October 11, 2004

Disclosure to an Audience with Limited Attention

David Hirshleifer*

Sonya Seongyeon Lim**

Siew Hong Teoh*

In our model, informed players decide whether or not to disclose, and observers allocate attention among disclosed signals, and toward reasoning through the implications of a failure to disclose. In equilibrium disclosure is incomplete, and observers are unrealistically optimistic. Nevertheless, regulation requiring greater disclosure can reduce observers' belief accuracies and welfare. A stronger tendency to neglect disclosed signals increases disclosure, whereas a stronger tendency to neglect failures to disclose reduces disclosure. Observer beliefs are influenced by the salience of disclosed signals, and disclosure in one arena can crowd out disclosure in other fundamentally unrelated arenas.

*Fisher College of Business, The Ohio State University, Columbus, OH 43210-1144;
Hirshleifer: hirshleifer_2@cob.osu.edu, 614-292-5174,
http://www.cob.ohio-state.edu/fin/faculty/hirshleifer/; Teoh: teoh_2@cob.osu.edu, 614-292-6547
**Department of Finance, DePaul University, Chicago, IL 60604, slim1@depaul.edu, 312-362-8825

We thank Dilip Abreu, Patrick Bolton, Paul Fischer, Jack Hirshleifer, Boyan Jovanovic, John Kagel, Dan Levin, Barbara Mellers, our discussants Amrita Nain and Reed Smith, Wei Xiong, and participants at seminars at Duke University, New York University, Ohio State University, Pennsylvania State University, Princeton University, the University of Michigan Summer Finance Conference at Estes Park, Colorado, and at the 14th Financial Economics and Accounting conference at Bloomington, Indiana for very helpful comments.

Disclosure to an Audience with Limited Attention

In our model, informed players decide whether or not to disclose, and observers allocate attention among disclosed signals, and toward reasoning through the implications of a failure to disclose. In equilibrium disclosure is incomplete, and observers are unrealistically optimistic. Nevertheless, regulation requiring greater disclosure can reduce observers' belief accuracies and welfare. A stronger tendency to neglect disclosed signals increases disclosure, whereas a stronger tendency to neglect failures to disclose reduces disclosure. Observer beliefs are influenced by the salience of disclosed signals, and disclosure in one arena can crowd out disclosure in other fundamentally unrelated arenas.

> So now we turn to the task of determining how to get more transparency – true transparency and not just more data with the unintended consequence of investor overload and the unnecessary reporting burden on companies.

—SEC Commissioner Cynthia Glassman, quoted by Paredes (2004), pp. 25-26.

1 Introduction

In the classic models of voluntary disclosure of verifiable information, observers exhibit *extreme skepticism* about those who do not reveal what they know (see Grossman (1981) and Milgrom (1981)). This skepticism is the rational response of observers to the incentive of a party with adverse information to withhold it. In practice, observers do tend to be skeptical of those who fail to disclose. However, the further implication of these models that there will be full disclosure is in practice often violated. Advertisers, con artists, firms, and politicians often do not disclose information adverse to their product or case.

Several extensions to the basic theory allow for the withholding of information.[1] Disclosure costs provide an innocent reason for non-disclosure, i.e., a reason other than the possession of an adverse signal. So disclosure costs make observers somewhat less skeptical of non-disclosure.[2] Thus, an informed player with a sufficiently favorable signal discloses, whereas if his signal is below some cutoff he withholds it. However, in these models full disclosure is still approached as disclosure costs become small.[3]

Interesting evidence about the degree of observer skepticism is provided by the market for salad dressing. Prior to the The Nutrition Labeling and Education Act of 1994, there was some voluntary provision of label information about fat content, mostly of low-fat brands. This legislation mandated quantitative disclosure of fat content. Subsequently, fattier dressings lost market share (Mathios (2000)). If consumers had been

[1] These include the models of Jovanovic (1982), Verrecchia (1983), Fishman and Hagerty (1989), Darrough and Stoughton (1990), and Teoh (1997).

[2] The rational skepticism of observers can even, in more complex signalling settings, pressure informed parties to reveal adverse information signals and withhold favorable ones (Teoh and Hwang (1991), Feltovich, Harbaugh, and To (2002)).

[3] Alternatively, Fishman and Hagerty (2003) examine a setting in which informed firms set prices and decide whether or not to disclose a signal about the quality of its product, and a subset of customers possess information complementary to the disclosed signal. This leads to differential updating by customers based upon the disclosed signal, and the possibility that in equilibrium firms do not disclose.

highly attentive and appropriately skeptical, without regulation they would have already inferred a high fat content among non-disclosing products.

There are several other strands of evidence that more directly cast doubt on the assumption underlying most existing models of disclosure that observers are fully attentive to publicly available information. For example, securities prices sometimes react strongly to irrelevant information (such as news about an unrelated firm whose abbreviation can be confused with the company's ticker symbol; see Rashes (2001)), and to the salient republication of information that is already publicly available (see Ho and Michaely (1988) and Huberman and Regev (2001)).

In this paper, observers are insufficiently skeptical because of limited attention and cognitive processing power. Limited attention sometimes causes observers to fail to take into account the implications of an absence of a signal— that an informed player has deliberately withheld relevant information. This credulity weakens the pressure on informed players to disclose. As a result, even when there are no exogenous costs of disclosure and the disclosing player surely is informed, in equilibrium disclosure is incomplete. On the other hand, limited attention in our model also sometimes causes a failure to process *disclosed* signals. This induces *greater* disclosure, by reducing the reputational penalty to a low type player of disclosing.

Our assumption of limited attention is intended to capture two stylized facts. First is the obvious fact that human information processing power is limited, which follows from the physical and design constraints of the human brain. A large literature in psychology studies limited attention, as discussed in Subsection 2.1.

The second is that people in certain contexts seem to be less skeptical about the incentives for strategic behavior of interested parties than rationality would seem to require. A body of evidence discussed in Subsection 2.2 suggests that the limited attention and credulity of investors about the motives of firms, analysts, and brokers potentially explains several general patterns in investor trading and capital market prices (these issues are also discussed in the review of Daniel, Hirshleifer, and Teoh (2002)). In our model such credulity is a natural consequence of limited attention. More broadly, we argue that limited attention has important effects on exchange between informed and uninformed parties in a range of settings, such as securities markets, consumer product markets, and non-market social interactions.

In our model, limited attention takes a simple form. Owing to cognitive resources, an individual sometimes fails to update based upon observable events, and especially based upon non-events such as the failure of an informed party to disclose some relevant information he possesses. Drawing a correct inference from non-disclosure requires both focusing attention on this non-event, and paying enough attention to the disclosure game to reason out its strategic implications. In our approach, an individual sometimes does not take these steps, and therefore fails to update his prior belief at all.

The focus of our analysis is on the equilibrium behavior of an informed player or players when the audience they face is subject to limited attention. We begin with a basic model with a single arena of possible disclosure. The informed player understands that exogenous fraction of his audience ignore disclosed signals—*cue neglect*; and ignore the implications of non-disclosure—*analytic failure*. As a result of analytic failure, in equilibrium there is a pool of non-disclosing types even though the cost of disclosure is zero. In contrast, cue neglect *encourages* disclosure, because the reputational blow received by the marginal type from disclosing is reduced.

The overall outcome is intuitive: owing to limited attention, in equilibrium there is only partial disclosure, and on average there is also excessive optimism about the quality of the informed player. However, this finding does not derive from the raw fact of limited attention, but from a tendency for observers to attend more fully to disclosed signals than to a failure to disclose. We further explore the effects of government imposed disclosure regulation, and of variation in observer attention, on equilibrium levels of disclosure, and on the precision and bias of average observer perceptions.

We extend the basic model to a setting in which individuals can choose in advance how carefully to attend to disclosed signals versus attending to a failure to disclose in an arena. The main insights of the basic setting extend to a setting with endogenous allocation of attention. In addition, the endogenous attention model allows us to examine how the degree of substitutability of attentional resources between attention to disclosed signals versus attention directed toward non-disclosure affects disclosure decisions.

Some key features of research in the psychology of attention are that individuals attend to only a limited set of cues, that there is competition between different environmental cues for attention, and that more salient or vivid cues (as defined by certain cue characteristics) capture greater attention (see subsection 2.1). We therefore examine a setting with two arenas of disclosure in which different informed players can compete for, or try to hide from, the attention of observers. Competing arenas of attention lead to effects which we call *cue competition* and *analytic interference*. Cue competition is the tendency for observation of a disclosure in one arena to distract observers' attention from disclosure in the other arena. For example, the announcement of an acquisition may distract investors from the fact that a firm has just missed an earnings forecast; for anecdotal evidence, see *Wall Street Journal*, "Heard on the Street," 12/18/98, p. C1).

Analytic interference is the tendency for disclosure in one arena to distract observers from taking into account appropriately the information implicit in the fact of non-disclosure in the other arena. For example, the announcement of large earnings surprises by firms in one industry may distract investor attention from a delay in the issuance of an earnings forecast by a firm in another industry.

We examine the implications of these effects for several issues: (1) Does regulation requiring disclosure in one arena cause the informed player in the other arena to disclose less often ('crowding out')? (2) Does requiring full disclosure in both arenas increases the accuracy of perceptions, and social welfare? (3) Is there cross-arena contagion of news announcements on observer perceptions (or market prices) in other fundamentally unrelated arenas? Finally, (4) What is the effect of disclosure regulation on observer welfare in different arenas?

Although our application of limited attention in this paper is to the theory of optimal disclosure, the simple modelling approach we provide is readily applicable to other problems in information economics. We suggest some further directions where the approach can be taken in the conclusion.

Previous work on limited attention and economic decisions has focused mainly on the imperfect rationality of managers or other organizational decisionmakers (see, e.g., the early discussion of March and Simon (1958)). As Simon (1976) remarks, "...the scarce resource is not information; it is processing capacity to attend to information." Several papers have analyzed the allocation of managerial attention across activities.[4] Our approach differs in focusing on a general audience of observers. In addition, we describe how to interpret our model assumptions in the context of a security market setting. Thus, our approach lends itself to the study of how limited attention affects the pricing of assets.

There are now a few studies which examine limited attention on the part of a general audience of observers. Gabaix and Laibson (2002) examine the implications of delayed information processing for the equity premium puzzle. Peng and Xiong (2002) and Peng (2004) examine the asset pricing implications of investors' need to allocate attention across different securities. Gabaix, Laibson, Moloche, and Weinberg (2003) and Gabaix and Laibson (2004) model the allocation of scarce information processing resources and verify the effects of attentional constraints experimentally. Our analysis differs from

[4] See Radner (1975), Radner and Rothschild (1975), Gifford (1992a, 1992b), and Gifford and Wilson (1995).

these studies in examining the two-sided problem of attention allocation by observers and manipulation of inattention by an informed party.

Hirshleifer and Teoh (2003) examine the consequences of limited investor attention for financial reporting. Their analysis takes as given that all relevant information is publicly available (either through disclosure or through spontaneous revelation). They focus on the effects of additionally reporting information as part of earnings in a firm's financial statements. In contrast, our analysis focuses on the decision to disclose information which otherwise will not be publicly available.

