$ ,  $S_i$ 's first-order condition is given by

$$d\alpha_i p_1(s_i, \hat{r}_i) = c'(s_i). \quad (2)$$

Because  $c'(s_i) \rightarrow 0$  as  $s_i \rightarrow 0$ , and because  $c'(s_i) \rightarrow \infty$  as  $s_i \rightarrow 1$ , the range of the right-hand side (RHS) is  $(0, \infty)$ . As  $d > 0$  and  $p_1(s_i, \hat{r}_i) > 0$  for all  $\hat{r}_i \in [0, 1]$ , the left-hand side (LHS) is strictly positive for  $\alpha_i \in (0, 1]$ . Since  $p(s_i, \hat{r}_i)$  is concave in  $s_i$  and  $c(s_i)$  is convex, LHS is decreasing in  $s_i$  and RHS is increasing in  $s_i$ . These observations imply that there exists a unique solution  $s_i^*(\hat{r}_i) \in (0, 1)$  to this equation. Clearly, the best reply function  $s_i^*(\hat{r}_i)$  is monotonically increasing if  $p_{12} > 0$ , and monotonically decreasing if  $p_{12} < 0$ . Because the function  $p(\cdot)$  is concave and  $c(\cdot)$  is convex, this solution is the solution to the maximization problem (the second-order condition holds).

The problem of the *DM* is given by

$$\begin{aligned} \max_{\{r_1, r_2\} \in [0, 1]^2} \bar{x} (\alpha_1 p(s_1, r_1) + \alpha_2 p(s_2, r_2)) \quad \text{s.t.} \quad r_1 + r_2 &= 1 \\ \Leftrightarrow \max_{r_1 \in [0, 1]} \bar{x} (\alpha_1 p(s_1, r_1) + \alpha_2 p(s_2, 1 - r_1)). \end{aligned}$$

For given (conjectured)  $\hat{s}_1$  and  $\hat{s}_2$ , her first-order conditions (FOCs) are given by

$$\bar{x} (\alpha_1 p_2(\hat{s}_1, r_1) - \alpha_2 p_2(\hat{s}_2, 1 - r_1)) = 0, \quad (3)$$

where  $\bar{x} > 0$ .

We substitute the experts' best reply functions from (2) into (3) and obtain

$$\alpha_1 p_2(s_1^*(r_1), r_1) = \alpha_2 p_2(s_2^*(1 - r_1), 1 - r_1). \quad (4)$$

An equilibrium which is interior must satisfy (4). We will now discuss the existence and uniqueness of equilibria in this communication game.

First, we show that  $p_2(s_1^*(r_1), r_1)$  is monotonically decreasing in  $r_1$ , i.e., that its derivative is negative. Differentiating  $p_2(s_1^*(r_1), r_1)$  with respect to  $r_1$  yields

$$p_{21}(s_1^*(r_1), r_1) s_1^{*'}(r_1) + p_{22}(s_1^*(r_1), r_1). \quad (5)$$

Differentiating sender 1's equilibrium condition, (2), with respect to  $r_1$  yields

$$\begin{aligned} d\alpha_i(p_{11}(s_1^*(r_1)) s_1^{*'}(r_1) + p_{12}(s_1^*(r_1))) &= c''(s_1^*(r_1)) s_1^{*'}(r_1), \\ s_1^{*'}(r_1) &= \frac{d\alpha_i p_{12}(s_1^*(r_1))}{c''(s_1^*(r_1)) - d\alpha_i p_{11}(s_1^*(r_1))} \end{aligned} \quad (6)$$

Inserting (6) into (5) yields that the derivative of  $p_2(s_1^*(r_1), r_1)$  is negative if and only if

$$\frac{d\alpha_i p_{12}(s_1^*(r_1)) p_{12}(s_1^*(r_1))}{c''(s_1^*(r_1)) - d\alpha_i p_{11}(s_1^*(r_1))} + p_{22}(s_1^*(r_1), r_1) < 0.$$

Using the fact that  $(c''(s_1^*(r_1)) - d\alpha_i p_{11}(s_1^*(r_1)))$  is strictly positive, we rearrange the formula to obtain

$$d\alpha_i p_{12}(s_1^*(r_1)) p_{12}(s_1^*(r_1)) < -p_{22}(s_1^*(r_1), r_1) (c''(s_1^*(r_1)) - d\alpha_i p_{11}(s_1^*(r_1)))$$

$$p_{12}(s_1^*(r_1)) p_{12}(s_1^*(r_1)) < p_{11}(s_1^*(r_1), r_1) p_{22}(s_1^*(r_1), r_1) - \frac{1}{d\alpha_i} p_{22}(s_1^*(r_1), r_1) c''(s_1^*(r_1)).$$

Because  $-\frac{1}{d\alpha_i} p_{22}(s_1^*(r_1), r_1) c''(s_1^*(r_1)) > 0$ , this condition is implied by global concavity. This establishes that  $p_2(s_1^*(r_1), r_1)$  is monotonically decreasing in  $r_1$ .

Second, we show that there exists a unique interior equilibrium. Defining  $g(r_1) \equiv p_2(s_1^*(r_1), r_1)$  and  $h(r_1) \equiv p_2(s_2^*(1-r_1), 1-r_1)$  we rewrite (4) as

$$\alpha_1 g(r_1) = \alpha_2 h(1-r_1). \quad (7)$$

Step 1 of this proof established that  $g(r_1)$  is decreasing. An analogous argument establishes that  $h(r_2) = h(1-r_1)$  is decreasing in  $r_2 = 1-r_1$  (increasing in  $r_1$ ).

