

Attention Manipulation and Information Overload*

Petra Persson[†]

9 December 2013

Abstract

When a decision-maker's attention is limited, her decisions depend on what she focuses on. This gives interested parties an incentive to manipulate not only the substance of communication but also the decision-maker's attention allocation. This paper models such attention manipulation. In its presence, competitive information supply can reduce the decision-maker's knowledge by causing information overload. Further, a single information provider may deliberately induce information overload to conceal information. These findings, pertinent to consumer protection, suggest a role for rules that restrict communication, mandate not only the content but also the format of disclosure, and regulate product design.

Keywords: Communication, Information Overload, Limited Attention, Persuasion, Disclosure, Complexity, Consumer Protection, Salience

JEL Codes: D82, D83, D18, M38

*I am grateful to Navin Kartik for his support at all stages of this project, as well as to Cass Sunstein for generously providing input into the theory's real-world consumer protection applications. I also thank Patrick Bolton, Yeon-Koo Che, Pierre-André Chiappori, Takakazu Honryo, Samuel Lee, Uliana Loginova, the participants of the Stanford Institute for Economic Policy Research Bi-Weekly Discussion Group, the Columbia University Theory Colloquium and Political Economy Breakfast, and the faculty and participants of the Russell Sage Foundation Summer Institute in Behavioral Economics in Trento for many useful comments.

[†]SIEPR and Department of Economics, Stanford University; Research Institute of Industrial Economics, Stockholm. Email: perssonp@stanford.edu.

1 Introduction

Before deciding whether to take an action, people often communicate with others who know more about it. A mother deciding whether to send her child to a school talks to the headmaster and teachers. An investor deciding whether to put her money in a fund reads its investment strategy or talks to its manager. A homeowner deciding whether to sign up for an insurance plan discusses it with the agent.

Such communication often requires time and devotion. Dewatripont and Tirole (2005) present a model in which the successful exchange of information between a sender and a receiver depends on effort choices of both communicating parties. This seems a plausible characterization of information exchange, in particular ahead of complex decisions: For a headmaster, fund manager, or insurance agent, being persuasive and comprehensible requires effort; for a parent, investor, or homeowner, processing the provided information requires attention.

But a headmaster, fund manager, or insurance agent seldom enjoys the prospective client's undivided attention; she may be faced with a plethora of schools, funds, or insurance products. Because the decision-maker's attention is limited, her decisions hinge on what she pays attention to. This gives the interested parties incentives to engage in *attention manipulation*, that is, strategic actions to influence how the decision-maker allocates her attention.

This paper proposes a framework to analyze attention manipulation and highlights some of its effects on communication. For example, interested parties steal each other's spotlight, which some may or may not like (crowding out); an individual may make worse decisions as more and more interested parties overwhelm her with access to information (information overload); or an interested party can divert an individual's attention from one issue by slanting communication toward another (distraction). In addition, the manipulator can hide specific details from an individual by inundating her with extraneous information (obfuscation) or can design decisions, actions, or products in a way that makes learning difficult (complexity). When at play, attention manipulation can benefit interested parties—such as headmasters, fund managers, or insurance agents—at the expense of the individual.

My point of departure from Dewatripont and Tirole (2005) is to introduce multitasking into their framework.¹ A decision-maker (DM, she) faces one or several binary decisions. Each decision concerns whether to take a certain action. Before making decisions, she can learn more about them. On each decision, she can communicate with one expert (he) about topics that may be relevant. Communication is a moral-hazard-in-teams problem: The more effort the expert expends explaining and the more attention the DM pays, the more likely information is exchanged. Dewatripont and Tirole (2005) refer to this as *issue-relevant communication* because the information exchanged concerns the actual benefits associated with

¹See also Caillaud and Tirole (2007).

the decision. 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.² The DM faces a multitasking problem because she must divide her limited attention between various topics (experts).

Dewatripont and Tirole (2005) contrast issue-relevant communication with *cue communication*. Cue communication does not concern the actual benefits associated with a decision, but rather the decision’s ex ante appeal, that is, the likelihood that (issue-relevant communication will show that) the decision is beneficial. In keeping with this distinction, in parts of the paper I add a stage before issue-relevant communication, in which the expert(s) can send cues—hard information about a decision’s appeal that are costly to send and read. When the DM can engage in issue-relevant communication on only a limited number of topics, cue communication helps the DM select which topics to devote attention to. For this reason, I say that cue communication takes place in the “selection stage” and issue-relevant communication in the “deliberation stage”.³

This paper studies two special cases within the framework. In the first (“multiple experts”), the DM considers several actions and, for each action, communicates with a distinct expert about a single topic. This fits, for example, an investor who considers several funds and talks with each fund manager about his trading philosophy. In the second (“one expert”), the DM communicates with a single expert about several topics, which may pertain to one or several actions. An example of this setting is a consumer who speaks to a retail banker about various details of the credit card contract the bank offers.

Section 2 begins the analysis with the multiple-experts setting. In the deliberation stage, experts expend effort to persuade the DM to make a distinct decision, while she decides how to allocate her attention among learning about the different decisions. Attention substitution leads to externalities I refer to as attention *crowding out*: Each expert ignores that his effort choice affects the attention given to other experts. Interestingly, an expert benefits or suffers from crowding out, depending on whether he welcomes or eschews attention, and an expert’s expected payoff is non-monotonic in the appeal of other experts.

Communication in the selection stage shapes the “supply” of actions and is crucial when the DM can only devote attention to a limited number of actions in the deliberation stage. I

²A fund manager can expend effort to explain the structure of the fund’s management fees, but he does not know how this will impact the prospective client’s decision to invest in the fund.

³The distinction between these two stages 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.” Note that, in my model, the cues represent the information upon which pre-attentive processing is performed.

ask how the DM’s welfare changes as the cost of sending cues—or proposing actions—falls and, consequently, more experts seek attention and more choices enter the picture. Initially, she benefits as she is offered more actions. But as entry becomes cheaper, it becomes profitable for experts who propose actions with a lesser appeal to enter. As a result, the average quality of the proposals 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, and she picks random proposals for deliberation.

Thus, at a certain point, as 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 invitations to communicate decreases with the quantity; worthy topics 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 continues the analysis with three one-expert settings. Even though there is only one expert, the DM still faces a multi-tasking problem, because the expert now communicates with her on several topics. The single-expert settings pick up on the previous results—crowding out and information overload—and show that a single expert may *induce* such outcomes. In the multiple-experts setting, these outcomes result from competition for attention. A single expert who communicates with the DM on multiple topics, however, may deliberately induce crowding out and information overload to divert attention away from a topic.

The starting point is to introduce the idea of an “inconvenient” topic, which arises naturally in this setting. If the DM is already inclined to take the action, the expert wants to minimize the chances that she learns more about it. To him, there is no upside; all that can happen is that she changes her mind and abandons the action. Hence, he wants to draw attention away from it.