Perhaps the most closely related paper to this one is that of Milgrom and Roberts (1986). They show that the extreme skepticism results of past literature extend to a setting in which the informed player can disclose a set to which his signal belongs (rather than the precise value of the signal)—if observers are rational, there is full disclosure. However, if there are unsophisticated observers who are insufficiently skeptical, disclosure can be incomplete.

A key difference in our approach from that of Milgrom and Roberts is that we analyze a specific source of unsophisticated behavior, limited attention. Thus, we model not just incomplete skepticism about nondisclosure, but also failure to incorporate disclosed signals.[5] Furthermore, limited attention leads to analysis of how disclosure carries over between different informational arenas (as with the 'crowding out' effect), and how regulation in one arena affects disclosure, beliefs, and welfare in another. Thus, a distinctive aspect of our approach is our analysis of competitive effects wherein a salient disclosure attracts attention away from another disclosure; and of interference between attention to a disclosed signal in one arena and to the implications of a failure to disclose in another.

Some readers may question whether limited attention affects market prices. What is hard to contest is that both the public comments of policymakers and actual regulations reflect concerns about protecting investors with limited attention and processing power, as reflected in the head quote of this article.[6] For example, there are rules specifying not

[5]The latter effect influences the nature of the results. For example, where in Milgrom and Roberts full disclosure in their basic model requires that the observer be smart enough to draw extreme skeptical inferences, in our setting there is full disclosure even when some observers are credulous, if inattention to disclosed signals is sufficiently strong.

[6]Similarly, during her tenure as acting chair of the Securities and Exchange Commission, Laura Unger commented upon Regulation FD (Fair Disclosure): "As a Commissioner of a disclosure-based agency, I believe that more information is generally better. But is that always the case? ... [W]hat if the proposals are adopted and result in significantly greater amounts of information coming out in the form of press release? Do we need to be concerned about potential 'information overload'? ... [W]e have to remember that information can only empower investors if they understand it and can effectively apply it." (As quoted by Paredes (2004), p. 26.)

just that certain information items be revealed in a firm's financial statement, but *where* on the financial statement these items must be placed (as with rules on the reporting of comprehensive income; see Hirst and Hopkins (1998)). Furthermore, a bitter continuing fight among regulators (the Securities and Exchange Commission, the Financial Accounting Statements Board), legislators, and firms concerns whether employee stock option compensation should be disclosed in a footnote, or should be integrated as part of reporting earnings (see, e.g., Mayer (2002) and Hof (2004)). Thus, at a minimum it seems useful to assess rigorously the implications of a view that forms part of the basis for existing policy.

2 Motivating Evidence

2.1 Psychological Findings about Limited Attention

Limited attention is a necessary consequence of the vast amount of information available in the environment, and of limits to information processing power. In the face of cognitive constraints, attention must be selective and requires effort (willful substitution of cognitive resources from other tasks); see, e.g., Kahneman (1973). Several well-known decision biases are probably closely related to limits to attention, such as the phenomenon of narrow framing (as reviewed in Read, Loewenstein, and Rabin (1999)), which involves analyzing problems in too isolated a fashion.

Attention is required both for the *encoding* of environmental stimuli (such as a corporate information disclosure), and the *processing* of ideas in conscious thought (as in the analysis of a corporate disclosure or of a failure of a company to disclose). As discussed in Fiske (1995), the encoding process involves taking external information and representing it internally in a way that enables its use. Conscious thought involves a focus on particular ideas or memories to the exclusion of others. For example, a sharp focus on understanding the implications for a firm of a disclosure by that firm may limit an individual's ability to study another firm at that time.

Some stimuli tend to be perceived and encoded more easily or retrievably than others. The *salience* of a stimulus is its 'prominence,' tendency to 'stand out', or its degree of contrast with other stimuli in the environment. For example, an unusually large earnings surprise is highly salient for investors. The effects of salience are "robust and wide-ranging" (Fiske and Taylor (1991), ch.7), with influence on judgments about causality, importance of the stimulus, and how extreme it is. We reflect salience in our model as influencing the probability that a signal will be attended to, and the probability that an individual will analyze correctly the implications of a failure to disclose.

Reasonably enough, a stimulus is also more salient if it is goal-related; e.g., an individual in a group becomes more salient if you learn that she is to be your new boss. However, attention to stimuli can be misdirected in many ways, and this affects judgments. Seemingly trivial manipulations of the salience of stimuli affect judgments substantially (see, e.g., Taylor and Fiske (1978) Sect. IV). Attention is also drawn to vivid stimuli.[7] In contrast, people tend to underweight abstract, statistical, and base-rate information (see, e.g., Kahneman and Tversky (1973) and Nisbett and Ross (1980)). In view of these findings, in our model we do not assume that the amount of attention that observers direct toward a signal corresponds perfectly with its economic importance. The occurrence of an event is more salient than non-occurrence—the 'dog that didn't bark.' There is indeed evidence that individuals tend to be more influenced by the information reflected in the occurrence of an event than the non-occurrence (see, e.g., Newman, Wolff, and Hearst (1980), and Nisbett and Ross (1980)).

These considerations suggest that in a business setting, disclosures by firms that are in the news a lot (larger firms or firms in 'fashionable' sectors), or are 'proximate' and affect-linked for observers who consume the firms' products (such as entertainment, sports, or automobile firms) may be particularly salient. Based on vividness, we would also expect more attention to simple disclosures than to those that are hard to process. We also expect disclosed information to be processed more readily than the information implicit in the fact of non-disclosure.

Paying attention to one thing leaves less attention available for other things. Owing to information overload, attention must be allocated selectively (Pashler (1998)), Riley and Roitblat (1978)). A literature in psychology has examined how subjects learn by observation over time to predict a variable that is stochastically related to multiple cues (see, e.g., Baker, Mercier, Valleettourangeau, Frank, and Pan (1993), Busemeyer, Myung, and McDaniel (1993) and Kruschke and Johansen (1999)). A consistent finding is that animals and people do not achieve correct understanding of the correlation structure. Instead, *cue competition* occurs: salient cues weaken the effects of less salient ones, and the presence of irrelevant cues causes subjects to use relevant cues and base rates (unconditional frequencies) less. The presence of multiple cues also causes people

[7]Vividness is greatest for concrete descriptions and scenarios, personal stories about individual experiences, information that falls into an easily summarized pattern, stimuli that trigger emotional responses, or which are more 'proximate in a sensory, temporal or spatial way' (Nisbett and Ross (1980), p. 45).

to make analytical errors, such as 'learning' over time to use irrelevant cues.

2.2 Evidence of Limited Attention and Credulity in Markets

Casual observation suggests that observers have limited attention, and are often too credulous about the strategic incentives of their information sources. Rather than focusing on detailed, careful analysis of issues, politicians and political pressure groups invest heavily in 'sound bites' or 'photo ops' designed to underscore a simple, vivid message. Relatedly, the notion of 'rational ignorance' of voters is consistent with limits to attention and processing power. Many product advertisements are designed to engage viewers' attention and emotions in support of a salient catchphrase, rather than to present a logical or evidentiary case in support of claims about quality and price.[8] Despite the evident possibility of interested motives, con artists recurrently seduce the foolish with get-rich-quick and other scams. So at a minimum there is an extreme tail of credulous individuals.

A strong indication of limited investor attention is provided by the evidence discussed in the introduction of stock price reactions to the republication of public information. There is other evidence that capital markets have a delayed response to public information arrival.[9] Furthermore, there is also evidence that market prices do not discount adequately for incentives of firms to act strategically to take advantage of investors. For example, there is evidence that the market fails to discount sufficiently for the incentive of firms to manage earnings (see, e.g., Sloan (1996), Teoh, Welch, and Wong (1998)), and for the incentive of firms with private information to sell shares when these shares are overvalued by the market (see Loughran and Ritter (1995)) and to buy shares when shares are undervalued (see Ikenberry, Lakonishok, and Vermaelen (1995)).

With regard to disclosure decisions, an interesting example is provided by pharmaceutical companies, which publicize hot blockbuster drug candidates but often remain silent for long periods about those that do not pan out.[10] More generally, firms do on

[8] Hanson and Kysar (1999) review evidence from the consumer marketing and consumer psychology literatures, which, they maintain, indicates that sellers successfully manipulate consumer perceptions of their products.

[9] There is a delayed market response to earnings surprises (Bernard and Thomas (1990)). There is evidence that market prices do not immediately reflect long-term information implicit in demographic data for future industry product demand (DellaVigna and Pollet (2003)), and that closed-end country fund prices react more strongly to changes in the values of their holdings when news about the country appears on the front page of *The New York Times* (Klibanoff, Lamont, and Wizman (1998)).

[10] In its 2001 Annual Report, Pfizer Inc. stated that a new anti-depressant, CP-122,721 "offers strong efficacy with fewer side effects." Over time analysts inferred that the drug may have been cancelled, but

occasion collapse when concealed adverse information comes to light.

3 The Economic Setting of the Basic Model

The Informed Player

The informed player observes a signal θ on the interval $[\underline{\theta}, \overline{\theta}]$. He decides either to disclose or withhold his signal; if he discloses he must be truthful. We follow the convention that a player who is indifferent always discloses.

Uninformed Observers

There is a continuum of uninformed observers. Limited attention has two effects. First, fraction α^W of the observers are rationally skeptical about the motives of a non-disclosing ('W' for Withholding) informed player, while the remaining fraction $1 - \alpha^W$ are inattentive, a phenenomon we call analytic failure. An individual who is inattentive in this fashion does not update his beliefs from his prior about the informed party's signal. This failure to update may occur because he simply does no cognitive processing. Alternatively, he may note the fact that information was withheld, but fail to take the further cognitive step of attributing this withholding to the strategic incentives of the informed party.

For most people, it is no great conceptual leap to recognize the possibility of strategic behavior by an informed party. When paying attention, people are often quick to recognize such possibilities. However, owing to limited attention, even intelligent people often neglect fairly obvious points. Time and cognitive resources are limited, there are many arenas requiring attention, and the universe of possible signals and considerations to attend to is large.[11]

The second effect of limited attention in the model is that a fraction α^D attend to information that is disclosed; the remaining fraction $1-\alpha^D$ of observers fail to attend to the disclosure, a phenomenon we call cue neglect. We assume that $\alpha^D \geq \alpha^W$, based on the notion that disclosure is salient, and therefore calls attention to itself more strongly than a failure to disclose. This is consistent with psychological evidence discussed in

this was not confirmed until 2004 upon queries from the news media. As Hensley and Landers (2004) put it, "The case is hardly exceptional. It is common practice among large pharmaceutical companies to keep silent when hot prospects in midstage testing stumble– a frequent occurrence."

[11] Broadly supportive of this argument (though not specifically a test of it) is evidence that people tend to underweight the probabilities of event contingencies that are not explicitly available for consideration. For example, people tend to understimate the probability of 'other causes' in a list of possible causes of an event (Fischoff, Slovic, and Lichtenstein (1978)).

Section 2.1.

The shared prior belief of observers about the informed player's type has density $f(\theta)$ and distribution function $F(\theta)$. The public information set ϕ is equal either to (D, θ) (knowledge that information was disclosed, and that the revealed value was θ) or else to W (knowledge that information was not disclosed).