Further, because  $g(r_1) = p_2(s_1^*(r_1), r_1)$ , the Indada condition

$$\text{for all } s_i \in [0, 1] : p_2(s_i, r_i) > 0 \text{ for all } r_i \in [0, 1] \text{ and } p_2(s_i, r_i) \rightarrow 0 \text{ as } r_i \rightarrow 1 \quad (8)$$

yields

$$\begin{aligned} g(r_1) &> 0 \text{ for all } r_1 \in [0, 1] \text{ and } g(r_1) \rightarrow 0 \text{ as } r_1 \rightarrow 1 \\ h(r_2) &> 0 \text{ for all } r_2 \in [0, 1] \text{ and } h(r_2) \rightarrow 0 \text{ as } r_2 \rightarrow 1, \end{aligned} \quad (9)$$

where the latter can be re-written as

$$h'(1 - r_1) > 0 \text{ for all } r_1 \in (0, 1] \text{ and } h'(1 - r_1) \rightarrow 0 \text{ as } r_1 \rightarrow 0, \quad (10)$$

By (9), when  $r_1$  tends to one, LHS of (7) tends to zero and RHS is strictly greater than zero. By (10), when  $r_1$  tends to zero, RHS tends to zero and LHS is strictly greater than zero. Thus, there exists an interior equilibrium  $r_1^* \in (0, 1)$ , as these must cross. Moreover, they cross at most once, so the solution  $r_1^*$  is unique. By arguments analogous to those given above, the second-order condition is satisfied.

From the above two steps, we conclude that the unique interior equilibrium is given by  $(r_1^*, s_1^*, s_2^*) \in (0, 1)^3$ , where  $r_1^* \in (0, 1)$  is the solution derived above,  $s_1^* = s_1^*(r_1^*)$ , and  $s_2^* = s_2^*(1 - r_1^*)$ .

Third, we show that there exists no equilibrium in which the *DM* devotes all her attention to only one of the experts. Suppose that there exists some equilibrium in which  $r_i^* = 0$  for some  $S_i$ , w.l.o.g. for  $S_1$ . However, (8) implies that, for any  $s_1$ ,  $\frac{\partial p(s_1, r_1)}{\partial r_1} \rightarrow 0$  as  $r_1 \rightarrow 1$  and  $\frac{\partial p(s_2, r_2)}{\partial r_2} > 0$  as  $r_1 \rightarrow 1$ . Thus,  $r_i^* = 0$  cannot be optimal for the *DM*. Note that this is the case even if  $s_1^* = 1$ .

**Claim A1.2** When  $\alpha_2 \leq \alpha^* < \alpha_1$ , there exists a unique equilibrium  $(r_2^*, s_2^*, 0)$ .

**Proof of Claim A1.2** When  $\alpha_1 > \alpha^*$  and  $\alpha_2 \leq \alpha^*$ ,  $S_1$ 's problem is given by

$$\max_{s_1 \geq 0} \{d - p(s_1, r_1)(1 - \alpha_1)d - c(s_1)\}.$$

Because the first-order derivative w.r.t.  $s_1$  is negative,  $s_1^* = 0$ .

The problem of  $S_2$  is identical to the experts' problem in the case when  $\alpha_1, \alpha_2 \leq \alpha^*$  as above. Thus,  $S_2$ 's (unique) best reply function  $s_2^*(\hat{r}_2)$  is monotonically increas-

ing if  $p_{12} > 0$  and monotonically decreasing if  $p_{12} < 0$ .

The problem of the *DM* is given by

$$\max_{r_1 \in [0,1]} \{ \alpha_1 \bar{x} + (1 - \alpha_1) \underline{x} - p(s_1, r_1) (1 - \alpha_1) \underline{x} + \bar{x} \alpha_2 p(s_2, 1 - r_1) \}.$$

For given (conjectured)  $\hat{s}_1$  and  $\hat{s}_2$ , her first-order conditions (FOCs) are given by

$$\Leftrightarrow -\underline{x} (1 - \alpha_1) p_2(\hat{s}_1, r_1) = \bar{x} \alpha_2 p_2(\hat{s}_2, 1 - r_1).$$

We substitute in the experts' best reply functions and obtain

$$-\underline{x} (1 - \alpha_1) p_2(0, r_1) = \bar{x} \alpha_2 p_2(s_2^*(1 - r_1), 1 - r_1). \quad (11)$$

Because  $-\underline{x} (1 - \alpha_1) > 0$  and  $p_2(0, r_1)$  is decreasing in  $r_1$ , LHS of (11) is decreasing in  $r_1$ . Replicating the steps in the proof of case 1 above establishes that RHS is increasing in  $r_1$ . As the Inada conditions stated in the proof of case 1 are defined for all  $s_1 \in [0, 1]$ , and hence for  $s_1 = 0$ , an analogous argument yields that there exists a unique solution  $r_1^* \in (0, 1)$  to (11). Thus, there exists a unique interior equilibrium  $(r_2^*, s_2^*, s_1^*) \in (0, 1)^2 \cup \{0\}$  of this game. Moreover, replicating the steps in the proof of case 1 establishes that there exists no equilibrium in which  $r_i^* = 0$  for some  $i$ .

We note that in this equilibrium, the *DM* engages in (one-sided) information acquisition relating to  $A_1$ , i.e., she devotes some attention to this project even though  $S_1$  makes no communication effort. In contrast, the *DM* and  $S_2$  engage in two-sided communication.

**Claim A1.3** When  $\alpha_1, \alpha_2 > \alpha^*$  for, there exists a unique equilibrium  $(r_1^*, 0, 0) \in (0, 1) \cup \{0\} \cup \{0\}$  of this game.