A first simple tactic the expert can employ is to fabricate, advertise, and communicate about another topic. I show that he is willing to incur expenses to do so, even if the other topic is irrelevant to him—so long as it is of interest to the DM. The rationale is simply to raise the DM’s marginal gain from paying attention to the other topic, and thereby to divert attention away from the inconvenient one. In short, he plays different topics off against each other to *distract* the DM.

In the second scenario, the DM is unsure which topics are relevant, that is, worth deliberation; in fact, she may even be unaware of some topics. In such a situation, the expert can conceal topics to keep the DM in the dark. In fact, if up to him, no inconvenient topic would be brought up. But sometimes the expert cannot or may not withhold such information. I

show that, if this is the case, he is inclined to inundate the DM with mostly irrelevant topics. This induces information overload and thus conceals inconvenient topics. In other words, the expert shares superfluous information to *obfuscate* the DM. This suggests that simple disclosure rules can be impotent. To have bite, laws may have to stipulate not only whether, but also how, information is disclosed.

Last, I consider a setting in which the DM's payoff from a single action comprises many components, and I allow the expert to manipulate that composition so long as the total payoff stays constant. If each component represents a topic, such payoff-equivalent variations amount to changing the number of relevant topics. This gives the expert yet another tactic to thwart learning. He 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*. He can use this tactic to ensure that an increasing amount of relevant information slips her 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. Even if she fully understands all the topics she can manage to deliberate on, she will always do what she would have done anyway. Note that complexity can lead to obfuscation. But complexity is a more delicate issue for regulation: Unlike obfuscation, it cannot be tackled at the level of communication, information, and disclosure; it may call for intervention in product design.

To summarize, the key mechanism in this framework is attention substitution: More attention devoted to one topic is less devoted to another. Attention substitution allows an expert to take strategic actions to steer the DM's attention towards certain topics and away from others. His choice of persuasion effort on a topic in the deliberation stage, as well as his decision to advertise a topic in the selection stage, affects how much attention is seized from other topics. It is these strategies targeted at the DM's attention allocation that I refer to as attention manipulation.

To my knowledge, this paper presents the first economic model of attention manipulation. It builds on and contributes to three strands of the economics literature.

First, I advance the recent work on two-sided communication as a moral-hazard-in-teams problem (Dewatripont and Tirole, 2005). Soft information can be misrepresented at no cost (Crawford and Sobel, 1982); hard information can be withheld but not misrepresented (Grossman, 1981; Milgrom, 1981). In Dewatripont and Tirole (2005), the softness of information is intermediate and endogenous: Communication conveys hard information with a probability that depends on effort by both sides; otherwise, information remains soft.⁴ I extend

⁴As Dewatripont and Tirole (2005) point out, moral hazard in teams arises even when the expert and the DM have perfectly aligned preferences, as in the literature on team theory (Marschak and Radner, 1972; Sah and Stiglitz, 1986; Radner, 1992; Bolton and Dewatripont, 1994). In the tradition of modeling the choice of communication type as endogenous, Loginova (2011) considers a choice between the two polar cases of soft and hard. Introducing a lying cost represents another way to bridge soft and hard information (see, e.g., Kartik et al., 2007; Kartik, 2009).

their framework by introducing multiple experts, or topics, that vie for the DM’s attention. Competition for attention leads to attention substitution. This in turn invites attention manipulation.

Second, I introduce a novel strategic aspect to the literature on rational inattention. The existing literature analyzes how a DM allocates her limited attention among many passive sources of information (e.g., Sims, 2003; Wiederholt, 2010). I allow the sources to be active by introducing senders who make communication (effort) choices, and in so doing expose a hitherto neglected aspect of rational inattention: Providers of information may strategically influence how the DM allocates her limited attention.⁵ My paper also relates to Gennaioli and Shleifer (2012), who build on Bordalo et al (2012) and model *to which* attributes 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 take actions *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.

Third, my paper is related 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 (2011) study competing experts in a hard information setting, and Chakraborty and Harbaugh (2007, 2010) study soft communication between a DM and one expert on several topics. In these papers, competition or multiplicity never reduces the amount of knowledge the DM gains. Key to the result in my paper that more information can reduce the 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 for whether individuals stand to benefit from more competitive, or greater, information supply.⁶

Following each theoretical result, I discuss practical examples and cite relevant empirical evidence.

⁵This highlights the key difference between standard rational inattention models and multitasking models in the vein of Holmstrom and Milgrom (1991). In the latter, the appeal of directing effort or attention to a specific task is shaped by counterparties (for example, through an incentive contract), and hence is endogenous.

⁶In these models, all the experts have information relevant to the same action. In contrast, I also analyze a case in which each expert has information about a different action. This relates my paper to a literature in organizational design in which multiple division managers communicate local information to a central management (Dessein and Santos, 2006; Alonso et al., 2008). This literature, however, deals with neither limited attention on part of the DM nor competition for such.

2 Multiple Experts

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

2.1 Crowding Out

I introduce multiple senders into the framework proposed by Dewatripont and Tirole (2005). A DM faces two simultaneous decisions, $i = 1, 2$.⁷ 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 α_i and otherwise the value $\underline{x} < 0$. The probability α_i is common knowledge. The larger the α_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 risk-neutral. In the absence of additional information, the DM takes A_i if and only if its expected payoff is positive: $\alpha_i > \alpha^* \equiv \frac{-\underline{x}}{\bar{x}-\underline{x}}$.⁸

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 $r_i \rightarrow 0$.⁹

Successful communication is more likely the more effort the expert exerts on 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

⁷It is not essential that the decisions are simultaneous; only that communication about both decisions is simultaneous.

⁸W.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.

⁹The subscripts refer to the derivative of a function with respect to the i th argument.

the DM's return from paying attention is higher when the expert makes a greater effort to explain.¹⁰ 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.

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 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 noncooperatively. 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 he is vested in. 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

¹⁰Dewatripont and Tirole (2005) study the particular complementary technology $p(s_i, r_i) = s_i r_i$.

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.

How well an expert fares in the deliberation stage not only depends 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, α_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 α_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, α_2 . Expert 2's expected utility is non-monotonic in the attractiveness of expert 1, α_1 .*

This follows from the fact that the attention the DM devotes to expert 1 in equilibrium, $r_1^*(\alpha_1)$, is nonmonotonic in α_1 : $r_1^*(\alpha_1)$ increases for $\alpha_1 \in (0, \alpha^*)$, falls at α^* , and decreases for $\alpha_1 \in (\alpha^*, 1)$. If expert 2 wants the DM's attention ($\alpha_2 \leq \alpha^*$), $\mathbf{EU}_{S_2}(\alpha_1)$ is negatively related to $r_1^*(\alpha_j)$; otherwise, the reverse holds. This is illustrated in Figure 1.