The average population belief about the type of the informed player is the average of the credulous/inattentive and the rational beliefs,

$$\begin{aligned} \hat{\theta}^D &\equiv (1-\alpha^D)E[\theta] + \alpha^D\theta \\ \hat{\theta}^W &\equiv (1-\alpha^W)E[\theta] + \alpha^W\hat{\theta}^\rho(W), \end{aligned} \tag{1}$$

where a hat denotes an average observer perception, and a ρ superscript indicates an attentive belief.[12]

The informed player's objective in deciding whether or not to disclose is to achieve the highest possible average perception among observers, the maximum of $\hat{\theta}^D$ and $\hat{\theta}^W$ above. For example, in a corporate disclosure context, this would amount to maximizing the current stock price. As has been found in several models of market equilibrium when some investors have imperfectly rational beliefs and others have rational beliefs, equilibrium stock prices reflect a weighted average of the beliefs of both groups of traders (see, e.g., Daniel, Hirshleifer, and Subrahmanyam (2001)), where the weights reflect the relative numbers, risk tolerances, and perceived risks of individuals with different beliefs.

4 Equilibrium in the Basic Model

4.1 Characterizing the Equilibrium

As is standard in several disclosure models, the behavior of the information recipients can be viewed in a very simple way. We assign the informed player the objective of inducing favorable average beliefs on the part of observers.[13] We propose a threshold

[12] Equation (1) has two possible interpretations. One is that observers are ex ante identical, but each has a probability of being attentive towards different information signals available in the environment and to the opportunities of the informed player to engage in strategic disclosure behavior. The other is that some non-stochastic fractions of individuals are by nature attentive or inattentive toward environmental information and toward the strategic incentives of others.

[13] This can be viewed as a reduced form of a setting in which observers take actions based upon their beliefs that affect the informed player. For example, in a corporate disclosure setting, investors would use incorrect beliefs in their security trading decisions. The resulting stock price would be of concern to the informed player (a corporate manager).

equilibrium in which an informed player discloses if and only if his type $\theta \geq \theta^*$, where θ^* is a critical signal value.

If the firm does not disclose, then the average expectation among the audience is a weighted average of the inattentive expectation $E[\theta]$, and the rational expectation $\hat{\theta}^{\rho}(W) = E[\theta|\theta < \theta^*]$. So the average perceptions given that the informed player withholds, or that he discloses his type θ, are

$$\hat{\theta}^D = (1 - \alpha^D)E[\theta] + \alpha^D\theta \tag{2}$$
$$\hat{\theta}^W = (1 - \alpha^W)E[\theta] + \alpha^W E[\theta|\theta < \theta^*]. \tag{3}$$

When an above-average type discloses, limited attention detracts from his reputation by (2), whereas limited attention enhances the reputation of a disclosing below-average type.

The equilibrium threshold value θ^* makes the informed player just willing to disclose,

$$(1 - \alpha^D)E[\theta] + \alpha^D\theta^* = (1 - \alpha^W)E[\theta] + \alpha^W E[\theta|\theta < \theta^*], \quad \text{or} \tag{4}$$
$$\theta^* = \gamma E[\theta] + (1 - \gamma)E[\theta|\theta < \theta^*], \tag{5}$$

where

$$\gamma \equiv \frac{\alpha^D - \alpha^W}{\alpha^D}. \tag{6}$$

The parameter γ measures the excess attention paid to an arena when a signal is disclosed rather than withheld. The possible equilibria are as follows.

Proposition 1 *For all parameter values, an equilibrium exists. If:*

1. $\alpha^W \geq \alpha^D > 0$*, then the unique equilibrium entails full disclosure;*
2. $0 < \alpha^W < \alpha^D$*, then in equilibrium there exists a threshold value* θ^**,* $\underline{\theta} < \theta^* < E[\theta]$*, such that the informed player discloses if his signal* $\theta \geq \theta^*$*, and withholds if* $\theta < \theta^*$*.*

To prove this, we will establish that the gain to an informed player of disclosing is a monotonic increasing function of θ for any given inference by attentive observers about the implications of non-disclosure. By (3) and (2), the difference

$$\hat{\theta}^D - \hat{\theta}^W = \alpha^D\theta + (\alpha^W - \alpha^D)E[\theta] - \alpha^W E^{\rho}[\theta|W] \tag{7}$$

is indeed monotonic in θ. Thus, there are up to three possible types of equilibrium: (i) All types disclose; (ii) No types disclose; and (iii) A player discloses if and only if his type equals or exceeds a critical value θ^*, $\underline{\theta} < \theta^* \leq \overline{\theta}$. In a proposed equilibrium with no disclosure (ii), the perception of a type that withheld would be $E[\theta]$, so any type $\theta > E[\theta]$ would prefer to disclose. This breaks the proposed equilibrium, so only (i) and (iii) are viable equilibrium candidates.

If $\alpha^W = \alpha^D$, then $\gamma = 0$, and there is full disclosure (i), because equation (5) can only be satisfied by $\theta^* = \underline{\theta}$. If $\alpha^W > \alpha^D$, then $\gamma < 0$, and there is no θ^* satisfying (5); the informed player always prefers to disclose. It remains to be shown that if $\alpha^W < \alpha^D$, full disclosure (i) is not an equilibrium, so that only possibility (iii) remains, and that equilibrium exists. The proof is in the Appendix. Intuitively, when $\alpha^W < \alpha^D$, the expected reputational penalty on a low type for failing to disclose is so small that such a type strictly prefers to withhold its signal. Finally, the critical value $\theta^* < E[\theta]$, because an above-average type would always prefer to disclose in the hope of being attended to, rather than being viewed as being below the threshold (and therefore, on average below $E[\theta]$).

This threshold equilibrium is analogous to those described in the models of Jovanovic (1982) and Verrecchia (1983). In their models, threshold behavior derives from a transaction cost of disclosure. Here, possible non-disclosure by low types results not from a disclosure cost, but from limited attention by observers to non-disclosure.

As discussed by DellaVigna and Pollet (2004), the cost of attending to a firm's earnings disclosure is likely to increase at the onset of weekends. Consistent with reduced attention to Friday disclosures, they find that the stock price sensitivity to earnings news is weaker to Friday disclosures than to Monday through Thursday disclosures, and that there is a catchup reaction to Friday earnings announcements which occurs over a period of weeks. They further find that firms are much more likely to disclose bad news on Fridays than on Mondays through Thursdays.

This evidence suggests that on non-Fridays firms disclose good news signals and withhold bad news for disclosure on Friday. Since earnings must be disclosed sometime, part of the attention toward the withholding of a signal on a given day is the attention paid to it when it is disclosed on a later day. If we interpret the lower attention paid to a non-Friday signal that is withheld until Friday as a lower α^W on the non-Friday, then for non-Fridays the premise of threshold equilibrium Part 2 of Proposition 1 that $\alpha^W < \alpha^D$ obtains, so that bad news is withheld. On Friday, in contrast, a signal that is withheld to be disclosed another day of the week will receive *more* attention when it is eventually revealed. If we interpret this greater attention to a signal withheld on Friday as a higher α^W , then on Friday α^W can exceed α^D, which encourages immediate disclosure of bad news on Friday. Thus, the evidence is consistent with Proposition 1.

Similarly, there is evidence that good news is disclosed early in the trading day, and bad news is deferred to later in the trading day (Patell and Wolfson (1982)). If information processing takes time, and if there is less attention after trading hours than during trading hours, then news disclosed late in the day receives less attention than news received early in the day. Consistent with the premise that there is less attention to news arriving after trading hours, Francis, Pagach, and Stephan (1992) document that after-hours earnings announcements are gradually impounded into price in the days after the disclosure. Thus, the non-disclosure of adverse earnings news early in the day is consistent with the threshold equilibrium of Proposition 1.

4.2 Comparative Statics on the Amount of Disclosure

Attention by observers to the withholding of information, α^W, and attention to disclosure, α^D, have opposing effects on the incentive of the informed player to disclose. Attention to a failure to disclose increases skepticism toward the informed player who withholds, encouraging disclosure. In contrast, attention to disclosure discourages disclosure by the marginal type. Since $\theta^* < E[\theta]$ (Proposition 1 Part 2), the marginal type is reevaluated adversely when observers attend to his disclosure.

Intuitively, the threshold value θ^* should decrease with α^W and increase with γ; less attention to withholding should accommodate more non-disclosure. Introducing some inattentiveness toward withholding creates a pool of non-disclosing types, and as $\alpha^W \to 0$, the pool of non-disclosing types eventually includes all below-average types (so $\theta^* = E[\theta]$).

To understand the effect of varying α^D, consider the critical type θ^*. Higher α^D increases the fraction of observers who, when the informed player discloses, perceive his type as $\theta^* < E[\theta]$ instead of the prior $E[\theta]$. This discourages disclosure, implying higher θ^*. This reasoning is consistent with (8).

To derive these results formally, note that by (6),

$$\frac{d\gamma}{d\alpha^W} < 0, \qquad \frac{d\gamma}{d\alpha^D} > 0. \tag{8}$$

Applying (3) and (2), let

$$G(t, \gamma) \equiv \frac{\hat{\theta}^D - \hat{\theta}^W}{\alpha^D} = t - \gamma E[\theta] - (1 - \gamma) E[\theta | \theta < t], \tag{9}$$

where $\hat{\theta}^D$ and $\hat{\theta}^W$ are the inferences if observers believe that the threshold is t. An equilibrium threshold θ^* satisfies $G(\theta^*, \gamma) = 0$. For a stable equilibrium, $G_1(\theta^*, \gamma) > 0$, so that a marginal increase in the perceived threshold encourages disclosure by the marginal type. Under the market perceptions associated with such a marginal increase, the firm then prefers to disclose at a critical threshold below the increased threshold. Since $G(\underline{\theta}, \gamma) < 0$ for a given $\gamma \in [0, 1]$ and $G(\overline{\theta}, \gamma) > 0$, there exists at least one stable equilibrium in the interval $[\underline{\theta}, \overline{\theta}]$.

To derive comparative statics of θ^* with respect to γ in the neighborhood of a stable equilibrium, we differentiate both sides of $G(\theta^*(\gamma), \gamma) \equiv 0$ with respect to γ:

$$\begin{aligned} 0 &= G_1(\theta^*, \gamma)\frac{d\theta^*}{d\gamma} + G_2(\theta^*, \gamma) \\ &= G_1(\theta^*, \gamma)\frac{d\theta^*}{d\gamma} - E[\theta] + E[\theta|\theta < \theta^*], \quad \text{so} \\ \frac{d\theta^*}{d\gamma} &= \frac{E[\theta] - E[\theta|\theta < \theta^*]}{G_1(\theta^*, \gamma)} > 0. \end{aligned} \tag{10}$$

The last inequality holds for stable equilibria ($G_1(\theta^*, \gamma) > 0$) since $E[\theta] > E[\theta|\theta < \theta^*]$.

Proposition 2 *Under the assumptions of the basic model, in stable equilibria:*

1. *The amount of disclosure increases with the fraction of observers who are attentive about the withholding of information α^W;*

2. *The amount of disclosure decreases with the fraction of observers who attend to disclosure α^D.*

4.3 Accuracy of Observer Perceptions

We now examine how attention affects the accuracy (bias and mean squared error) of observers' average perception.