**Proof of Claim A1.3** Both experts' problems are given by the problem of  $S_1$  in the proof of Claim 2. Hence,  $s_1^* = s_1^* = 0$ . The problem of the *DM* is given by

$$\max_{r_1 \in [0,1]} \{ \alpha_1 \bar{x} + (1 - \alpha_1) \underline{x} - p(s_1, r_1) (1 - \alpha_1) \underline{x} + \alpha_2 \bar{x} + (1 - \alpha_2) \underline{x} - p(s_2, 1 - r_1) (1 - \alpha_2) \underline{x} \}.$$

For given (conjectured)  $\hat{s}_1$  and  $\hat{s}_2$ , her first-order conditions (FOCs) are given by

$$\Leftrightarrow -\underline{x} (1 - \alpha_1) p_2(\hat{s}_1, r_1) = -\underline{x} (1 - \alpha_2) p_2(\hat{s}_2, 1 - r_1).$$

We substitute in the experts' best reply functions and obtain

$$(1 - \alpha_1)p_2(0, r_1) = (1 - \alpha_2)p_2(0, 1 - r_1). \quad (12)$$

An argument that is analogous to those presented in the proofs of Claim A1.1 and Claim A1.2 yields that there exists a unique solution  $r_1^* \in (0, 1)$  to (12).

## B.2 Proof of Proposition 1

**Claim A2.1** Fix the attractiveness of expert 2's action,  $\alpha_2$ . The DM's attention devoted to Expert 1,  $r_1^*(\alpha_1)$ , is non-monotonic in  $\alpha_1$ .

**Proof of Claim A2.1** When  $\alpha_1 \leq \alpha^*$ , an increase in  $\alpha_1$  affects the DM (and Expert 1) in two ways. First, it becomes more likely that the DM benefits from  $A_1$ . This direct effect makes communication more attractive, for both the DM and Expert 1. Second, the increase in  $\alpha_1$  has an indirect effect on the DM through its effect on Expert 1, and vice versa. Due to complementarity, an increase in one team member's effort raises the marginal productivity of the counterpart's effort. The direct and indirect effects thus reinforce each other, so both  $r_1^*(\alpha_1)$  and  $s_1^*(\alpha_1)$  are increasing in  $\alpha_1$ .

When  $\alpha_1 > \alpha^*$ , as  $\alpha_1$  increases, the DM becomes more convinced that  $\tilde{x}_1 = \bar{x}$ , so the marginal value of acquiring information decreases. Hence,  $r_1^*(\alpha_1)$  is decreasing in  $\alpha_1$ . From Lemma 1, we know that  $s_1^*(\alpha_1) = 0$  in this region.

At  $\alpha^*$ , the DM's default choice changes from not taking  $A_1$  to taking  $A_1$ , so the expert's communication effort drops to zero. Due to complementarity, this lowers the marginal benefit of the DM's effort, so her attention drops discontinuously.

**Claim A2.2** Fix the attractiveness of expert 2's action,  $\alpha_2$ . The expected utility of Sender 1 in equilibrium increases continuously with  $\alpha_1$  for  $\alpha_1 \in (0, \alpha^*)$ , increases discontinuously at  $\alpha^*$ , and increases continuously for  $\alpha_1 \in (\alpha^*, 1)$ .

**Proof of Claim A2.2** For any  $\alpha_1 \in (0, 1)$ , an increase in  $\alpha_1$  has a positive direct effect on the utility of Expert 1: for given effort levels on the part of Expert 1 and the DM, an increase in  $\alpha$  raises the probability that a trade will occur. In addition to this direct effect, an increase in  $\alpha_1$  affects Expert 1 because the optimal efforts change. I show that this second effect reinforces the direct effect.

We start from  $\alpha_1 = \alpha_L < \alpha^*$ , and the associated equilibrium  $(r_1^*(\alpha_L), s_1^*(\alpha_L), s_2^*(\alpha_L)) \in (0, 1)^3$  (for a given  $\alpha_2$ ). I compare Expert 1's expected utility in this equilibrium to that in an equilibrium where  $\alpha_1 = \alpha_H = \alpha_L + \varepsilon$ ,  $\alpha_H < \alpha^*$ . The equilibrium associated with  $\alpha_H$ ,  $(r_1^*(\alpha_H), s_1^*(\alpha_H), s_2^*(\alpha_H)) \in (0, 1)^3$ , satisfies  $r_1^*(\alpha_H) > r_1^*(\alpha_L)$  and  $s_1^*(\alpha_H) > s_1^*(\alpha_L)$ . When  $\alpha_1 \leq \alpha^*$ , for a given level of effort on the part of Expert 1, his expected utility is increasing with the attention that he gets from the DM. Thus, even if Expert 1's effort were held fixed at  $s_1^*(\alpha_L)$  when  $\alpha_1 = \alpha_H$ , Expert 1 would be strictly better off getting attention  $r_1^*(\alpha_H)$  from the receiver than getting attention  $r_1^*(\alpha_L) < r_1^*(\alpha_H)$ . Clearly, then, Expert 1 is strictly better off in the equilibrium associated with  $\alpha_H$ —where the DM devotes attention  $r_1^*(\alpha_H)$  to him and he plays his best reply,  $s_1^*(\alpha_H)$ —than in the equilibrium associated with  $\alpha_L$ . Hence, the expected utility of Expert 1 in equilibrium increases with  $\alpha_1$  for  $\alpha_1 \in (0, \alpha^*)$ . Because all best reply functions and utility functions are continuous, the expected utility increases continuously.

When  $\alpha_1 > \alpha^*$ , his expected utility is decreasing with the attention that he gets from the DM. Sender 1's effort is fixed at zero when  $\alpha_1 > \alpha^*$ ; and the DM's attention  $r_1^*(\alpha_1)$  is decreasing with  $\alpha_1$ . Thus, as  $\alpha_1$  increases, Expert 1's expected utility increases because he gets less (undesirable) attention from the receiver. Because all best reply functions and utility functions are continuous, the expected utility decreases continuously.

At  $\alpha^*$ , Expert 1's effort cost drops discontinuously (to zero); moreover, the attention he receives drops discontinuously as the DM's decision rule changes from an opt-in to an opt-out rule. Both of these changes raise Expert 1's expected utility discontinuously.

**Claim A2.3** When Expert 2 wants the DM's attention ( $\alpha_2 \leq \alpha^*$ ), Expert 2's expected utility is a strictly decreasing function of the attention given to the other expert,  $r_1^*(\alpha_1)$ .