Figure 1 here

Examples: Public Attention and Crowding Out A *New York Times* article gives an example of how one information provider, Sarah Palin, crowded out attention to another, Mitt Romney, on the political stage where candidates compete for public attention:

Ms. Palin had breached campaign decorum by showing up in New Hampshire last week on the very day Mitt Romney was formally announcing his presidential campaign there. ... He had designated Thursday as his "announcement day," and, the decorum police felt, the rest of the field was obliged to stay out of the way in

deference to the “unwritten rule” that says Mr. Romney should have the stage to himself on these special occasions.¹¹

Similarly, in a recent incident—dubbed Speechgate—President Barack Obama scheduled an important speech to Congress for the same day on which Republican candidates, hosted by MSNBC and Politico, were set to have a presidential election debate. This caused a quarrel. As a *Washington Post* blog pointed out, the main issue was crowding out:

[S]cheduling the speech during the GOP debate, even if Boehner had immediately acceded, is the one way the White House could guarantee a) that fewer voters would be watching and that b) viewers and pundits would pay less attention to the speech’s content and more to the theatrics around it. In other words, it’s the easiest way to lessen the speech’s chances at success.¹²

Such issues arise because public attention—or media coverage for that matter—is limited; what makes front page news shapes not only public opinion but also determines its agenda. For example, Eisensee and Strömberg (2007) document that natural disasters that concur with major sports events elicit less relief aid because they are crowded out in the news.

In the above examples, crowding out represents a negative externality. This is not always the case: Those who shun attention on certain topics can certainly benefit from crowding out. In a 2011 *New York Times* column, Thomas L. Friedmans begins with such an example:

Citigroup is lucky that Muammar el-Qaddafi was killed when he was. The Libyan leader’s death diverted attention from a lethal article involving Citigroup that deserved more attention because it helps to explain why many average Americans have expressed support for the Occupy Wall Street movement.¹³

2.2 Information Overload

I assume that the DM can only pay attention to a limited number of things, two in this case.¹⁴ When more than two experts seek the DM’s attention, selection becomes an important issue.

¹¹Leibovich, Mark. 2011. “Sarah Palin and the Politics of Winging It.” *New York Times*, June 4. <http://www.nytimes.com/2011/06/05/weekinreview/05palin.html?pagewanted=all>

¹²James Downie, “Obama-Boehner Speech Spat Should Worry Democrats,” *Partisan Post* (blog), *Washington Post*, August 31, 2011, http://www.washingtonpost.com/blogs/post-partisan/post/obama-boehner-speech-spat-should-worry-democrats/2011/08/31/gIQAlfgssJ_blog.html

¹³Friedman, Thomas L. 2011. “Did You Hear the One about the Bankers?” *New York Times*, October 29. <http://www.nytimes.com/2011/10/30/opinion/sunday/friedman-did-you-hear-the-one-about-the-bankers.html>

¹⁴The assumption that the DM can only devote attention to 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 one topic, $r_i \in \{0\} \cup [\underline{r}, 1]$ for all i . This limits the number of topics she can communicate with to $t \equiv t(\underline{r}) \in \mathbb{N}$. 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 .

To capture this, I add a selection stage in which the DM must select at most two experts for the deliberation stage. $N = N_{\bar{\alpha}} + N_{\underline{\alpha}}$ experts can enter the competition to be selected. Of these, $N_{\bar{\alpha}}$ propose actions of high quality ($\alpha = \bar{\alpha} \leq \alpha^*$) 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, $\underline{\alpha} < \bar{\alpha} < \alpha^*$.

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 which actions to deliberate on, that is, which experts to communicate with. I allow for this in the form of *cue communication*, as in Dewatripont and Tirole (2005): At cost $q_S > 0$, an expert can send a cue 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.

As would be expected, when the cost of entry decreases, the supply of experts—and hence the number of actions the DM can choose among—increases. Proposition 2 states a key result: As the DM's choice set grows, communication first improves but then deteriorates.

Proposition 2 (Information overload). *As $q_S \rightarrow 0$, the DM receives more cues but eventually processes less. 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. In the beginning, cues are so expensive that no expert enters. As q_S falls, it becomes 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 selects her communication partner(s) for the deliberation stage from the pool of entrants at random.¹⁵ 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 she can communicate with in the deliberation stage; otherwise, it is constant.

As q_S falls further, 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 low enough for the DM to engage in active

¹⁵The above 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.

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 meet low quality in the deliberation stage. I refer to this phenomenon—the more cues the DM gets, the less she processes, and the worse she fares—as *information overload*.¹⁶

If I instead assume that the DM cannot observe that a cue was sent unless she assimilates 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 types enter and decreases with the number of low types.

Figure 2 here

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—that 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 and a low 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.

¹⁶The signaling and the screening outcomes arise independent 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.

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 attained. 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.

Last but not least, note that information overload in this setting arises due to externalities: Each expert ignores how his own entry affects the communication environment as a whole. If the expert were identical for all actions, he would only send cues 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. 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, amenable to efficiency-improving intervention.

Example: Information Overload Redux Recent advances in information technology have rapidly increased information production, duplication, and transmission. People 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—also referred to as information deluge, cognitive overload, or information pollution—such an abundance of information can be confusing and counterproductive.

A business research firm nominated information overload as the “problem of the year” for 2008, and claimed that it caused “a \$650 Billion drag on the economy” by way of “lost productivity and innovation.” A productivity study at Intel estimates “the impact of information overload on each knowledge worker at up to eight hours a week.” The main concern is that an escalating quantity of information comes with a declining average quality and that this inverse relationship between amount and relevance—that is, a decreasing signal-to-noise-ratio—makes it harder to find “good” information.¹⁷ This makes selection, as in my model, a daunting issue.

¹⁷The quotes in this paragraph are from Lohr (2007).

Indeed, in 2008, several technology firms—including Microsoft, Intel, Google, and IBM—formed a nonprofit organization, the Information Overload Research Group, to study the problem.¹⁸ Some of them are already in the business of dealing with it, by designing technological solutions for information selection. Google’s success formula, its *ranking* algorithm, is pre-selection. And its ubiquity in the Internet, as gateway and gatekeeper, betrays the import of information overload.¹⁹ The logic of pre-selection also underlies solutions such as email filters. While spam filters are now standard in most email programs, there are more advanced software solutions, such as 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. It can arise in a variety of decision situations, and leads to poor decision-making. 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 with 192 subjects shows that the ability to pick the best 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. Also, both papers report that how information is organized on package displays affects decision quality. O’Reilly (1980) documents similar effects with respect to decision-making in organizations, and stresses the role of organizing communication. Similarly, (Iyengar, 2011) provides empirical evidence that a reduction in choices can benefit decision-makers. 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 (Thaler and Sunstein, 2008). Choice architecture describes how the presentation of choices affects decisions. For example, Cronqvist and Thaler (2004) study the retirement savings plan introduced in Sweden in 1993. Eligible Swedes were encouraged to choose five out of 456 funds, to which their savings would be allocated. The study reports that one third did not make any active choice; their savings were instead allocated to a default fund. The default, a pre-

¹⁸The Information Overload Research Group’s web site is <http://iorgforum.org/>.