4.3.1 Optimism

Observers on average tend to be optimistic about the quality of the informed party ($E[\hat{\theta}] > E[\theta]$). On the one hand, for a given threshold θ^*, credulity about non-disclosure increases the average perception— inattentive individuals perceive the type of a non-disclosing informed player as on average $E[\theta]$ instead of $\theta < \theta^*$. On the other hand, inattention to disclosure tends to decrease average perception— inattentive individuals perceive the type of a disclosing informed player as on average $E[\theta]$ instead of the disclosed $\theta > \theta^*$, which on average must be greater than $E[\theta]$. However, so long as $\alpha^W < \alpha^D$, the first effect dominates (see the Appendix). We therefore have:

Proposition 3 *If $\alpha^D > \alpha^W$, then observers are on average overoptimistic about the informed signal ($E[\hat{\theta}] > E[\theta]$). If $\alpha^D = \alpha^W$, then observers on average correctly assess the quality of the informed player.*

Although there is an average tendency toward optimism, owing to cue neglect average investor perceptions are too pessimistic when θ is sufficiently high.

Similar points may apply more broadly in settings where an informed player has some discretion in what he tells an observer with limited attention. Psychologists have found that individuals on average tend to exhibit unrealistic optimism about the likelihood of experiencing favorable personal outcomes. Adam Smith (1776), in regard to "the greater part of mankind," referred to "Their absurd presumption in their own good fortune..." Psychological research has also confirmed that individuals are subject to unrealistic optimism (Weinstein (1980, 1982)).

Unrealistic optimism may result in part from limited attention. Life events are substantially influenced by the strategic revelation policies of interested, informed parties. Participants in business and personal relationships often conceal their lack of commitment; as a result, all too often people are shocked when they lose their jobs or life partners 'out of the blue.' Our analysis suggests that credulity about the strategic incentives of others may be a source of unrealistic optimism.[14,15]

Differentiating the average optimism (38) with respect to α^W and α^D, it is not hard to verify that more attention to disclosed signals discourages disclosure, and thereby increases optimism; and that greater attention to non-disclosure decreases optimism.

[14]Kennedy and Dimick (1987) find that 48% of college athletes in revenue-producing sports expect to play professionally, while the actual figure is 2%. Colleges may have an incentive to allow athletes to believe they have a real shot at going professional, rather than disclosing adverse information about likelihood of success.

[15]To give a conjectural hint about how such an issue could be modelled, consider a social exchange setting in which two individuals play a repeated Prisoners' Dilemma over an infinite number of periods. In such a setting there may be a trigger strategy equilibrium enforcing cooperation (C) as opposed to defection (D). However, in round t one party I receives private information about the value of an external option that will become available in round $t+k$, where $k > 0$ is known to all. To exploit the external option he must, at time $t+k$, abandon the existing relationship, i.e., he must play D at round $t+k$ and at all later rounds. If the external option is sufficiently favorable, it pays for him to do so. I's signal provides him with superior information in round t about whether he will later abandon the current relationship. If the uninformed party U infers that I is sufficiently likely to abandon, then the trigger strategy equilibrium breaks down and U defects immediately as well. Such breakdown of cooperation is costly to I, who prefers to reap the rewards of the old relationship longer. Thus, under limited attention, I may benefit from concealing favorable external options, thereby encouraging U to be optimistic about I's commitment to the relationship.

Proposition 4 *In the neighborhood of a stable equilibrium, average optimism* $E[\hat{\theta}-\theta]$ *increases as attention to disclosure increases* ($dE[\hat{\theta}-\theta]/d\alpha^D > 0$) *and decreases with attention to non-disclosure* ($dE[\hat{\theta}-\theta]/d\alpha^W < 0$) *.*

In addition, greater uncertainty by observers about the informed player's information increases optimism; it is not hard to show that if the density of θ is horizontally stretched by a factor of K, then average optimism is also multiplied by K.

4.3.2 Mean Squared Error

The mean squared deviation of the average perception held by observers about the informed party from the actual type,

$$E[(\hat{\theta}-\theta)^2] = \int_{\underline{\theta}}^{\theta^*} (\hat{\theta}^W - \theta)^2 f(\theta) d\theta + \int_{\theta^*}^{\overline{\theta}} (\hat{\theta}^D - \theta)^2 f(\theta) d\theta, \tag{11}$$

is a measure of the inaccuracy of observer perceptions.

Proposition 5 *If there is a higher probability that observers attend to disclosed information than to the fact that information is withheld,* $\alpha^D > \alpha^W$, *then in the neighborhood of a stable equilibrium:*

1. *The mean squared error of the average observer perceptions as an estimate of the true type decreases in* α^W.

2. *The mean squared error of the average observer perceptions as an estimate of the true type can either increase or decrease as* α^D *increases.*

The proof is contained in the Appendix.

Intuitively, by Proposition 2, the more attentive observers are about non-disclosure, the more disclosure occurs. Greater disclosure makes beliefs on average more accurate, implying a lower mean squared error. It is true that perceptions are sometimes inaccurate even after disclosure so long as $\alpha^D < 1$. However, since $\alpha^D > \alpha^W$, the frequency of inaccurate perceptions is lower when the informed player discloses than when he does not disclose. Furthermore, greater α^W implies the average perception of observers about a non-disclosing type is more accurate as $E[\theta|\theta < \theta^*]$ is a better estimate of a non-disclosing type than $E[\theta]$

In contrast, the comparative statics with respect to α^D is ambiguous. The more attentive observers are to disclosure, the less disclosure occurs (Proposition 2). For reasons discussed in the preceding paragraph, less disclosure tends to increase the mean squared error of the average investor perception if $\alpha^D > \alpha^W$. However, a countervailing force is that higher α^D increases the probability that an individual incorporates the disclosed information into his beliefs.

5 Competing Attentional Demands and Salience

When individuals have limited cognitive capacity it is impossible to attend to all decision-relevant information of different forms and derived from different sources. The allocation of attention across strategic considerations or information signals will in general be biased by the salience of different aspects of the decision environment. As discussed in Section 2.1, there is extensive evidence from the psychology literature that attention is selective, and that the allocation of attention across stimuli often does not correspond very well to differences in the informativeness or usefulness of different signals. We now explore a setting in which disclosure or non-disclosure in each arena influences perceptions and disclosure behavior in the other.

In this setting there are two arenas of disclosure, $i = A$ or B. The informed player in arena i observes a signal θ_i, where $\underline{\theta}_i \leq \theta_i \leq \overline{\theta}_i$. Disclosure decisions in each arena are taken simultaneously. An observer who attends to disclosure or to non-disclosure in a given arena updates his belief rationally, whereas an observer who fails to attend to arena i holds to his prior about θ_i.

We consider a very simple form of limited attention: individuals can attend to one or the other arena, but not both. This specification captures the notion of information overload in a simple way. In more realistic settings with multiple arenas, it is likely that attention can be complementary within some groups of arenas. In attentionally related arenas, paying attention to one arena may increase attention to the others in the same group. Nevertheless, limited attention implies that an increase of attention in some arenas *must* be offset by decreased attention in some other arena. For example, news about a particular biotech company may draw investors' attention to other biotech firms while reducing attention to firms in other industries. Our two-arena approach is based upon the unavoidability of attentional substitutions. Our model captures the competitive nature of attention, as supported by the experimental studies on selective attention; our results in two-arena cases can be interpreted more broadly as involving competition between attentional sectors.

Borrowing from the literature in experimental psychology on multiple cue learning (see Section 2.1), we call the tendency for an information disclosure in one arena to distract observers from attending to a disclosure in the other arena *cue competition*. We call the tendency for an information disclosure in one arena to distract observers from inferring the reason for an action or failure to act in another arena *analytic interference*.

5.1 The Basic Model with Competing Information Sources

For each arena A and B, we will first show that an equilibrium of the sort described in the preceding section applies. Limited attention determines the fraction of the observers who are credulous with respect to each arena, α_A^W, α_A^D, α_B^W and α_B^D, endogenously.

If the individual is faced with no disclosure in either arena, we assume that he attends to one or the other arena with equal probability. If there is disclosure about one arena but not another, then the effect of the disclosure on attention to the other arena is assumed to be related to the *salience* of the information disclosure.

We allow different arenas of disclosure to have different levels of salience or vividness. For simplicity the amount of attentional interference between arenas depends only on whether disclosure occurred, not on the signal realization. The saliences of arenas A and B, denoted s_A or $s_B > 0$, help determine the probability that individuals will attend to each arena.

In the absence of any attention-grabbing events, the probability of an individual attending to A versus B would be .5. If the arena A signal is withheld and the arena B signal is disclosed, arena B is likely to capture a greater share of observer attention. To reflect the higher salience of occurrence than non-occurrence of an event, we assume that the probability that an individual attends to arena A or to B is

$$\begin{aligned} \alpha_A(W_A, D_B) &= .5(1 - s_B) \\ \alpha_B(W_A, D_B) &= .5(1 + s_B). \end{aligned} \tag{12}$$

Thus, as the salience of disclosure in arena B increases, it robs more attention from the non-disclosing arena A. If the salience of the disclosed information is 0 this effect vanishes (B gets only its 50:50 share of attention). However, as salience rises to 1 the probability of attending to A diminishes to zero. Symmetrically, we assume that

$$\begin{aligned} \alpha_A(D_A, W_B) &= .5(1 + s_A) \\ \alpha_B(D_A, W_B) &= .5(1 - s_A). \end{aligned} \tag{13}$$

If there is disclosure in both arenas, then it is assumed that the probability that an individual attends to A or to B is

$$\begin{aligned}\alpha_A(D_A, D_B) &= .5(1 + s_A - s_B) \\ \alpha_B(D_A, D_B) &= .5(1 - s_A + s_B). \end{aligned} \tag{14}$$

There is a greater tendency to attend to the more salient arena, and if the difference in salience between the two arenas is maximal ($1 - 0 = 1$), then an observer attends to the more salient disclosure with certainty.

In equilibrium, the informed player in arena i takes the strategy of the other informed player (i.e., player i''s threshold value $\theta^*_{i'}$) as given. Each informed player therefore treats the fraction of observers who will attend to disclosure, α^D_i, or to non-disclosure, α^W_i, in his arena as given. Thus, we can apply the equilibrium of the previous section to each of the arenas, to derive the threshold value in arena i, θ^*_i, as a function of the proposed critical value in the other arena i', $\theta^*_{i'}$. Given critical value θ^*_B, the probability that an observer attends to the fact that information is withheld in arena A is

$$\begin{aligned}\alpha^W_A &= \alpha_A(W_A, W_B)Pr(W_B) + \alpha_A(W_A, D_B)Pr(D_B) \\ &= .5[Pr(\theta_B < \theta^*_B)] + .5(1 - s_B)[Pr(\theta_B > \theta^*_B)] \\ &= .5[1 - s_B + s_B F(\theta^*_B)]. \end{aligned} \tag{15}$$

Similarly, for arena B,

$$\alpha^W_B = .5[1 - s_A + s_A F(\theta^*_A)]. \tag{16}$$

Given a proposed threshold in arena B, θ^*_B, the probability that a given observer attends to disclosure in arena A is

$$\begin{aligned}\alpha^D_A &= \alpha_A(D_A, W_B)Pr(W_B) + \alpha_A(D_A, D_B)Pr(D_B) \\ &= .5(1 + s_A)F_B(\theta^*_B) + .5(1 - s_B + s_A)[1 - F_B(\theta^*_B)] \\ &= .5(1 + s_A - s_B[1 - F_B(\theta^*_B)]). \end{aligned} \tag{17}$$

The probability that an observer is attentive to disclosure in B is derived similarly:

$$\alpha^D_B = .5(1 + s_B - s_A[1 - F_A(\theta^*_A)]). \tag{18}$$

From (15)-(18), $\alpha^D_i > \alpha^W_i$ when $s_i > 0$. We propose an equilibrium in which each informed player follows a threshold disclosure rule, with cutoffs θ^*_A and θ^*_B. We determine the equilibrium in each arena taking the cutoff in the other arena as given. We seek a set of self-confirming cutoff values that satisfy the basic model equilibrium conditions together with (15)-(18).