**Proof of Claim A2.3** This follows immediately from the facts that (i)  $U_{S_2}(\alpha_1)$  is increasing in  $r_2^*(\alpha_1)$  for  $\alpha_2 \leq \alpha^*$ , and (ii)  $r_2^*(\alpha_1) = 1 - r_1^*(\alpha_1)$ . Here, (i) follows from Claims 1 and 2, and (ii) is the DM's budget constraint.

**Claim A2.4** When Expert 2 does not want the DM's attention ( $\alpha_2 > \alpha^*$ ), Expert 2's expected utility is a strictly increasing function of the attention given to the other



expert,  $r_1^*(\alpha_1)$ .

**Proof of Claim A2.4** This follows immediately from the facts that (i)  $U_{S_2}(\alpha_1)$  is decreasing in  $r_2^*(\alpha_1)$  for  $\alpha_2 > \alpha^*$ , and (ii)  $r_2^*(\alpha_1) = 1 - r_1^*(\alpha_1)$ . Here, (i) follows from Claims 1 and 2, and (ii) is the DM's budget constraint.

### B.3 Proof of Corollary 1

This follows from Claims 1, 3, and 4 of the proof of Proposition 1.

### B.4 Proof of Proposition 2

I first establish a preliminary result:

**Lemma** (Symmetric information outcome). *Assume that the DM faces a cognitive constraint such that there exists a lower bound on the amount of (non-zero) attention that she can give to any one sender;  $r_i \in \{0\} \cup [\underline{r}, 1]$  for all  $i$ . Denote by  $t(\underline{r})$  the highest number of senders that the DM can split her attention between if she splits her attention equally among them, given the cognitive constraint  $\underline{r}$ . Assume that there are  $N_{\bar{\alpha}}$  high-quality types ( $\alpha = \bar{\alpha}$ ) and  $N_{\underline{\alpha}}$  low-quality types ( $\alpha = \underline{\alpha}$ ), with  $t(\underline{r}) < N_{\bar{\alpha}} \ll N_{\underline{\alpha}}$  and  $\alpha < \bar{\alpha} < \alpha^*$ . Under symmetric information, there is an (essentially unique) equilibrium in which the DM communicates with exactly  $t(\underline{r})$  high-quality types.*

I establish this Lemma in three steps.

**Claim A4.1** Assume that there are  $N_{\bar{\alpha}} \geq 2$  identical experts with  $\alpha = \bar{\alpha} < \alpha^*$ . Then, there exists a unique equilibrium of this game, in which  $r_i^* = \frac{1}{N_{\bar{\alpha}}}$ .

**Proof of Claim A4.1** The problem of  $S_i$ ,  $i \in \{1, 2, \dots, N_{\bar{\alpha}}\}$  is characterized in the proof of Lemma 1, and  $S_i$ 's unique best reply function is given by  $s_i^*(\hat{r}_i) \in (0, 1)$ . By symmetry,  $s_1^*(\cdot) = \dots = s_{N_{\bar{\alpha}}}^*(\cdot) \equiv s^*(\cdot)$ . The problem of the DM is given by

$$\max_{\{r_1, r_2, \dots, r_{N_{\bar{\alpha}}}\} \in [0, 1]^{(N_{\bar{\alpha}}-1)}} \left\{ \bar{x}\alpha \left( p(s_1, r_1) + \dots + p\left(s_{N_{\bar{\alpha}}}, 1 - \sum_{i=1}^{i=N_{\bar{\alpha}}-1} r_i\right) \right) \right\}$$

For given (conjectured)  $\widehat{s}_1, \dots, \widehat{s}_{N_{\bar{\alpha}}}$ , her first-order conditions yield

$$\frac{\partial p(\widehat{s}_1, r_1)}{\partial r_1} = \dots = \frac{\partial p(\widehat{s}_{N_{\bar{\alpha}}-1}, r_{N_{\bar{\alpha}}-1})}{\partial r_{N_{\bar{\alpha}}-1}} = \frac{\partial p\left(\widehat{s}_{N_{\bar{\alpha}}}, 1 - \sum_{i=1}^{i=N_{\bar{\alpha}}-1} r_i\right)}{\partial\left(1 - \sum_{i=1}^{i=N_{\bar{\alpha}}-1} r_i\right)}.$$

Substituting the experts' best reply functions into this condition yields

$$\frac{\partial p(s^*(r_1), r_1)}{\partial r_1} = \dots = \frac{\partial p\left(s^*\left(1 - \sum_{i=1}^{i=N_{\bar{\alpha}}-1} r_i\right), 1 - \sum_{i=1}^{i=N_{\bar{\alpha}}-1} r_i\right)}{\partial\left(1 - \sum_{i=1}^{i=N_{\bar{\alpha}}-1} r_i\right)}. \quad (13)$$

Clearly,  $r_1^* = r_2^* = r_3^* = \dots = r_{N_{\bar{\alpha}}}^* \equiv r^*$  satisfies (13). Because the *DM* exhausts her attention constraint in any equilibrium, there exists a unique symmetric equilibrium of this game, given by  $(r^*, \mathbf{s}^*) = \left(\frac{1}{N_{\bar{\alpha}}}, \mathbf{s}^*\left(\frac{1}{N_{\bar{\alpha}}}\right)\right)$ , where  $\mathbf{s}^*\left(\frac{1}{N_{\bar{\alpha}}}\right)$  is a vector  $(s_1^*, \dots, s_{N_{\bar{\alpha}}}^*)$  such that  $s_i^* = s^*\left(\frac{1}{N_{\bar{\alpha}}}\right)$  for all  $i$ .