¹⁹The quotes in this paragraph are from Lohr (2007). The Information Overload Research Group’s web site is <http://iorgforum.org/>.

selection by the government, was thus important. Information overload seems a likely reason for so many Swedes to 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 are (for exogenous reasons) better represented in the fund catalogue—that is, have better “menu exposure”—receive more active contributions.²⁰

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

3 One expert

In the previous section, crowding out and information overload result from the strategic interaction between multiple experts. In this section, I change the previous model in various ways to study a single expert who has incentives to suppress communication. I formalize three forms of attention manipulation: In the first two, distraction and obfuscation, the expert *induces* crowding out and information overload. In the third, he varies the complexity of the action, keeping its expected payoff constant, only to manipulate communication.

3.1 Distraction

Like before, the DM considers two actions, A_1 and A_2 . But she now faces only one expert. The expert chooses efforts s_1 and s_2 to communicate with the DM about the two actions at cost $c(s_1) + c(s_2)$. I further change the sequence of moves in the deliberation stage: The expert chooses his persuasion effort(s) first, whereas the DM chooses her attention allocation after observing the expert’s efforts. In the selection stage, the expert can send a cue at cost $q_S \geq 0$, and the DM incurs $q_R \geq 0$ to process a cue. For simplicity, let $q_R = 0$.²¹ In addition, I entertain the possibility that only one action exists at the outset but that the expert can “fabricate” A_2 at cost $f > 0$.

To set the stage for distraction, I restrict attention to the case in which the expert deems A_1 an inconvenient topic: The expert wants the action taken ($d_1 > 0$), but because $\alpha_1 > \alpha^*$, he wishes no exchange of information about it. In fact, he wants the DM to devote as little

²⁰Studying retirement savings plans in the United States, ? show that making the problem less complex—a multidimensional problem was collapsed into a binary choice—increases enrollment in the plan. That is, simplification renders active participation more attractive. In a similar study, Choi et al. (2011) report that sending short email cues that draw attention to selective details of the savings program significantly affects participation.

²¹This assumption is not crucial, as I explain below.

attention as possible to A_1 . By contrast, I assume that A_2 is irrelevant to the expert ($d_2 = 0$). With no stake in A_2 , he has no *direct* reason to communicate about it or to fabricate it.

Proposition 3 (Distraction). *The expert (i) fabricates A_2 for some $f > 0$, (ii) sends cues about A_2 , if he privately observes $\alpha_2 \in \{\underline{\alpha}, \bar{\alpha}\}$, for some $q_S > 0$, and (iii) chooses $s_2 > 0$.*

To see this, consider first the subgame perfect equilibrium of the deliberation stage. The expert exerts no effort on the inconvenient topic, $s_1 = 0$, but may exert effort on the irrelevant topic, $s_2 \geq 0$. When the DM chooses her attention allocation, she knows the expert's efforts. Thus, she solves

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

The first-order condition is

$$\frac{\alpha_1}{\alpha_2} p_2(0, r_1) = p_2(s_2, 1 - r_1)$$

Because $p_{12} > 0$, the DM finds it optimal to put more attention on A_2 —that is, to increase $1 - r_1$, or equivalently, to decrease r_1 —when s_2 is larger. The solution to the DM's problem, $r_1^*(s_2, \alpha_1, \alpha_2)$, will thus be decreasing in s_2 : $\frac{\partial r_1^*}{\partial s_2} < 0$.

This means that the expert can manipulate the attention given to the inconvenient topic, A_1 , by varying his effort on the irrelevant topic, s_2 . His problem is

$$\max_{s_2} \{ d_1 - (1 - \alpha_1) p(0, r_1^*(s_2, \alpha_1, \alpha_2)) - c(s_2) \}.$$

We see where the expert's incentive to increase s_2 comes from: He can reduce the attention that the DM pays to A_1 , $r_1^*(s_2, \alpha_1, \alpha_2)$, and thereby the probability that the DM opts out of A_1 , $(1 - \alpha_1) p(0, r_1^*(s_2, \alpha_1, \alpha_2))$. The first-order condition is

$$-\frac{\partial r_1^*}{\partial s_2} (1 - \alpha_1) p_2(0, r_1^*(s_2, \alpha_1, \alpha_2)) = c_1(s_2)$$

The left-hand side is positive. Thus, $s_2 > 0$.

Next, suppose the expert privately knows whether $\alpha_2 = \underline{\alpha}$ or $\alpha_2 = \bar{\alpha}$, and consider the perfect Bayesian equilibrium of the game with the selection stage. Note from the first-order condition of the DM's problem that the DM puts more attention on A_2 when α_2 is larger. Thus, $\frac{\partial r_1^*}{\partial \alpha_2} < 0$. This gives the expert an incentive to reveal $\alpha_2 = \bar{\alpha}$ and conceal $\alpha_2 = \underline{\alpha}$. Clearly, for low enough q_S , there exists a fully revealing equilibrium with pessimistic posture: The DM receives a cue when $\alpha_2 = \bar{\alpha}$ and infers $\alpha_2 = \underline{\alpha}$ when no cue is sent.²²

Finally, in the absence of A_2 , the DM focuses all of her attention on A_1 : $r_1^* = 1$. Thus, for low enough f , the expert will fabricate A_2 to divert some of that attention.

²²If $q_R > 0$, there exists a separating equilibrium in which the high type sends a cue. The DM reads this cue with some positive probability, which is high enough to deter the low type from sending a cue. The equilibrium is constructed such that the DM is indifferent between opening and not opening the cue.

Example: Public Relations and Spin A 2009 *New York Times* article speculated whether Vladimir Putin, then Russian Prime Minister, used his public persona as distraction:

Mr. Putin may be encouraging speculation about his political future to enhance his influence—or to divert attention from more important matters like economic reform or the demographic crisis facing Russia.²³

In a similar vein, a *Times* article suggests that Mahmoud Ahmadinejad, then Iranian President, strategically diverted Western media coverage away from the country’s domestic situation:

President Ahmadinejad also sought to grab the headlines and divert attention from the protests by announcing that Iran had produced its first stock of 20 percent-enriched uranium. He declared that Iran was now a “nuclear state.”²⁴

Tactics like these, used by politicians to manipulate public opinion, have a name in public relations: spin. Two other well-known spin tactics are “bury bad news” and “wag the dog.” The former refers to the practice of releasing bad news in the shadow of other important news events. A revealing anecdote involves an UK government press officer, Jo Moore, who notoriously wrote, “It’s now a very good day to get out anything we want to bury,” in an email on September 11, 2001, following the terrorist attacks on the World Trade Center.²⁵

The latter refers to fashioning a salient issue for the purpose of diverting attention from another (minor) issue. Hard evidence for this is elusive, but allegations abound. For example, the satire “American Hero” claims that George H. W. Bush used the first Iraq War, Operation Desert Storm, to divert attention from domestic issues and ensure reelection (Beinhart, 1994); the book was also the inspiration for “Wag the Dog,” a movie in which a government spin-doctor fabricates a war to take the spotlight off a political sex scandal.