However, in general there may be multiple equilibria, possibly asymmetric. High disclosure in one arena can distract from the other, leading to lower disclosure in the other. Later we will show uniqueness by direct calculation in the case of a uniform distribution of types.

The equilibrium condition for informed player i to be just willing to disclose, as in (5), is that

$$\theta_i^* = \gamma_i E[\theta] + (1-\gamma_i) E[\theta | \theta < \theta_i^*], \quad \text{where} \tag{19}$$

$$\gamma_i \equiv \frac{\alpha_i^D - \alpha_i^W}{\alpha_i^D}, \qquad i = A \text{ or } B, \tag{20}$$

and where α_i^W, α_i^D are the probabilities that individuals attend to either the withholding of information, or the disclosure of information, in arena i.

Equations (19) and (20) describe θ_A^* in terms of α_A^W and α_A^D. But these in turn are both functions of θ_B^*. Thus, we can solve for a reaction curve $\theta_A^*(\theta_B^*)$. Similarly, we can solve for the reaction curve $\theta_B^*(\theta_A^*)$. Together these reaction curves determine equilibrium values for the two disclosure thresholds.

We consider the case of uniform distributions, $f(\theta) = 1/(\overline{\theta} - \underline{\theta})$. By (5),

$$\theta^* = \frac{\gamma\overline{\theta} + \underline{\theta}}{\gamma + 1}. \tag{21}$$

Without loss of generality set $\underline{\theta}_A = \underline{\theta}_B = 0$, and $\overline{\theta}_A = \overline{\theta}_B = 1$ (a rescaling and translation). By (6) and (21),

$$\theta_i^*(\alpha_i^W, \alpha_i^D) = \frac{\alpha_i^D - \alpha_i^W}{2\alpha_i^D - \alpha_i^W}, \qquad i = A, B. \tag{22}$$

We now solve for the attention parameters in arena A (the α's) in terms of the tendency for disclosure in arena B, as measured by the threshold in that arena, θ_B^*. From (15)-(18), this gives the reaction curve for the informed player in arena A, and a similar derivation gives the curve for arena B:

$$\theta_A^*(\theta_B^*) = \frac{s_A}{1 + 2s_A - s_B + s_B\theta_B^*} \tag{23}$$

$$\theta_B^*(\theta_A^*) = \frac{s_B}{1 + 2s_B - s_A + s_A\theta_A^*}. \tag{24}$$

Figure 1 shows reaction curves for different parameter values, and the equilibria determined by the intersections of these curves.

Insert Figure 1 Here

Combining equations (23) and (24) gives the equilibrium disclosure threshold θ_A^* as a root of the quadratic equation

$$s_A(1+2s_A-s_B){\theta_A^*}^2+[(1+2s_A-s_B)(1+2s_B-s_A)+{s_B}^2-{s_A}^2]\theta_A^*-s_A(1+2s_B-s_A)=0. \tag{25}$$

This equation has only one root between 0 and 1. Since $0<\theta_A^*<1$, by (24), θ_B^* is also between 0 and 1. We therefore have:

Proposition 6 *If the types of the informed players in two arenas are distributed uniformly on* $[0,1]$*, then there is a unique equilibrium. In this equilibrium, there is partial disclosure.*

By (23), $d\theta_A^*/d\theta_B^*<0$, so the reaction curves are downward sloping. In other words, disclosures in the two arenas are strategic substitutes. It follows that regulation that forces more disclosure in one arena than would have occurred in equilibrium crowds out disclosure in the other arena. Intuitively, reduced attention to B reduces the pressure on a marginal B player to disclose.[16]

Proposition 7 *Under the assumptions of this section, regulation that forces greater disclosure in one arena (by reducing* θ_i^* *in arena* i*) causes the informed player in the other arena to disclose less information (i.e., in arena* $\sim i$ *the disclosure threshold* $\theta_{\sim i}^*$ *increases).*

5.2 Cross-Arena Contagion of News Announcements

This subsection examines the effect of disclosure versus non-disclosure in arena A on the expected perception by observers of the signal value in a fundamentally-unrelated arena B. Applied in a stock market setting, the results we derive describe how an announcement about one stock, such as an earnings forecast or dividend announcement, affects the price of another stock.

[16] Under alternative assumptions, the two arenas could be attentional complements so that disclosure in one arena increases attention to the other arena. But there is still a crowding out effect when disclosure in one arena results in much larger increase in attention to disclosure than increase in attention to non-disclosure. Since greater attention to disclosure discourages disclosure, disclosure in one arena can crowd out disclosure in the other arena by increasing attention to disclosure.

Thus, we calculate $E[\hat{\theta}_B|D_A, \theta_A]$ and $E[\hat{\theta}_B|W_A]$, where the expectation is taken with respect to the possible outcome for θ_B. The average perception in arena B if the informed player in arena B withholds, $\hat{\theta}_B^W$, is, in analogy to equation (3),

$$\hat{\theta}_B^W = (1 - \alpha_B^W)E[\theta_B] + \alpha_B^W \int_{\underline{\theta}_B}^{\theta_B^*} \theta_B \frac{f(\theta_B)}{F(\theta_B^*)} d\theta_B. \tag{26}$$

If he discloses, then the average perception is

$$\hat{\theta}_B^D = (1 - \alpha_B^D)E[\theta_B] + \alpha_B^D \theta_B. \tag{27}$$

Taking expectations over the possible values of θ_B, we find the expected perception by observers of the informed player in arena B conditional on the behavior of the informed player in arena A. Recalling our notation $\phi = W$ or (D, θ), this is

$$\begin{aligned} E[\hat{\theta}_B|\phi_A] &= \int_{\underline{\theta}_B}^{\theta_B^*(\phi_A)} \hat{\theta}_B^W f(\theta_B) d\theta_B + \int_{\theta_B^*(\phi_A)}^{\overline{\theta}_B} \hat{\theta}_B^D f(\theta_B) d\theta_B \\ &= [\alpha_B^D(\phi_A) - \alpha_B^W(\phi_A)] \int_{\underline{\theta}_B}^{\theta_B^*(\phi_A)} \{E[\theta_B] - \theta_B\} f(\theta_B) d\theta_B + E[\theta_B]. \end{aligned} \tag{28}$$

The effect of ϕ_A on perceptions in arena B comes from ϕ_A's effect on α_B^W, α_B^D, and thereby on θ_B^*; attention in arena B depends on the disclosure choice and outcome in arena A.

Comparing $E[\hat{\theta}_B|D_A, \theta_A]$ with $E[\hat{\theta}_B|W_A]$, we obtain:

Proposition 8 *Under the assumptions of the model, disclosure in one arena causes the expected perception of the type in the other arena to increase.*

Two further empirical implications follow from this approach by specializing the analysis to a setting in which there is an exogenous probability of news arrival in one of two arenas, instead of disclosure in both arenas. As in Proposition 8, it follows that news arrival in one arena implies a positive revision of perceptions by observers in a fundamentally unrelated arena, owing to a lack of skepticism about non-disclosure. For example, a major general news event is predicted to distract attention from an unrelated non-event such as a failure of a firm to make an earnings announcements, leading on average to overvaluation. Second, the model implies that a major general news event will distract attention from an unrelated news event such as a firm's earnings announcement (as reflected in equation (14)), leading to underreaction in investor perceptions to the earnings surprise.

The two arenas in our model have no fundamental relationship (θ_A and θ_B are independent). Cross-effects here are induced by the attentional relationship between the two arenas. More generally, if there are many arenas, disclosure in any single arena may have little effect on another arena unless there is some kind of attentional linkage between the two. The attentional linkage could derive from a fundamental relationship, or could be entirely superficial (as with two firms with similar-sounding names).[17]

When there are many arenas, the cross effect between attentionally related ones can be positive. Instead of distracting, disclosure in a given arena can call further attention to a few attentionally related arenas, while distracting slightly from a large number of more distantly related arenas. Thus, in applying Proposition 8 to a stock market setting, it is best to view Arenas A and B not as individual stocks, but as entire industries or sectors. For example, if big news attracts attention to the high-tech sector, this may distract investors from attending to disclosure or non-disclosure by firms in the energy sector.

6 Welfare Effects of Disclosure Regulation

We examine here how disclosure regulation affects the accuracy of investor perceptions, the degree of optimism, and welfare. If government simply mandates disclosure and that mandate is always obeyed, in effect the disclosure threshold is set below $\underline{\theta}$. However, often a more realistic description of the legal/regulatory environment is that only partial disclosure is enforced, implying a threshold at an intermediate value between $\underline{\theta}$ and $\overline{\theta}$.

The threat of liability for the failure to disclose can encourage a firm to do so (see, e.g., Skinner (1994)), but may not enforce complete disclosure. Even when disclosure is clearly mandated, in deciding whether to disclose firms may balance the risk of liability against the costs of publicizing bad news.

Furthermore, there may sometimes be legal uncertainty as to whether disclosure of the information item is mandatory. In either case, firms will be pressured to disclose more bad news when there is a higher probability of legal liability for non-disclosure, and when the expected penalties are higher. Thus, the legal/regulatory environment involves different effects which can be viewed as adjusting the level of the threshold indirectly.

We assume that welfare is measured by the accuracy of investor perceptions. If there

[17]Rashes (2001) provides evidence of stock market pricing errors based upon investors being confused by similarities in names and ticker symbols of different stocks, a clear symptom that limited attention causes inappropriate contagion across arenas.

is only a single arena, we define welfare as the negative of the mean squared error,

$$W \equiv -E[(\hat{\theta} - \theta)^2], \tag{29}$$

where $\hat{\theta}$ is the average perception of type conditional on the disclosure decision.

When there are multiple arenas, the value of belief accuracy will generally vary depending on the arena's size and characteristics.[18] We therefore consider a social welfare function that allows for unequal weights on perception errors between arenas A and B,

$$W \equiv -\lambda E[(\hat{\theta}_A - \theta_A)^2] - (1 - \lambda) E[(\hat{\theta}_B - \theta_B)^2], \tag{30}$$

where $\hat{\theta}_A$ and $\hat{\theta}_B$ are the average perceptions of type in the two arenas conditional on the disclosure decision, and λ measures the relative importance of the two arenas.

We begin by describing the effects of regulation when there is only a single arena.