There exists no asymmetric interior equilibrium (where  $r_i^* > 0$  for all  $i$  and  $r_i^* \neq r_j^*$  for some  $i, j$  such that  $i \neq j$ ). To see this, define  $g(r_i) \equiv p_2(s^*(r_i), r_i)$ . By the proof of Lemma 1,  $g(r_i)$  is strictly increasing in  $r_i$ . Hence, if  $r_i^* \neq r_j^*$  for some  $i, j$  such that  $i \neq j$ , (13) must be violated.

There exists no equilibrium such that  $r_i^* = 0$  for some  $i$ . This follows directly from (i) for all  $s_i \in [0, 1]$ :  $\frac{\partial p(s_i, r_i)}{\partial r_i} > 0$  for all  $r_i \in (0, 1)$ , (ii)  $\frac{\partial p(s_i, r_i)}{\partial r_i} \rightarrow 0$  as  $r_i \rightarrow 1$ , and (iii)  $\frac{\partial p(s_i, r_i)}{\partial r_i} \rightarrow \infty$  as  $r_i \rightarrow 0$ .

Thus, the symmetric equilibrium is the unique equilibrium of this game.

**Claim A4.2** Assume that the *DM* faces a cognitive constraint such that there exists a lower bound on the amount of (non-zero) attention that she can give to any one sender;  $r_i \in \{0\} \cup [\underline{r}, 1]$  for all  $i$ . Then, there exists a unique equilibrium of this game, in which  $r_i^* = \frac{1}{t(\underline{r})}$ , where  $t(\underline{r})$  is the highest number of senders that the *DM* can split her attention between, given the cognitive constraint  $\underline{r}$ .

**Proof of Claim A4.2** In the unconstrained optimum derived in the proof of Claim 1, as  $N_{\bar{\alpha}}$  increases,  $r^* = \frac{1}{N_{\bar{\alpha}}} \equiv r^*(N_{\bar{\alpha}})$  decreases monotonically. Thus, there exists some integer  $t(\underline{r})$  such that  $r^*(t(\underline{r})) > \underline{r} > r^*(t(\underline{r}) + 1)$ . By the proof of Claim 1, the *DM* strictly prefers to communicate with  $t(\underline{r})$  senders over communicating

with strictly fewer senders. Because the *DM* exhausts her attention constraint in any optimum, there exists a unique symmetric equilibrium of this game, given by  $(r^*, \mathbf{s}^*) = (\frac{1}{t}, \mathbf{s}^*(\frac{1}{t}))$ , where  $\mathbf{s}^*(\frac{1}{t})$  is a vector  $(s_1^*, \dots, s_t^*)$  such that  $s_i^* = s^*(\frac{1}{t})$  for all  $i$ . This implies that the *DM* fares worse in an equilibrium where less than  $t$  high type experts enter than in an equilibrium where  $t$  (or more) high-quality types enter.

**Claim A4.3** Assume that the *DM* faces a cognitive constraint  $\underline{r}$  and that there are  $N_{\bar{\alpha}}$  high-quality types ( $\alpha = \bar{\alpha}$ ) and  $N_{\underline{\alpha}}$  low-quality types ( $\alpha = \underline{\alpha}$ ), with  $t(\underline{r}) < N_{\bar{\alpha}} \ll N_{\underline{\alpha}}$ . Under symmetric information, the *DM* communicates with  $t(\underline{r})$  high-quality types.

**Proof of Claim A4.3** Because  $\bar{\alpha} > \underline{\alpha}$ , and because the *DM*'s expected utility from communication with an expert is increasing with the expert's type ( $\alpha$ ), the *DM*'s expected utility from devoting attention  $r = 1/t(\underline{r})$  to a high-quality type is higher than her expected utility from devoting the same amount of attention to a low-quality type. Because  $t(\underline{r}) < N_{\bar{\alpha}}$ , the *DM* only communicates with high types. By the proof of Claim 2, the *DM* communicates with exactly  $t(\underline{r})$  high-quality types.

Having established the Lemma, the proof now proceeds in four steps.

**Claim A4.4** When  $q_S \in (\underline{q}_S, \bar{q}_S)$ , there exists a fully revealing equilibrium where only high-quality experts approach the *DM*. She obtains the same expected decision payoff as under perfect information.

**Proof of Claim A4.4** We derive conditions under which equilibria with cue communication exist. We postulate an equilibrium such that  $P$  high-quality types send cues to the *DM*, where  $t \leq P \leq N_{\bar{\alpha}}$ , zero low-quality types send a cue to the *DM*, and the *DM* devotes  $r_t^* = 1/t$  to  $t$  experts chosen randomly among the  $P$  high-quality types who send a cue, where  $t \equiv t(\bar{r})$ , and zero attention to all other experts. In such an equilibrium, a low-quality type refrains from sending a cue iff

$$q_S > t \frac{1}{P+1} \left[ d\alpha p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) - c(s_{\underline{\alpha}}^* \left( \frac{1}{t} \right)) \right]. \quad (14)$$

Exactly  $P$  high-quality types send a cue iff

$$t \frac{1}{P+1} [d\bar{\alpha}p(s_{\bar{\alpha}}^*(\frac{1}{t}), \frac{1}{t}) - c(s_{\bar{\alpha}}^*(\frac{1}{t}))] < q_S < t \frac{1}{P} [d\bar{\alpha}p(s_{\bar{\alpha}}^*(\frac{1}{t}), \frac{1}{t}) - c(s_{\bar{\alpha}}^*(\frac{1}{t}))] \quad (15)$$

The fact that  $s_{\bar{\alpha}}^*(\frac{1}{t})$  is a high-quality type's best reply implies that

$$d\bar{\alpha}p\left(s_{\bar{\alpha}}^*\left(\frac{1}{t}\right), \frac{1}{t}\right) - c\left(s_{\bar{\alpha}}^*\left(\frac{1}{t}\right)\right) < d\bar{\alpha}p\left(s_{\bar{\alpha}}^*\left(\frac{1}{t}\right), \frac{1}{t}\right) - c\left(s_{\bar{\alpha}}^*\left(\frac{1}{t}\right)\right). \quad (16)$$

By (16), (15) implies (14). Hence, if (15) is satisfied, the postulated equilibrium is incentive compatible for all experts. By (16), there exists a nonempty range of  $q_S$  such that (15) is satisfied.