In rhetoric, a tactic of distraction is called a “red herring.” A red herring is an irrelevant detail or issue raised to sidetrack a discussion. If successful, the discussants debate the red herring and disregard the original topic (Gula, 2007). A *Guardian* article claims that Alastair Campbell, then Director of Communications and Strategy in the UK government, used one:

Campbell has managed to turn an argument about the way the government presented its case for war in Iraq into an entirely different dispute about the way the BBC covered what was going on in Whitehall at the time. [...] Brilliant or

²³“Mr. Putin in Perpetuity?” *New York Times*, December 5, 2009, <http://www.nytimes.com/2009/12/06/opinion/06sun2.html>

²⁴“Iranian regime ships in support for anniversary celebrations” *Times*, February 11, 2010.

²⁵Sparrow, Andrew. 2001. “Sept. 11: ‘A Good Day to Bury Bad News,’” *The Telegraph*, October 10. <http://www.telegraph.co.uk/news/uknews/1358985/Sept-11-a-good-day-to-bury-bad-news.html>

not, what Mr Campbell has achieved is largely a classic use of a very pungent red herring. The BBC’s reporting, though important, is not in fact the real issue.²⁶

3.2 Obfuscation

In the next three settings, the DM communicates with one expert about just one action A . Multiple topics pertain to the decision under consideration. Attention manipulation still manifests in the way the expert steers the DM’s attention towards, or away from, certain topics.

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 α_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 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 regard to communication, there exist $N = N_R + N_\emptyset$ topics related to the decision, but only N_R of them are relevant to the DM’s payoff. Crucially, I assume that the DM does not know which topics are relevant, and more, is potentially unaware of topics per se. This might seem peculiar, but it captures a plausible situation: Inclined towards a particular choice, the DM might yet discover aspects that change her opinion. Thus, the DM is faced with two sets of questions: What topics exist, and which ones are relevant (selection stage)? How much attention should a given topic receive (deliberation stage)?

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 she revises her beliefs upwards, and so opts in, he seeks to draw her attention to those topics that are most likely to yield favorable information. That is, he sends cues about, and exerts persuasion effort on, the topics with the highest α_i .

By contrast, if the DM follows an opt-out rule, the expert wants to withhold the relevant information. On his own accord, he would neither advertise nor explain anything of relevance, lest she change her mind and abandon the action after all. If the DM were bound to allocate attention to some relevant topics, the expert might advertise and exert effort on those relevant topics that are least likely to alter her decision. Intuitively, he would lure her attention towards topics that are—to him—the lesser evil. The fact that he prefers to talk about all and only irrelevant topics, in turn, implies that he has an incentive to *supply* irrelevant topics in the selection stage.

To show this, I assume $N_R = 1$ and N_\emptyset is infinite, and that the DM can at most deliberate

²⁶“Labour’s Phoney War.” 2003. The Guardian, June 28. <http://www.guardian.co.uk/media/2003/jun/28/bbc.politicsandthemedia>

on two topics. These assumptions are merely simplifying. Indeed, I obtain the following result as long as the DM can devote attention only to a limited number of topics in the deliberation stage and N_\emptyset is sufficiently large relative to N_R . Finally, 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.

Proposition 4 (Obfuscation). *Let $\bar{\alpha} > \alpha^*$. Suppose the relevant cue is sent out. As $q_S \rightarrow 0$, the expert sends a swarm of irrelevant cues, and the DM’s expected utility decreases.*

The intuition behind this result is the same as for information overload, though the expert in the current setting benefits and hence induces it. 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 a relevant cue, by mandate or by design, must be sent. In this case, the expert is afraid that the DM, by paying attention to the relevant topic, might discover unfavorable information about the action and abandon it.

To reduce the odds that the DM identifies, that is, selects the relevant topic, the expert now has an incentive to send out irrelevant cues (even though it is costly for him). A swarm of mostly irrelevant cues thwarts the DM’s chance, and hence her incentives, to pinpoint the relevant topic. In other words, the expert intentionally evokes information overload to conceal relevant information. From a normative point of view, therefore, too much disclosure can be a concern as much as too little. A surfeit of details can prove as uninformative as a dearth thereof.

3.3 Complexity

Using a similar setting, I consider a related form of manipulation. Instead of obfuscating the DM with cues, the expert can now “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 α_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, $\{\tilde{x}_i\}$.²⁷ In any case, absent communication, such variation does not matter; it affects neither the DM’s decision nor her welfare.

²⁷Alternatively, 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 ones.

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, 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 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 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—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 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 to manipulate the DM’s learning process. How the expert uses such means hinges, like 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 5 (Complexity). *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 only deliberate on 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 that 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 to prevent the DM from getting the full picture.

Example: Mandatory Disclosure and Product Regulation *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...²⁸

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. Ordinary customers would be hard-pressed to know not only *where* to look for relevant items but also *what* items to look for.

The financial products market seems rife with such practices. Credit cards are perhaps the most widely debated example. 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.²⁹ According to Edward L. Yingling, president and chief executive of the American Bankers Association, Obama urged the 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.³⁰

Similar debates are waged 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

²⁸Tilghman, 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>

²⁹Alexander, 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>

³⁰Labaton, 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>

the real price.³¹

The UK financial regulator, the Financial Services Authority (FSA), is also quoted 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.

Note that the two quotes refer, respectively, to what I call obfuscation and complexity. The first talks about manipulating the complexity of *information*, whereas the second talks about manipulating the complexity of the *product* itself. In either case, says the article, the danger is that misguided decisions affect “[consumers’] risk of getting into debt.”

These examples along with the theory highlight two interesting aspects that elude previous communication models. First, 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”). Imagine the banks must 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 attention manipulation—conceal what is relevant.

Second, not even the most elaborate disclosure rule may be enough. Jacoby et al. (1974) begin their study on information overload with the question of whether simple disclosure rules like the Truth in Lending Act are effective if information overload poses a concern; this, I add, is a *fortiori* questionable given the threat of obfuscation. Even more unsettling, though, is that *any* regulation at the communication level proves futile if the sender 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, regulation might also want to target the object of communication, which is, as the examples illustrate, a much thornier issue.

4 Conclusion

This paper presents a theoretical framework that captures basic aspects of communication: seeking, shunning, and vying for attention and, deliberately or not, distracting or confusing others. The main results—crowding out, information overload, distraction, obfuscation, and complexity—resonate with empirical findings in various social sciences. The key to micro-modeling these phenomena is to proceed from the idea of attention allocation, namely split-

³¹Daily 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>

ting one's own attention between competing demands, to the idea of attention manipulation, namely influencing what others pay attention to.