Proposition 9 *Under the assumptions of the basic model, if there is a higher probability that observers attend to the disclosed information than to the fact that information is withheld, $\alpha^D > \alpha^W$, then:*

1. *Suppose that a regulation imposes a disclosure threshold level below that implied by the indifference condition (5). Then, as the regulated threshold decreases (greater disclosure), the mean squared error of the average observer perception as an estimate of the true type decreases, and social welfare increases.*

2. *If the regulated disclosure threshold $\theta^* < E[\theta]$, then an exogenous decrease in the regulated threshold reduces optimism.*

So long as regulation encourages disclosure (perhaps by imposing a risk of liability on a non-disclosing firm), the condition $\theta^* < E[\theta]$ holds (by Proposition 1), because this condition holds even when there is no regulation.

The accuracy of observers' beliefs increases with the amount of disclosure. Consider now a regulation that imposes a level of disclosure θ^*. Intuitively, higher θ^* (a milder

[18] For example, one arena may have greater payoff variability than the other, so that the benefits of accurate information differs across arenas. The validity of investor perceptions of the paper clip industry may matter less than perceptions of the steel industry. Also, an incorrect perception of a state of the world that is not very relevant for the observer's actions may matter less than a highly action-relevant state of the world. Alternatively, the informed player's 'type' in our model could be viewed as a noisy indicator of value, where the noise variance can differ across arenas. Other things equal, the social value of accurate communication of a noisier signal is smaller than communication of a more accurate one. Also, the qualitative nature of the information disclosed is important for welfare; see, e.g., the analysis of Boot and Thakor (2001).

disclosure requirement) increases the probability that observers hold to their prior beliefs (since $\alpha^D > \alpha^W$) instead of updating appropriately. It also tends to make even fully rational inferences of type more noisy by increasing the set of types that are not revealed (see the Appendix, Proof of Proposition 9, Part 1).

Policymakers have sometimes expressed concern that investors can be overwhelmed by information with limited usefulness, which may prevent them from processing effectively other information that is more useful. For example, a Supreme Court ruling on TSC industries, Inc. v. Northway states that "Some information is of such dubious significance that insistence on its disclosure may accomplish more harm than good. ...[M]anagement's fear of exposing itself to substantial liability may cause it simply to bury the shareholders in an avalanche of trivial information – a result that is hardly conducive to informed decision making." (As quoted by Paredes (2004), p.27.) The court's concern was with disclosure of low quality information; we show here that there are related problems even with the disclosure of accurate and relevant information.

When there are multiple arenas, we can examine the effects of a change in the disclosure rule in one of the arenas, or in both simultaneously.

Proposition 10

1. *If there is disclosure in two arenas, then an exogenous increase in the disclosure threshold (implying a lower probability of disclosure) in one arena causes average beliefs to be more accurate (lower mean-squared-error) in the other arena in the neighborhood of a stable equilibrium.*

2. *An exogenous increase in the disclosure threshold in one arena reduces optimism in the other arena in the neighborhood of a stable equilibrium.*

3. *Forced disclosure in one arena can reduce welfare by discouraging disclosure in the other arena.*

4. *If there is disclosure in two arenas, then a simultaneous exogenous increase in the disclosure threshold in both arenas ($\theta^* = \theta^*_A = \theta^*_B$) can cause average beliefs in each area to become either more or less accurate, i.e., $\partial E[(\hat{\theta} - \theta)^2]/\partial\theta^*$ can be negative or positive. Thus, even though there are no costs of disclosure, forced disclosure in both arenas can reduce welfare.*

The proof of Parts 1, 2, and 3 are in the Appendix. Intuitively, there is a crowding out effect of disclosure. For example, if regulators impose higher disclosure in arena A, then this distracts observers from arena B. The lower attention to either disclosure or non-disclosure in arena B makes beliefs in arena B less accurate, which can reduce welfare, especially if arena B is more important than arena A.

This suggests an important limitation to disclosure regulation. It is impossible to regulate disclosure in all arenas. Even if high disclosure were required of all firms, there is private information in other parts of economic and social life. The effect of imposing disclosure requirements in some arenas but not others may be to redirect attention rather than to improve the accuracy of all perceptions. This suggests that a relevant input for regulatory decisions is the importance of different arenas for the decisions of observers.

To prove Part 4, we consider the case of symmetric salience and calculate the mean squared error of beliefs as a function of the common cutoff θ^*. When the θ_i's are distributed uniformly on $[0,1]$, by direct calculation the mean squared error can increase or decrease with the common cutoff θ^*. For example, when $s_A = s_B = 0.99$ and $\theta^* = 0.4$, $\partial E[(\hat{\theta}-\theta)^2]/\partial\theta^* \approx -0.021$, but when $s_A = s_B = 0.1$, the derivative is positive (≈ 0.025).

Part 3 indicates that disclosure in a less important arena may distract so much from disclosure in the important arena that welfare declines. To gain further insight about this finding, consider the symmetric case where $s_A = s_B = s$. Using the attention probabilities given in equations (15)-(18), we can differentiate welfare with respect to the disclosure threshold for A (holding constant θ_B^*); details are provided in the Appendix, proof of Part 3. In this case, $\partial W/\partial\theta_A^*$ as a function of s and λ is shown in Figure 2.

Insert Figure 2 Here

As seen in Figure 2, $\partial W/\partial\theta_A^*$ is positive when the salience s is high and the weight on arena A is low. Forced disclosure in arena A (lower θ_A^*) can decrease welfare when the importance of arena B is high (lower λ) and the salience is high. A high salience of disclosure in arena A, combined with greater disclosure in A, makes withholding more attractive in arena B. Therefore, forced disclosure in A can have a negative overall effect by discouraging disclosure in the more important arena.

In Part 4 there are countervailing effects. On the one hand, greater disclosure in one arena has a direct tendency to increase the precision of observer beliefs in that arena. On the other hand, it distracts attention from the other arena, which tends to reduce precision there. Overall, jointly forcing increased disclosure can reduce belief precisions and therefore welfare, as illustrated in Figure 3.

Letting $\theta_A^* = \theta_B^* \equiv \theta^*$, as illustrated in Figure 3, $\partial W/\partial\theta^*$ is positive for high values of s and negative for low values of s (calculations in the Appendix).

Insert Figure 3 Here

Intuitively, although forcing disclosure in both arenas directly increases the amount of information available, it also makes people less attentive to a given datum. If salience is high enough, the second effect (lower attention) can outweigh the first (more information/disclosure).

7 Optimal Allocation of Attention

Even boundedly rational observers can try hard to attend to those signals that offer high return to attention. This section generalizes the basic single-arena model of Section 3 to allow individuals to decide *ex ante* how much attention to allocate to either disclosed information or the strategic implications of non-disclosure. We first examine whether the qualitative and comparative statics implications the the basic model survive in a setting with endogenous allocation of attention. We then examine how the technology for allocating attention affects comparative statics. Specifically, we examine how the degree of complementarity or substitutability between attention to disclosed signals versus attention to non-disclosure affects individuals' decisions.

The focus of the analysis here is on the first arena A, but we include a second arena B to give individuals an opportunity cost of attending to arena A. Individuals ex ante also have a choice within arena A as to how much attention to devote to disclosed information versus the failure of information to be disclosed.

There are two stages in the model. In the first stage identical observers choose attention probabilities α_A^W, α_A^D, as defined earlier, and the probability of attending to a second independent arena B, α_B. This choice is not observable to the informed player, although in equilibrium he knows what the choice will be.

In the second stage, attention outcomes are realized, so each observer either attends or does not attend to each arena and forms beliefs accordingly. The informed player observes his private signal about arena A, and decides whether or not to disclose. At this point, the decision problem of the informed player is identical to that of the informed player in the basic model. So the disclosure threshold θ_A^* is determined as in the basic model as a function of α_A^W and α_A^D.

Also at the second stage, each observer makes a project choice based on his beliefs at that time. Thus, at the first stage the individual allocates attention so as to increase the quality of his later project choice. Let θ_A be an observer's net payoff from adopting the arena A project. We assume that if he is indifferent, he adopts the project. Thus, he adopts if and only if $\theta_A \geq 0$. For algebraic simplicity, we assume that $E[\theta_A] = 0$, so that an individual who does not attend to arena A is just willing to undertake the project. Similar results apply more generally.

Also for simplicity, we assume that there is no disclosure game in arena B. Instead, an information signal becomes public spontaneously and with certainty. We therefore assume that the B component of the observers' payoffs is equal to $\alpha_B K$, where $K > 0$ is a constant.[19]

From the analysis of the basic model, $\theta_A^* < E[\theta_A] = 0$. When an individual attends to a disclosure in arena A, he adopts the project if and only if $\theta_A > 0$. When he attends to a failure to disclose, he rejects the project since $E[\theta_A|\theta_A < \theta_A^*] < 0$. As discussed above, whenever an individual fails to attend (either to a disclosure or to a failure to disclose), he adopts. Thus, his expected payoff from his arena A project choices given his attention probabilities α_A^D and α_A^W is

$$\begin{aligned}
\Pi &= (1-\alpha_A^D)Pr(\theta_A > \theta_A^*)E[\theta_A|\theta_A > \theta_A^*] \\
&+ \alpha_A^D Pr(\theta_A > 0)E[\theta_A|\theta_A > 0] + (1-\alpha_A^W)Pr(\theta_A < \theta_A^*)E[\theta_A|\theta_A < \theta_A^*] + \alpha_A^W \cdot (0) \\
&= (1-\alpha_A^D)\int_{\theta_A*}^{\bar{\theta}_A} \theta_A f(\theta_A)d\theta_A + \alpha_A^D \int_0^{\bar{\theta}_A} \theta_A f(\theta_A)d\theta_A + (1-\alpha_A^W)\int_{\underline{\theta_A}}^{\theta_A^*} \theta_A f(\theta_A)d\theta_A \\
&= \int_{\theta_A*}^{\bar{\theta}_A} \theta_A f(\theta_A)d\theta_A - \alpha_A^D \int_{\theta_A^*}^{0} \theta_A f(\theta_A)d\theta_A + (1-\alpha_A^W)\int_{\underline{\theta}_A}^{\theta_A^*} \theta_A f(\theta_A)d\theta_A \qquad (31)
\end{aligned}$$

Adding to this the expected profit from the arena B project, an observer's overall first stage optimization problem is

$$\max_{\alpha_A^W,\alpha_A^D,\alpha_B} \int_{\theta_A*}^{\bar{\theta}_A} \theta_A f(\theta_A)d\theta_A - \alpha_A^D \int_{\theta_A^*}^{0} \theta_A f(\theta_A)d\theta_A + (1-\alpha_A^W)\int_{\underline{\theta}_A}^{\theta_A^*} \theta_A f(\theta_A)d\theta_A + \alpha_B K,$$

subject to the attention allocation constraint,

$$G(\alpha_A^W, \alpha_A^D, \alpha_B) \leq 1, \qquad (32)$$

where $G(\cdot,\cdot,\cdot)$ is weakly increasing in each of its arguments.

The function G reflects the degree to which the individual is able to substitute attention between different arenas, and between attention toward disclosure in Arena

[19]It is easy to endogenize this form by introducing an investment project related to arena B with net value $k\theta_B$, where $E[\theta_B] = 0$ and $k > 0$. Attending to B allows an observer to obtain on average $kE[\theta_B|\theta_B \geq 0]$ by investing when doing so is profitable, instead of always investing for an average payoff of 0. In this setting, arena B contributes $\alpha_B kE[\theta_B|\theta_B \geq 0]$ to the observer's expected profits, so that $K = kE[\theta_B|\theta_B \geq 0]$.