Because only high-quality types send cues, an expert's cue communication decision reveals his type, so the  $DM$  need not assimilate the cues. By the proof of Claim 3, the  $DM$ 's preferred attention allocation is to communicate with  $t$  high-quality types. She is indifferent between the  $P$  high-quality types who send cues to her. Thus, randomizing between all experts who send her a cue, and devoting zero attention to all experts who do not, is incentive compatible for the  $DM$ . Hence, the postulated FRE exists if (15) holds.

The FRE is such that all  $N_{\bar{\alpha}}$  high-quality types send cues in equilibrium (but, if there were  $N_{\bar{\alpha}} + 1$  high-quality types, the last one would not enter) when  $q_S$  satisfies

$$t \frac{1}{N_{\bar{\alpha}}+1} [d\bar{\alpha}p(s_{\bar{\alpha}}^*(\frac{1}{t}), \frac{1}{t}) - c(s_{\bar{\alpha}}^*(\frac{1}{t}))] < q_S < t \frac{1}{N_{\bar{\alpha}}} [d\bar{\alpha}p(s_{\bar{\alpha}}^*(\frac{1}{t}), \frac{1}{t}) - c(s_{\bar{\alpha}}^*(\frac{1}{t}))] \quad (17)$$

The left-hand side of (17) gives the expected utility from entry in the presence of  $N_{\bar{\alpha}}$  high-quality types for a (hypothetical)  $(N_{\bar{\alpha}} + 1)$ th high-quality type. The expected utility from entry in the presence of  $N_{\bar{\alpha}}$  high-quality types is strictly smaller for a low-quality type. Thus, denoting this expected utility by  $\underline{q}_S$ , we have

$$\underline{q}_S < t \frac{1}{N_{\bar{\alpha}} + 1} \left[ d\bar{\alpha}p\left(s_{\bar{\alpha}}^*\left(\frac{1}{t}\right), \frac{1}{t}\right) - c\left(s_{\bar{\alpha}}^*\left(\frac{1}{t}\right)\right) \right].$$

There exists a FRE such that all  $N_{\bar{\alpha}}$  high-quality types send cues in equilibrium (but

no low-quality type) when  $q_S$  satisfies

$$\underline{q}_S < q_S < t \frac{1}{N_{\bar{\alpha}}} \left[ d\bar{\alpha}p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) - c \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right) \right) \right]. \quad (18)$$

The FRE is such that exactly  $t$  high-quality types send cues in equilibrium (but no low-quality type) when  $q_S$  satisfies

$$t \frac{1}{t+1} \left[ d\bar{\alpha}p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) - c \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right) \right) \right] < q_S < \left[ d\bar{\alpha}p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) - c \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right) \right) \right]$$

Thus, the FRE in which  $t$  or more high-quality types (but no low-quality type) send cues to the *DM* exist iff  $q_S > \underline{q}_S$  and  $q_S < \left[ d\bar{\alpha}p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) - c \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right) \right) \right] \equiv \bar{q}_S$ . The *DM*'s expected utility, in any of these equilibria, is given by  $U_R^* = \bar{x}\bar{\alpha}tp \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right)$ , which is the same decision payoff as she obtains in the perfect information case.

**Claim A4.5** When  $q_S$  falls below  $\underline{q}_S$ , low-quality experts also approach the receiver. She must either accept a lower expected decision payoff or intensify her search for high-quality senders. In either case, her expected utility is strictly lower than when  $q_S \in (\underline{q}_S, \bar{q}_S)$ .

**Proof of Claim A4.5** We first show that when  $q_S$  falls below  $\underline{q}_S$ , there exist only equilibria that make the *DM* strictly worse off than when  $q_S \in (\underline{q}_S, \bar{q}_S)$ .

When  $q_S$  falls below  $\underline{q}_S$ , (18) is violated. Thus, in any equilibrium with cue communication, at least one low-quality type sends a cue (and all high-quality types). In such an equilibrium, the *DM* either (i) opens zero cues but randomly chooses to communicate with  $t$  experts, or (ii) opens at least one cue. We show that both of these may be consistent with equilibrium play. We show this in the context of an equilibrium in which exactly one low-quality type sends a cue.

If the *DM* plays strategy (i), the equilibrium must satisfy

$$\begin{aligned} \frac{t}{N_{\bar{\alpha}}+2} \left[ s\alpha p \left( x_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) - c \left( x_{\bar{\alpha}}^* \left( \frac{1}{t} \right) \right) \right] < q_S \\ < \frac{t}{N_{\bar{\alpha}}+1} \left[ s\alpha p \left( x_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) - c \left( x_{\bar{\alpha}}^* \left( \frac{1}{t} \right) \right) \right] \end{aligned}$$

where we use the fact that the *DM* will devote the same amount of attention to every expert with whom she communicates (as the experts' types are their private information). We denote by  $U_R^{ran}(N_{\bar{\alpha}}, 1, t)$  the *DM*'s ex ante utility in this random-

ization equilibrium where  $N_{\bar{\alpha}}$  high-quality types and one low-quality type send cues to the  $DM$  who chooses  $t$  experts among them randomly. We have

$$U_R^{ran}(N_{\bar{\alpha}}, 1, t) = \bar{x}\mu(N_{\bar{\alpha}}, 1, t) \left[ \bar{\alpha}p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) + (t-1)\bar{\alpha}p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) \right] \\ + \bar{x}\bar{\alpha}(1 - \mu(N_{\bar{\alpha}}, 1, t))tp \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right)$$

where  $\mu(N_{\bar{\alpha}}, 1, t)$  is the probability that the (only) low-quality type is among the  $t$  experts that the  $DM$  randomly picks from the  $N_{\bar{\alpha}} + 1$  available experts. Because  $\mu(N_{\bar{\alpha}}, 1, t) > 0$ ,  $U_R^{ran}(N_{\bar{\alpha}}, 1, t) < U_R^* = \bar{x}\bar{\alpha}tp \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right)$ .