A particularly interesting avenue, not pursued in the present paper, is that heterogeneity in attention constraints, or cognitive capacity, might provoke different 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. This paper's findings suggest that the problem may be even worse: The poor may not only start out with tighter attention constraints but also find their limited attention exploited, more so 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

- 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), 489–493.
- Battaglini, Marco**, "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, 2002, *70* (4), 1379–1401.
- Beinhart, L.**, *American Hero*, Century, 1994.
- Bolton, Patrick and Mathias Dewatripont**, "The firm as a communication network," *The Quarterly Journal of Economics*, 1994, *109* (4), 809–839.
- Caillaud, Bernard and Jean Tirole**, "Consensus Building: How to Persuade a Group," *American Economic Review*, 2007, *97* (5), 1877–1900.
- 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.

- Choi, James J., Emily Haisley, Jennifer Kurkoski, and Cade Massey**, “Small Cues Change Savings Choices,” *Unpublished manuscript*, 2011.
- 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), 424–428.
- Dessein, Wouter and Tano Santos**, “Adaptive organizations,” *Journal of Political Economy*, 2006, *114* (5), 956–995.
- 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.
- Eisensee, Thomas and David Strömberg**, “News droughts, news floods, and US disaster relief,” *The Quarterly Journal of Economics*, 2007, *122* (2), 693–728.
- 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.
- Gentzkow, Matthew and Emir Kamenica**, “Competition in persuasion,” *Unpublished manuscript*, 2011.
- Grossman, Stanford J.**, “Informational Role of Warranties and Private Disclosure about Product Quality,” *Journal of Law & Economics*, 1981, *24* (3), 461–483.
- Gula, Robert J.**, “Red Herrings, Straw Men and Sacred Cows: How We Abuse Logic in Our Everyday Language,” *Axios*, 2007.
- Holmstrom, Bengt and Paul Milgrom**, “Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design,” *Journal of Law, Economics, and organization*, 1991, *7* (Special Issue), 24–52.
- 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,” *Unpublished manuscript*, 2006.
- Kartik, Navin**, “Strategic communication with lying costs,” *Review of Economic Studies*, 2009, 76 (4), 1359–1395.
- , **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.
- Loginova, Uliana**, “Strategic communication in networks,” *Unpublished manuscript*, 2011.
- 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.
- Marschak, Jacob and Roy Radner**, “Economic Theory of Teams,” *New Haven, CT: Yale Univ. Press.*, 1972.
- Milgrom, Paul and John Roberts**, “Relying on the information of interested parties,” *The RAND Journal of Economics*, 1986, 17 (1), 18–32.
- Milgrom, Paul R**, “Good news and bad news: Representation theorems and applications,” *The Bell Journal of Economics*, 1981, 12 (2), 380–391.
- O’Reilly, Charles A.**, “Individuals and Information Overload in Organizations: Is More Necessarily Better?,” *The Academy Management Journal*, 1980, 23 (4), 684–684696.
- Radner, Roy**, “Hierarchy: The Economics of Managing,” *Journal of economic literature*, 1992, 30 (3), 1382–1415.
- Sah, Raaj Kumar and Joseph E. Stiglitz**, “The Architecture of Economic Systems: Hierarchies and Polyarchies,” *American Economic Review*, 1986, 76 (4), 716–727.

Simon, HA, “Designing organizations for an information-rich world,” 1971.

Sims, Christopher A., “Implications of rational inattention,” *Journal of Monetary Economics*, 2003, 50 (3), 665–690.

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.

A Figures

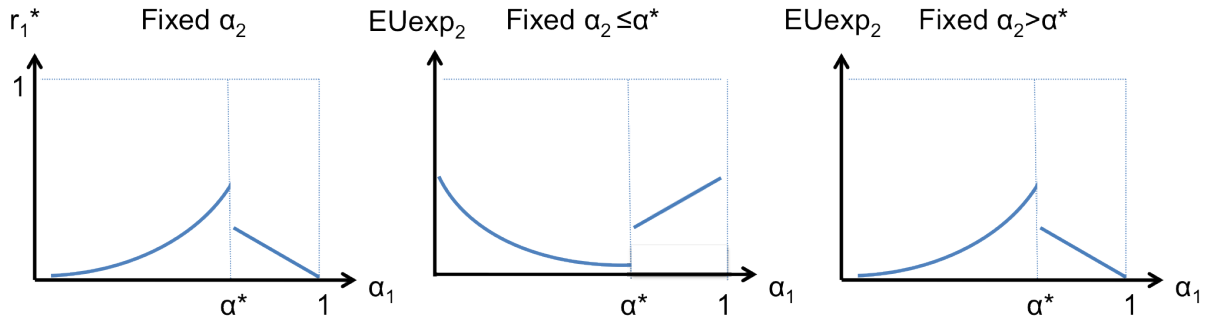


Figure 1: **Crowding out.** The DM’s attention devoted to Expert 1 in equilibrium, $r_1^*(\alpha_1)$, for a given α_2 , is nonmonotonic in Expert 1’s attractiveness (left panel). This makes Expert 2’s utility nonmonotonic in α_1 : When 2 desires attention ($\alpha_2 \leq \alpha^*$), $r_1^*(\alpha_1)$ 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 negative externality on Expert 2, so $\mathbf{EU}_{Exp_2}(\alpha_1)$ is positively related to $r_1^*(\alpha_1)$ (right panel).

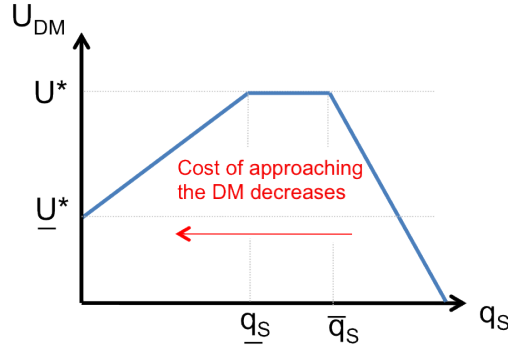


Figure 2: **Information overload.** A movement from right to left on the x-axis represents a decrease in the cost of entry. As q_S decreases, the DM first benefits from an expansion in information supply ($q_S > \bar{q}_S$). Then, the DM has access to (at least) t high-quality experts, but no low-quality experts, so she obtains (decision) payoff U^* . As the cost of entry falls below \underline{q}_S , low types join the battle for access, and the DM's expected utility falls below U^* . When $q_S \rightarrow 0$, the number of low types who enter tends to infinity, so the DM ceases to screen experts. Her decision payoff falls, as she relies on lower quality of information 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

Proofs of Lemma 1, Proposition 1, Corollary 1 and Proposition 4: See Appendix C.

B.1 Proposition 2 (Information Overload)

I first establish a preliminary result, proven in Appendix C:

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 types ($\alpha = \bar{\alpha}$) and $N_{\underline{\alpha}}$ low 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 types.*

The proof now proceeds in four steps.