A and attention toward withholding of information in Arena A. For example, If the individual has good control over his attention, the α's may be highly substitutable. On the other hand, if vividness and salience grab people's attention without conscious volition, the different α's may be highly complementary in G, so that it is hard for the individual to shift the proportions between the α's from some natural ratio.

We consider a tractable functional form for G,

$$G(\alpha_A^W, \alpha_A^D, \alpha_B) = \left[s^W(\sigma^W \alpha_A^W)^p + s^D(\sigma^D \alpha_A^D)^p + (1 - s^W - s^D)\alpha_B\right]^{-1/p}, \qquad (33)$$

where $p > 1$ and σ^W and σ^D are exogenous parameters which measure the opportunity cost of directing attention to a particular target (the opportunity cost of directing attention to arena B is normalized to 1). s^W and s^D correspond to relative factor shares in the constant elasticity of substitution production function, where s^W, s^D, and $1 - s^W - s^D$ all range from 0 to 1. It can be easily shown that the elasticity of substitution between α_A^W and α_A^D is

$$\frac{\partial ln(\alpha_A^W/\alpha_A^D)}{\partial ln(G_{\alpha_A^W}/G_{\alpha_A^D})} = \frac{1}{p-1},$$

which is a decreasing function of p.

Proposition 11 *If individuals decide how much attention to devote to either disclosed information or the implications of non-disclosure subject to the attention transformation constraint (33), and an interior equilibrium obtains with* $0 < \alpha_A^W < \alpha_A^D < 1$, *then*

1. *The amount of disclosure increases with the opportunity cost of attending to disclosure* ($d\theta_A^*/d\sigma^D < 0$)*;*

2. *The amount of disclosure decreases with the opportunity cost of attending to non-disclosure* ($d\theta_A^*/d\sigma^W > 0$)*;*

3. *If* $\sigma^W\alpha^W < \sigma^D\alpha^D$, *then the amount of disclosure decreases with the elasticity of substitution between* α_A^W *and* α_A^D ($d\theta_A^*/dp < 0$)*. If* $\sigma^W\alpha^W > \sigma^D\alpha^D$, *then the amount of disclosure increases with the elasticity of substitution* ($d\theta_A^*/dp > 0$)*.*

If attention is highly substitutable (p close to 1) the condition that an interior equilibrium obtain may fail. However, if substitutability is sufficiently limited (p sufficiently large), then an interior equilibrium exists.

The comparative statics when individuals have an allocation choice (Proposition 11 Parts 1 and 2) are similar to those with exogenous α's, as described in Section 4.2, Proposition 2. Since an increase in σ^D increases the opportunity cost of attending to disclosure, it causes less attention to disclosure, and a lower threshold (more disclosure). An increase in σ^W makes attending to non-disclosure more costly, which reduces attention to withholding, and thereby increases the disclosure threshold (less disclosure). Thus, variation in the cost of attending to different arenas (the σ's) leads to comparative statics on the amount of disclosure essentially identical to those in the basic model varying the α's.

To build intuition for Part 3 of Proposition 11, observe that $\sigma^W\alpha^W - \sigma^D\alpha^D$ is positive if the attention expenditure share on α^W is greater than the expenditure share on α^D. In this situation, the balance of costs and benefit has caused the individual to substitute resources toward α^W. So when attention becomes more substitutable, he substitutes even more attention toward α^W, which increases disclosure and decreases θ^* (Proposition 2). The reverse happens when $\sigma^W\alpha^W < \sigma^D\alpha^D$.

Example: Perfect Complementarity:

The analysis matches that of the basic model even more closely in the extreme special case of perfect complementarity between between withholding and disclosing. The perfect complementarity case is obtained when the elasticity of substitution goes to zero ($p \to \infty$) in the attention allocation constraints. Since $G \leq 1$, it is optimal to allocate attention so as to equate the three components such that

$$\sigma^W\alpha_A^W = \sigma^D\alpha_A^D = \alpha_B = 1;$$

Thus, the optimal attention levels are

$$\alpha^W = \frac{1}{\sigma^W}, \quad \alpha^D = \frac{1}{\sigma^D}. \tag{34}$$

Each of the α's in (34) is a function only of its corresponding exogenous σ. It follows that all propositions of the model with exogenous attention levels are consistent with the perfect complementarity case, with comparative statics on α's interpreted as corresponding variations in σ's as in (34). For example, the amount of disclosure increases with the opportunity cost of attending to disclosure ($d\theta^*/d\sigma^D < 0$) and the amount of disclosure decreases with the opportunity cost of attending to non-disclosure ($d\theta^*/d\sigma^W > 0$).

8 Summary and Conclusion

This paper models limited attention as incomplete usage of publicly available information. Informed players decide whether or not to disclose information to an audience of observers who sometimes neglect either disclosed signals or the implications of non-disclosure. In equilibrium, we find that observers are unrealistically optimistic, disclosure is incomplete, neglect of disclosed signals increases disclosure, and neglect of a failure to disclose reduces disclosure. We also find that these insights extend to a setting in which observers choose ex ante how to allocate their limited attention. In a setting with multiple arenas of disclosure, we find that disclosure in one arena affects perceptions in fundamentally unrelated arenas, owing to cue competition, salience, and analytic interference; and that disclosure in one arena can crowd out disclosure in another.

We consider the implications of limited attention and the resulting credulity of observers for disclosure regulation. Law and regulation in the U.S. require firms to reveal information in financial reports, and to disclose other relevant information. Such regulation is not needed in the classic unravelling models of disclosure, in which rational observers, through appropriate skepticism, induce full disclosure, and full disclosure is a good thing. Extensions with costly disclosure can create a rationale for regulation, but the recommendation is very simple: force additional disclosure only if the social benefits exceed the costs.

In contrast, limited attention suggests that the balance of considerations is more complex, even when disclosure is costless. On the one hand, informed parties may conceal information in the hope of exploiting the inattention and credulity of observers. This puts regulation of disclosure on the table. However, we find, paradoxically, that regulations designed to force greater disclosure can make perceptions less accurate. For example, we find that forced disclosure in one arena can crowd out disclosure in another, and thereby can reduce welfare. Thus, even if forced disclosure in one arena creates benefits to observers in that arena, there is no presumption that forced disclosure is socially desirable.

Furthermore, forcing simultaneous disclosure in multiple arenas can also reduce welfare, for two reasons. First, greater forced disclosure can increase what we call analytic interference, wherein a disclosed signal in one arena distracts observers from analyzing the reasons for a failure of a player to disclose in the other arena. Second, even if *complete* disclosure be enforced in both arenas, greater disclosure can cause greater cue competition between disclosures— observers have trouble attending to both signals. Even though more information is publicly available, observer perceptions may on average be less accurate.

Thus, to determine whether forcing greater disclosure in one or in many arenas will improve welfare, policymakers have a challenging task. Evaluating such a policy requires an assessment of the relative importance of the different arenas, the precision of the information that might be disclosed, and the salience of the different arenas.

An issue not captured in our model that deserves further exploration is that in reality an informed party can misrepresent by 'disclosing' an incorrect value for his information signal. This issue is highlighted by U.S. and global corporate accounting scandals such as those involving Enron, WorldCom, and Parmalat. Limits to investor attention presumably affect the incentives for firms to engage in fraud, or milder shading of the truth. An interesting further direction would be to analyze how different regulatory policies influence the incentives for firms to exploit inattentive observers through misrepresentation as well as simply not disclosing.

We close by emphasizing that the general approach to limited attention offered here may be applicable to a variety of human transactions. As argued in Subsection 2.2, there are various interactions in which informed players seem to take advantage of inattention to manipulate the perceptions of others. Some possible directions that merit further exploration include the advertising of products to consumers, the reporting of firms' financial condition to investors, the presentation of information by political activists, and the presentation of personal information by individuals in their everyday lives. The approach described here—in which observers have limited ability to attend to public signals or to features of the strategic environment, salience parameters influence which signals or environmental features are attended to more, and the occurence of an event is more salient than non-occurence—may be helpful in capturing parsimoniously the effects of limited attention in a range of contexts.

Appendix

Proof of Proposition 1: We now show that if $\alpha^W < \alpha^D$ there is no equilibrium with full disclosure. The most skeptical inference in any equilibrium that could be drawn about non-disclosure would be $\hat{\theta} = \underline{\theta}$. Thus, observers' perception of a withholding player satisfies

$$\hat{\theta}^W \;\geq\; (1-\alpha^W)E[\theta] + \alpha^W \underline{\theta}. \tag{35}$$

The perception upon disclosing for a type $\theta = \underline{\theta} + \epsilon$ is

$$\hat{\theta}^D \;=\; (1-\alpha^D)E[\theta] + \alpha^D(\underline{\theta} + \epsilon). \tag{36}$$

Thus,

$$\begin{aligned} \hat{\theta}^W - \hat{\theta}^D &\geq (\alpha^D - \alpha^W)\left(E[\theta] - \underline{\theta}\right) - \alpha^D \epsilon \\ &> 0, \end{aligned} \tag{37}$$

where the last inequality holds, for given α^W and α^D, by choosing ϵ to satisfy

$$0 < \epsilon < \frac{(\alpha^D - \alpha^W)(E[\theta] - \underline{\theta})}{\alpha^D}.$$

Thus, in any equilibrium there exists a set of types with $\theta < \underline{\theta} + \epsilon$, ϵ small that prefer not to disclose. Furthermore, so long as $\gamma < 1$, $\theta^* < E[\theta]$. Otherwise, there would be an above-average type ($\theta > E[\theta]$) who prefers not to disclose. This is impossible; if he does not disclose, the average perception of his type is below $E[\theta]$, by (3), whereas if he does disclose, he is correctly perceived as having information $\theta > E[\theta]$. $\|$

Proof of Proposition 3:

$$\begin{aligned} E[\hat{\theta} - \theta] &= \int_{\underline{\theta}}^{\theta^*} \{(1-\alpha^W)E[\theta] + \alpha^W E[\theta|\theta < \theta^*] - \theta\} f(\theta) d\theta \\ &+ \int_{\theta^*}^{\bar{\theta}} \{(1-\alpha^D)E[\theta] + \alpha^D \theta - \theta\} f(\theta) d\theta \\ &= \{(1-\alpha^W)E[\theta] + \alpha^W E[\theta|\theta < \theta^*]\} F(\theta^*) - E[\theta|\theta < \theta^*] F(\theta^*) \\ &\quad +(1-\alpha^D)E[\theta][1 - F(\theta^*)] - (1-\alpha^D)\left\{E[\theta] - E[\theta|\theta < \theta^*]F(\theta^*)\right\} \\ &= (\alpha^D - \alpha^W) F(\theta^*) \left(E[\theta] - E[\theta|\theta < \theta^*]\right). \end{aligned} \tag{38}$$

Since $\alpha^D > \alpha^W$ and $E[\theta] > E[\theta|\theta < \theta^*]$, $E[\hat{\theta} - \theta] > 0$.