If the  $DM$  instead plays strategy (ii), and if she commits to opening *exactly one* cue, her expected utility satisfies

$$U_R^{cue}(N_{\bar{\alpha}}, 1, t) \geq -q_R + \frac{1}{N_{\bar{\alpha}} + 1} \bar{x}\bar{\alpha}tp \left( x_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) \\ + \left( \frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + 1} \right) \left( \bar{x}\bar{\alpha}p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) + U_R^{ran}(N_{\bar{\alpha}} - 1, 1, t - 1) \right)$$

With probability  $\frac{1}{N_{\bar{\alpha}} + 1}$ , the  $DM$  opens the cue sent by the (only) low-quality type, in which case she communicates with the  $t$  high-quality types and gets her preferred attention allocation. With probability  $\frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + 1}$ , the cue was sent by a high-quality type, so the  $DM$  communicates with this expert and randomly picks  $(t-1)$  others. In this case, her ex ante expected decision utility is greater than or equal to  $\bar{x}\bar{\alpha}p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) + U_R^{ran}(N_{\bar{\alpha}} - 1, 1, t - 1)$  (where the inequality is strict if the  $DM$  chooses to treat the identified high-quality type preferentially, at the expense of dropping one expert of unknown type). Because the  $DM$  does not obtain her preferred attention allocation with probability one,  $U_R^{cue}(N_{\bar{\alpha}}, 1) < U_R^*$ .

We now show that both (i) and (ii) may, depending on the parameter values, be preferred by the  $DM$ :

In expectation, choosing  $t$  experts at random among  $N_{\bar{\alpha}} + 1$  is strictly worse than observing one high-quality type and choosing the other  $(t-1)$  at random, i.e.,

$$\frac{1}{N_{\bar{\alpha}} + 1} \bar{x}\bar{\alpha}tp \left( x_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) + \frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + 1} \left( \bar{x}\bar{\alpha}p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) + U_R^{ran}(N_{\bar{\alpha}} - 1, 1, t - 1) \right) \\ - U_R^{ran}(N_{\bar{\alpha}}, 1, t) > 0 \quad (19)$$

Equation (19) implies that there exists a non-empty range of  $q_R$  such that  $U_R^{cue}(N_{\bar{\alpha}}, 1, t) > U_R^{ran}(N_{\bar{\alpha}}, 1, t)$ , given by

$$\frac{1}{N_{\bar{\alpha}} + 1} \bar{x} \bar{\alpha} t p \left( x_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) + \frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + 1} \left( \bar{x} \bar{\alpha} p \left( s_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) + U_R^{ran}(N_{\bar{\alpha}} - 1, 1, t - 1) \right) - U_R^{ran}(N_{\bar{\alpha}}, 1, t) \geq q_R. \quad (20)$$

We now note that when the *DM* prefers a strategy in which she commits to opening *exactly one cue* to a strategy in which she randomizes, the *DM* also prefers a strategy in which she opens at least one cue to randomization. Thus, whenever (20) is satisfied, the *DM* opens at least one cue. Because the left-hand side of (20) is finite, the reverse is true for large enough  $q_R$ , i.e.,  $U_R^{cue}(N_{\bar{\alpha}}, 1, t) < U_R^{ran}(N_{\bar{\alpha}}, 1, t)$ .

**Claim A4.6** When  $q_S \rightarrow 0$ , the number of low-quality experts who approach the *DM* becomes so large that she ceases to screen experts for quality. The *DM*'s expected decision payoff is strictly smaller than that obtained in any equilibrium where cue communication takes place.

**Proof of Claim A4.6** Suppose that  $N_{\bar{\alpha}}$  is infinite. As  $q_S \rightarrow 0$ , the number of low-quality types that wish to send a cue to the *DM*,  $n$ , approaches infinity, which implies that  $\frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + n} \rightarrow 1$  and  $\frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + n} \rightarrow 0$ . Thus,  $U_R^{cue}(N_{\bar{\alpha}}, n, t) \rightarrow \bar{x} \bar{\alpha} t p \left( x_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right) - q_R$ . Because  $U_R^{ran}(N_{\bar{\alpha}}, n, t) \rightarrow \bar{x} \bar{\alpha} t p \left( x_{\bar{\alpha}}^* \left( \frac{1}{t} \right), \frac{1}{t} \right)$ , the *DM* strictly prefers not to open any cue in the limit. In an equilibrium in which she randomizes, she is worse off, the larger the share of low-quality experts. Thus, she is clearly worse off than in any equilibrium where she reads cues.

**Claim A4.7** The decrease in the *DM*'s expected utility is monotonic for  $q_S < \underline{q}_S$ .

**Proof of Claim A4.7** When the *DM* randomly chooses experts, this follows directly from the fact that  $\frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + n}$  and  $\frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + n}$  change monotonically with the number of entering low-quality types,  $n$ . So long as the *DM* opens cues, the value of opening one cue is decreasing with  $\frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + n}$ , which is monotonically increasing with  $n$ .

## B.5 Proof of Proposition 3

**Claim A5.1** If the DM assimilates one cue, she continues to assimilate cues until she identifies the relevant topic.

**Proof of Claim A5.1** Suppose that the DM has launched  $k$  topics. Consider the first cue that the DM assimilates. She incurs the cost  $q_R$ . With probability  $1/k$ , she finds the relevant topic, and devotes all of her attention to this topic. With probability  $(k-1)/k$  she does not find the relevant topic. In this situation, the DM always assimilates a second cue: the cost of assimilation is still  $q_R$ ; however, the probability that she identifies the relevant topic is  $1/(k-1) > 1/k$ . Hence, if the DM assimilated the first cue, she assimilates a second cue in the event that the first topic is irrelevant. Repeating this argument yields that, if she assimilates one cue, she continues to assimilate cues until she finds the relevant topic.