Claim B1.1 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 B1.1 We derive conditions under which equilibria with cue communication exist. We postulate an equilibrium such that P high types send cues to the DM, where $t \leq P \leq N_{\bar{\alpha}}$, zero low types send a cue to the DM, and the DM devotes $r_t^* = 1/t$ to t experts chosen randomly among the P high types who send a cue, where $t \equiv t(\bar{r})$, and zero attention to all other experts. In such an equilibrium, a low type refrains from sending a cue iff

$$q_S > t \frac{1}{P+1} \left[d\bar{\alpha}p \left(s_{\underline{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) - c \left(s_{\underline{\alpha}}^* \left(\frac{1}{t} \right) \right) \right]. \quad (2)$$

Exactly P high types send a cue iff

$$t \frac{1}{P+1} \left[d\bar{\alpha}p \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) - c \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right) \right) \right] < q_S < t \frac{1}{P} \left[d\bar{\alpha}p \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) - c \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right) \right) \right] \quad (3)$$

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

$$d\bar{\alpha}p \left(s_{\underline{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) - c \left(s_{\underline{\alpha}}^* \left(\frac{1}{t} \right) \right) < d\bar{\alpha}p \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) - c \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right) \right). \quad (4)$$

By (4), (3) implies (2). Hence, if (3) is satisfied, the postulated equilibrium is incentive compatible for all experts. By (4), there exists a nonempty range of q_S such that (3) is satisfied.

Because only high 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 types. She is indifferent between the P high 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 (3) holds.

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

$$t \frac{1}{N_{\bar{\alpha}}+1} \left[d\bar{\alpha}p \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) - c \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right) \right) \right] < q_S < t \frac{1}{N_{\bar{\alpha}}} \left[d\bar{\alpha}p \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) - c \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right) \right) \right] \quad (5)$$

The left-hand side of (5) gives the expected utility from entry in the presence of $N_{\bar{\alpha}}$ high types for a (hypothetical) $(N_{\bar{\alpha}} + 1)$ th high type. The expected utility from entry in the presence of $N_{\bar{\alpha}}$ high types is strictly smaller for a low type. Thus, denoting this expected utility by

\underline{q}_S , we have that

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

There exists a the FRE is such that all $N_{\bar{\alpha}}$ high types send cues in equilibrium (but no low types) when q_S satisfies

$$\underline{q}_S < q_S < t \frac{1}{N_{\bar{\alpha}}} \left[d\bar{\alpha}p \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) - c \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right) \right) \right]. \quad (6)$$

The FRE is such that exactly t high types send cues in equilibrium (but no low types) 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, FRE in which t or more high types (but no low types) 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 B1.2 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 B1.2 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 , (6) is violated. Thus, in any equilibrium with cue communication, at least one low type sends a cue (and all high 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 type sends a cue.

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

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

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 randomization equilibrium where $N_{\bar{\alpha}}$ high

types and one low type send cues to the *DM* who chooses t experts among them randomly. We have that

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

where $\mu(N_{\bar{\alpha}}, 1, t)$ is the probability that the (only) low 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 type, in which case she communicates with the t high types and gets her preferred attention allocation. With probability $\frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + 1}$, the cue was sent by a high 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 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 type and choosing the other $(t-1)$ at random, i.e.,

$$\frac{1}{N_{\bar{\alpha}} + 1} \bar{x}\bar{\alpha}tp \left(x_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) + \frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + 1} \left(\bar{x}\bar{\alpha}p \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) + U_R^{ran}(N_{\bar{\alpha}} - 1, 1, t - 1) \right) \quad (7) \\ - U_R^{ran}(N_{\bar{\alpha}}, 1, t) > 0$$

Equation (7) 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

$$\frac{1}{N_{\bar{\alpha}} + 1} \bar{x} \bar{\alpha} t p \left(x_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) + \frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + 1} \left(\bar{x} \bar{\alpha} p \left(s_{\bar{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) + U_R^{ran}(N_{\bar{\alpha}} - 1, 1, t - 1) \right) \quad (8)$$

$$-U_R^{ran}(N_{\bar{\alpha}}, 1, t) \geq q_R.$$

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 (8) is satisfied, the *DM* opens at least one cue. Because the left-hand side of (8) 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 B1.3 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 B1.3 Suppose that $N_{\underline{\alpha}}$ is infinite. As $q_S \rightarrow 0$, the number of low types that wish to send a cue to the *DM*, n , approaches infinity, which implies that $\frac{N_{\underline{\alpha}}}{N_{\bar{\alpha}} + n} \rightarrow 1$ and $\frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + n} \rightarrow 0$. Thus, $U_R^{cue}(N_{\bar{\alpha}}, n, t) \rightarrow \bar{x} \underline{\alpha} t p \left(x_{\underline{\alpha}}^* \left(\frac{1}{t} \right), \frac{1}{t} \right) - q_R$. 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 type experts. Thus, she is clearly worse off than in any equilibrium where she reads cues.

Claim B1.4 The decrease in the *DM*'s expected utility is monotonic for $q_S < \underline{q}_S$.

Proof of Claim B1.4 When the *DM* randomly chooses experts, this follows directly from the fact that $\frac{N_{\underline{\alpha}}}{N_{\bar{\alpha}} + n}$ and $\frac{N_{\bar{\alpha}}}{N_{\bar{\alpha}} + n}$ change monotonically with the number of entering low types, n . So long as the *DM* opens cues, the value of opening one cue is decreasing in $\frac{N_{\underline{\alpha}}}{N_{\bar{\alpha}} + n}$, which is monotonically increasing in n .

ONLINE APPENDIX

Attention Manipulation and Information Overload

Petra Persson

C Omitted Proofs

C.1 Lemma 1

Claim C1.1 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 C1.1 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

$$d\alpha_i p_1(s_i, \hat{r}_i) = c'(s_i). \quad (9)$$

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 yield 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

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

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

$$\bar{x} (\alpha_1 p_2(\hat{s}_1, r_1) - \alpha_2 p_2(\hat{s}_2, 1 - r_1)) = 0, \quad (10)$$

where $\bar{x} > 0$.

We substitute the experts best reply functions from (9) into (10) and obtain

$$\alpha_1 p_2(s_1^*(r_1), r_1) = \alpha_2 p_2(s_2^*(1 - r_1), 1 - r_1). \quad (11)$$

An equilibrium which is interior must satisfy (11). We will now discuss 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

$$p_{21}(s_1^*(r_1), r_1) s_1^{*'}(r_1) + p_{22}(s_1^*(r_1), r_1). \quad (12)$$

Differentiating sender 1's equilibrium condition, (9), with respect to r_1 yields

$$\begin{aligned} d\alpha_i (p_{11}(s_1^*(r_1)) s_1^{*'}(r_1) + p_{12}(s_1^*(r_1))) &= c''(s_1^*(r_1)) s_1^{*'}(r_1), \\ s_1^{*'}(r_1) &= \frac{d\alpha_i p_{12}(s_1^*(r_1))}{c''(s_1^*(r_1)) - d\alpha_i p_{11}(s_1^*(r_1))} \end{aligned} \quad (13)$$

Inserting (13) into (12) 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 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 (11) as

$$\alpha_1 g(r_1) = \alpha_2 h(1-r_1). \quad (14)$$

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

$$\text{for all } s_i \in [0, 1] : p_2(s_i, r_i) > 0 \text{ for all } r_i \in [0, 1) \text{ and } p_2(s_i, r_i) \rightarrow 0 \text{ as } r_i \rightarrow 1 \quad (15)$$

yields that

$$\begin{aligned} g(r_1) &> 0 \text{ for all } r_1 \in [0, 1) \text{ and } g(r_1) \rightarrow 0 \text{ as } r_1 \rightarrow 1 \\ h(r_2) &> 0 \text{ for all } r_2 \in [0, 1) \text{ and } h(r_2) \rightarrow 0 \text{ as } r_2 \rightarrow 1, \end{aligned} \quad (16)$$

where the latter can be re-written

$$h'(1 - r_1) > 0 \text{ for all } r_1 \in (0, 1] \text{ and } h'(1 - r_1) \rightarrow 0 \text{ as } r_1 \rightarrow 0, \quad (17)$$

By (16), when r_1 tends to one, LHS of (14) tends to zero and RHS is strictly greater than zero. By (17), 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 above, second-order condition is satisfied.

From the above two steps, we have 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, (15) yields 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 C1.2 When $\alpha_2 \leq \alpha^* < \alpha_1$, there exists a unique equilibrium $(r_2^*, s_2^*, 0)$.

Proof of Claim C1.2 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 $\alpha_1, \alpha_2 \leq \alpha^*$ above. Thus, S_2 's (unique) best reply function $s_2^*(\hat{r}_2)$ is monotonically increasing 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) \widehat{s}_1 and \widehat{s}_2 , her first-order conditions (FOCs) are given by

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

We substitute in the experts' best reply functions and obtain

$$-\underline{x}(1 - \alpha_1)p_2(0, r_1) = \bar{x}\alpha_2 p_2(s_2^*(1 - r_1), 1 - r_1). \quad (18)$$

Because $-\underline{x}(1 - \alpha_1) > 0$ and $p_2(0, r_1)$ is decreasing in r_1 , LHS of (18) 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 (18). 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 exist 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 does not make any communication effort. In contrast, *DM* and S_2 engage in two-sided communication.

Claim C1.3 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 C1.3 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 *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) \widehat{s}_1 and \widehat{s}_2 , her first-order conditions (FOCs) are given by

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

We substitute in the experts' best reply functions and obtain

$$(1 - \alpha_1)p_2(0, r_1) = (1 - \alpha_2)p_2(0, 1 - r_1). \quad (19)$$

An analogous argument to those in the proofs of Claim 1 and Claim 2 yields that there exists a unique solution $r_1^* \in (0, 1)$ to (19).

C.2 Proposition 1

Claim C2.1 Fix the attractiveness of expert 2's action, α_2 . The DM's attention devoted to Expert 1, $r_1^*(\alpha_1)$, is non-monotonic in α_1 .

Proof of Claim C2.1 When $\alpha_1 \leq \alpha^*$, an increase in α_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 α_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 counterparty'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 α_1 .

When $\alpha_1 > \alpha^*$, as α_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 α_1 . From Lemma 1, we know that $s_1^*(\alpha_1) = 0$ in this region.

At α^* , 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 C2.2 Fix the attractiveness of expert 2's action, α_2 . The expected utility of Sender 1 in equilibrium increases continuously with α_1 for $\alpha_1 \in (0, \alpha^*)$, increases discontinuously at α^* , and increases continuously for $\alpha_1 \in (\alpha^*, 1)$.

Proof of Claim C2.2 For any $\alpha_1 \in (0, 1)$, an increase in α_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 α raises the probability that trade will occur. In addition to this direct effect, an increase in α_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 α_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 α_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 in 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 α_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 α_L . Hence,

the expected utility of Expert 1 in equilibrium increases with α_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 in 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 in α_1 . Thus, as α_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 α^* , 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 C2.3 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 C2.3 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 C2.4 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 C2.4 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.

C.3 Corollary 1

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

C.4 Proposition 2

I establish the **Lemma** in three steps.

Claim C4.1 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 C4.1 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) $\hat{s}_1, \dots, \hat{s}_{N_{\bar{\alpha}}}$, her first-order conditions yield

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

Substituting the experts' best reply functions into this condition yields

$$\frac{\partial p(s^*(r_1), r_1)}{\partial r_1} = \dots = \frac{\partial p\left(s^*\left(1 - \sum_{i=1}^{i=N_{\bar{\alpha}}-1} r_i\right), 1 - \sum_{i=1}^{i=N_{\bar{\alpha}}-1} r_i\right)}{\partial \left(1 - \sum_{i=1}^{i=N_{\bar{\alpha}}-1} r_i\right)}. \quad (20)$$

Clearly, $r_1^* = r_2^* = r_3^* = \dots = r_{N_{\bar{\alpha}}}^* \equiv r^*$ satisfies (20). 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 $i \neq j$, (20) 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 C4.2 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; $r_i \in \{0\} \cup [r, 1]$ for all i . Then, there exists a unique equilibrium of this game, in which $r_i^* = \frac{1}{t(r)}$, where $t(r)$ is the highest number of senders that the *DM* can split her attention between, given the cognitive constraint r .

Proof of Claim C4.2 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 types enter.

Claim C4.3 Assume that the *DM* faces a cognitive constraint \underline{r} and that there are $N_{\bar{\alpha}}$ high types ($\alpha = \bar{\alpha}$) and $N_{\underline{\alpha}}$ low 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 types.

Proof of Claim C4.3 Because $\bar{\alpha} > \underline{\alpha}$, and because the *DM*'s expected utility from communication with an expert is increasing in the expert's type (α), the *DM*'s expected utility from devoting attention $r = 1/t(\underline{r})$ to a high type is higher than her expected utility from devoting the same amount of attention to a low 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 types.

C.5 Proposition 4

Claim C5.1 If the *DM* assimilates one cue, she continues to assimilate cues until she identifies the relevant topic.

Proof of Claim C5.1 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 even that the first topic is irrelevant. Repeating this argument yields that she, if she assimilates one cue, continues to assimilate cues until she finds the relevant topic.

Claim C5.2 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 C5.2 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

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

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 C5.3 The expert either launches only one topic, or at least k^{**} topics. If q_S is small, he launches at least k^{**} topics.

Proof of Claim C5.3 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 C5.4 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 C5.4 When $q_S \rightarrow \infty$, the number of topics that the expert launches satisfies $k \rightarrow \infty$. The probability that the DM opts out satisfies $\lim_{k \rightarrow \infty} \left[\frac{t}{k}p(0, \frac{1}{t})(1 - \alpha) \right] = 0$. 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.