**Claim A5.2** There exists a number of cues (topics)  $k^*$  such that, if the DM obtains more than  $k^*$  topics, then she assimilates no cue. Instead, she randomly chooses  $t$  topics that she divides her attention between (equally) in the deliberation stage.

**Proof of Claim A5.2** Consider the DM's expected utility if she assimilates cues. If the first cue that she assimilates is the relevant one, which happens with probability  $1/k$ , then her expected payoff is  $(\pi - q_R)$ , where  $\pi = \alpha\bar{x} + (1-\alpha)\underline{x} - p(0,1)(1-\alpha)\underline{x}$ . That is, her expected payoff is the expected payoff from the action, adjusted for the fact that she may find out, through her information acquisition on the relevant topic, that the product quality is low (and opt out). When she devotes all of her attention to this topic, and the expert devotes zero effort, the probability that she obtains such information is given by  $p(0,1)$  in the event that the product quality indeed is low, which happens with probability  $(1-\alpha)$ . If the first cue that she assimilates is not the relevant one, which happens with probability  $(k-1)/k$ , then she assimilates a second cue.

If the second cue that she assimilates is the relevant one, which happens with probability  $1/(k-1)$ , then her expected payoff is  $(\pi - 2q_R)$ . If the second cue is not the relevant one, then she continues. Repeating this argument yields that her



expected payoff from assimilating cues (until she finds the relevant one) is given by

$$\begin{aligned} & \frac{1}{k}(\pi - q_R) + \frac{(k-1)}{k} \frac{1}{(k-1)}(\pi - 2q_R) + \frac{(k-1)(k-2)}{k} \frac{1}{(k-1)(k-2)}(\pi - 3q_R) + \\ & \dots + \frac{1}{k}(\pi - kq_R) = \frac{1}{k} \sum_{i=1}^{i=k} (\pi - iq_R) = \pi - \frac{q_R}{k}(1 + 2 + \dots + k) = \\ & \pi - \frac{q_R k(1+k)}{2} = \pi - q_R \frac{(1+k)}{2}. \end{aligned}$$

Because  $q_R \frac{(1+k)}{2}$  increases in  $k$  without bound, there exists a  $k^*$  such that

$$\pi - q_R \frac{(1+k^*)}{2} > 0 > \pi - q_R \frac{(1+(k^*+1))}{2}.$$

If the DM does not assimilate any cue, but instead randomly chooses  $t$  out of the  $k$  cues available to her, her expected utility is given by  $\pi' = \alpha \bar{x} + (1-\alpha)\underline{x} - \frac{t}{k}p(0, \frac{1}{t})(1-\alpha)\underline{x}$ , since she chooses  $t/k$  out of the topics available, and hence picks the relevant topic with probability  $t/k$ . Among the  $t$  topics that she randomly chooses, she devotes  $1/t$  of her attention to each of them. Because  $\alpha \bar{x} + (1-\alpha)\underline{x} > 0$ , we have that  $\pi' > 0$ . Clearly, the DM strictly prefers to randomize over assimilating cues if the expert makes more than  $k^*$  topics available. The DM prefers to randomize when her expected payoff from randomization exceeds her expected payoff from opening cues, i.e., when

$$\alpha \bar{x} + (1-\alpha)\underline{x} - \frac{t}{k}p(0, \frac{1}{t})(1-\alpha)\underline{x} > \alpha \bar{x} + (1-\alpha)\underline{x} - p(0, 1)(1-\alpha)\underline{x} - q_R \frac{(1+k)}{2}.$$

We know that this holds when  $k > k^*$ . We denote the smallest number of topics such that the DM prefers to randomize by  $k^{**}$ . Clearly,  $k^{**} \leq k^*$ .

**Claim A5.3** The expert either launches only one topic or launches at least  $k^{**}$  topics. If  $q_S$  is small, he launches at least  $k^{**}$  topics.

**Proof of Claim A5.3** If the expert launches only one topic (the relevant one), then the DM devotes all of her attention to this topic. Thus, she opts out with probability  $p(0, 1)(1-\alpha)$ .

If he launches more than one but fewer than  $k^{**}$  topics, the DM assimilates cues until she finds the relevant topic. Then, she devotes all of her attention to this

topic. Hence, she opts out with the same probability; however, the expert incurred a higher cost of making the (additional) topics available. Thus, the expert strictly prefers launching one topic to launching strictly more than one but fewer than  $k^{**}$ , topics.

If he launches at least  $k^{**}$  topics, the DM randomly chooses  $t$  out of the  $k^{**}$  topics, and devotes attention  $1/t$  to each of the selected topics. In this case, she opts out with probability  $\frac{t}{k}p(0, \frac{1}{t})(1 - \alpha) < p(0, 1)(1 - \alpha)$ . Clearly, if the cost of launching a topic,  $q_S$ , is small enough, the expert strictly prefers to launch at least  $k^{**}$  topics.

**Claim A5.4** When  $q_S = 0$ , the mandate to disclose the relevant topic has no effect on the DM's expected utility; she does not process the relevant information at all.

**Proof of Claim A5.4** When  $q_S \rightarrow \infty$ , the number  $k$  of topics launched goes to  $\infty$ , and the probability that the DM opts out goes to  $\lim_{k \rightarrow \infty} [\frac{t}{k}p(0, \frac{1}{t})(1 - \alpha)] = 0$ . Hence, the mandate to disclose the relevant topic has no effect on the DM's expected utility; the expected utility is simply given by  $\alpha\bar{x} + (1 - \alpha)\underline{x}$ , which is her expected utility in the absence of any mandate.