Proof of Proposition 4:

$$\begin{aligned}
\frac{dE[\hat{\theta} - \theta]}{d\alpha^W} &= -(E[\theta] - E[\theta|\theta < \theta^*])F(\theta^*) + (\alpha^D - \alpha^W)f(\theta^*)(E[\theta] - \theta^*)\frac{d\theta^*}{d\alpha^W} \\
\frac{dE[\hat{\theta} - \theta]}{d\alpha^D} &= (E[\theta] - E[\theta|\theta < \theta^*])F(\theta^*) + (\alpha^D - \alpha^W)f(\theta^*)(E[\theta] - \theta^*)\frac{d\theta^*}{d\alpha^D}.
\end{aligned}$$

Since $E[\theta] > E[\theta|\theta < \theta^*]$, $\alpha^D > \alpha^W$, $E[\theta] > \theta^*$, $d\theta^*/d\alpha^W < 0$ and $d\theta^*/d\alpha^D > 0$, $E[\hat{\theta} - \theta]/d\alpha^W < 0$ and $E[\hat{\theta} - \theta]/d\alpha^D > 0$ ∥

Proof of Proposition 5: By (3)–(4), we can rewrite the MSE as

$$\begin{aligned}
E[(\hat{\theta} - \theta)^2] &= \int_{\underline{\theta}}^{\theta^*} \left\{(1 - \alpha^D)E[\theta] + \alpha^D\theta^* - \theta\right\}^2 f(\theta)d\theta \\
&+ \int_{\theta^*}^{\overline{\theta}} \left\{(1 - \alpha^D)\left(E[\theta] - \theta\right)\right\}^2 f(\theta)d\theta. \qquad (39)
\end{aligned}$$

Since α^W affects $E[(\hat{\theta} - \theta)^2]$ only through θ^*,

$$\frac{dE[(\hat{\theta} - \theta)^2]}{d\alpha^W} = \frac{dE[(\hat{\theta} - \theta)^2]}{d\theta^*}\frac{d\theta^*}{d\alpha^W}, \qquad (40)$$

where

$$\begin{aligned}
\frac{dE[(\hat{\theta} - \theta)^2]}{d\theta^*} &= 2\alpha^D F(\theta^*)\left\{\alpha^D\theta^* + (1 - \alpha^D)E[\theta] - E[\theta|\theta < \theta^*]\right\} \\
&> 0. \qquad (41)
\end{aligned}$$

Thus,

$$Sign\left(\frac{dE[(\hat{\theta} - \theta)^2]}{d\alpha^W}\right) = Sign\left(\frac{d\theta^*}{d\alpha^W}\right).$$

By (6) and (10),

$$\frac{d\theta^*}{d\alpha^W} = -\frac{d\theta^*/d\gamma}{\alpha^D} < 0$$

for a stable equilibrium.

To see how the mean squared error varies with α^D, substitute θ^* of the uniform $[0,1]$ distribution case from (21) into the MSE formula, (11), and differentiating with respect to α^D. This yields

$$\begin{aligned}
\frac{dE[(\hat{\theta} - \theta)^2]}{d\alpha^D} &\equiv \left[\frac{\alpha^D}{6(2\alpha^D - \alpha^W)^4}\right]\left\{8(1 - \alpha^D)^4 - 3(\alpha^W)^2(1 - \alpha^W)^2 - 8(1 - \alpha^D)^3(3 - 2\alpha^W)\right. \\
&+ \left. 4(1 - \alpha^D)^2\left[6 - 8\alpha^W + 3(\alpha^W)^2\right] - (1 - \alpha^D)\left[8 - 16\alpha^W + 21(\alpha^W)^2 - 6(\alpha^W)^3\right]\right\},
\end{aligned}$$

which can be either positive or negative: $dE[(\hat{\theta}-\theta)^2]/d\alpha^D = -0.0354938$ when $(\alpha^W, \alpha^D) = (0.4, 0.5)$, and $dE[(\hat{\theta}-\theta)^2]/d\alpha^D = 0.000708395$ when $(\alpha^W, \alpha^D) = (0.4, 0.95)$. $\|$

Proof of Proposition 8: To condition on D_A, we substitute α_B^W from (16) and α_B^D from (18) when $\theta_A^* = \underline{\theta}$ (implying certainty of disclosure in arena A). This yields

$$\alpha_B^W(D_A) = .5(1 - s_A), \quad \alpha_B^D(D_A) = .5(1 + s_B - s_A). \tag{42}$$

Similarly, to condition on W_A, we substitute α_B^W from (16) and α_B^D from (18) when $\theta_A^* = \overline{\theta}$ (implying no disclosure in arena A). This yields

$$\alpha_B^W(W_A) = .5, \quad \alpha_B^D(W_A) = .5(1 + s_B). \tag{43}$$

From (42) and (43), the difference

$$\alpha_B^D(D_A) - \alpha_B^W(D_A) = \alpha_B^D(W_A) - \alpha_B^W(W_A) = .5 s_B$$

does not depend on the disclosure decision in arena A.

By Proposition 1, $E[\theta_B] > \theta_B^* > 0$. Therefore, if $\theta_B < \theta_B^*$, $E[\theta_B] - \theta_B > 0$, which implies that the integral on the RHS of the final equation in (28) is increasing with θ_B^*.

Thus, substituting $\phi_A = W_A$ or D_A into (28), we see that $E[\hat{\theta}_B|D_A] > E[\hat{\theta}_B|W_A]$ if and only if $\theta_B^*(D_A) > \theta_B^*(W_A)$. Let $\gamma_B = (\alpha_B^D - \alpha_B^W)/(\alpha_B^D)$. Since

$$\gamma_B(D_A) = s_B/(1 + s_B - s_A) > s_B/(1 + s_B) = \gamma_B(W_A),$$

$\theta_B^*(D_A) > \theta_B^*(W_A)$. Thus $E[\hat{\theta}_B|D_A] > E[\hat{\theta}_B|W_A]$. $\|$

Proof of Proposition 9:
Part 1:

$$
\begin{aligned}
& E[(\hat{\theta}-\theta)^2] \\
= \; & \int_{\underline{\theta}}^{\theta^*} \left\{(1-\alpha^W)E[\theta] + \alpha^W E[\theta|\theta<\theta^*] - \theta\right\}^2 f(\theta)d\theta + \int_{\theta^*}^{\overline{\theta}} \left\{(1-\alpha^D)\left(E[\theta]-\theta\right)\right\}^2 f(\theta)d\theta \\
= \; & \left\{(1-\alpha^W)E[\theta] + \alpha^W E[\theta|\theta<\theta^*]\right\}^2 \int_{\underline{\theta}}^{\theta^*} f(\theta)d\theta - 2\left\{(1-\alpha^W)E[\theta] + \alpha^W E[\theta|\theta<\theta^*]\right\} \int_{\underline{\theta}}^{\theta^*} \theta f(\theta)d\theta \\
+ \; & \int_{\underline{\theta}}^{\theta^*} \theta^2 f(\theta)d\theta + \int_{\theta^*}^{\overline{\theta}} \left\{(1-\alpha^D)\left(E[\theta]-\theta\right)\right\}^2 f(\theta)d\theta.
\end{aligned} \tag{44}
$$

$$
\begin{aligned}
\frac{\partial E[(\hat{\theta}-\theta)^2]}{\partial \theta^*} &= 2\left\{(1-\alpha^W)E[\theta]+\alpha^W E[\theta|\theta<\theta^*]\right\}\alpha^W \frac{\partial E[\theta|\theta<\theta^*]}{\partial \theta^*}F(\theta^*) \\
&+ \left\{(1-\alpha^W)E[\theta]+\alpha^W E[\theta|\theta<\theta^*]\right\}^2 f(\theta^*) - 2\alpha^W \frac{\partial E[\theta|\theta<\theta^*]}{\partial \theta^*}\int_{\underline{\theta}}^{\theta^*}\theta f(\theta)d\theta \\
&- 2\left\{(1-\alpha^W)E[\theta]+\alpha^W E[\theta|\theta<\theta^*]\right\}\theta^* f(\theta^*) \\
&+ (\theta^*)^2 f(\theta^*) - \left\{(1-\alpha^D)\left(E[\theta]-\theta^*\right)\right\}^2 f(\theta^*). \qquad (45)
\end{aligned}
$$

Since $\alpha^W < \alpha^D$, by

$$
\begin{aligned}
\frac{\partial E[\theta|\theta<\theta^*]}{\partial \theta^*} &= \frac{f(\theta^*)}{F(\theta^*)}(\theta^* - E[\theta|\theta<\theta^*]) \quad \text{and} \\
\int_{\underline{\theta}}^{\theta^*}\theta f(\theta)d\theta &= E[\theta|\theta<\theta^*]F(\theta^*),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial E[(\hat{\theta}-\theta)^2]}{\partial \theta^*} &= f(\theta^*)[1-(1-\alpha^W)^2](\theta^* - E[\theta|\theta<\theta^*])^2 \\
&+ f(\theta^*)[(1-\alpha^W)^2-(1-\alpha^D)^2](\theta^*-E[\theta])^2 > 0.
\end{aligned}
$$

Part 2:

$$
\begin{aligned}
E[\hat{\theta}-\theta] &= (\alpha^D-\alpha^W)F(\theta^*)(E[\theta]-E[\theta|\theta<\theta^*]) \\
\frac{\partial E[\hat{\theta}-\theta]}{\partial \theta^*} &= (\alpha^D-\alpha^W)f(\theta^*)(E[\theta]-\theta^*) > 0, \text{ when } E[\theta] > \theta^*. \qquad (46)
\end{aligned}
$$

‖

Proof of Proposition 10, Part 1: The mean squared error in arena i is

$$
\begin{aligned}
E[(\hat{\theta}_i-\theta_i)^2] &= \int_{\underline{\theta}_i}^{\theta_i^*}\left\{(1-\alpha_i^W)E[\theta_i]+\alpha_i^W E[\theta_i|\theta_i<\theta_i^*]-\theta_i\right\}^2 f(\theta_i)d\theta_i \\
&+ \int_{\theta_i^*}^{\overline{\theta}_i}\left\{(1-\alpha_i^D)\left(E[\theta_i]-\theta_i\right)\right\}^2 f(\theta_i)d\theta_i, \qquad (47)
\end{aligned}
$$

where α_i^W, α_i^D, and θ_i^* are functions of $\theta_{\sim i}^*$, the critical value in the other arena. The derivative of mean squared error in arena i with respect to $\theta_{\sim i}^*$ can be written as:

$$
\begin{aligned}
\frac{\partial E[(\hat{\theta}_i-\theta_i)^2]}{\partial \theta_{\sim i}^*} &= -2\frac{\partial \alpha_i^W}{\partial \theta_{\sim i}^*}(1-\alpha_i^W)(E[\theta_i]-E[\theta_i|\theta_i<\theta_i^*])^2 Pr[\theta_i<\theta_i^*] \\
&- 2\frac{\partial \alpha_i^D}{\partial \theta_{\sim i}^*}(1-\alpha_i^D)\int_{\theta_i^*}^{\overline{\theta}_i}\left(E[\theta_i]-\theta_i\right)^2 f(\theta_i)d\theta_i \\
&+ \frac{\partial \theta_i^*}{\partial \theta_{\sim i}^*}f(\theta_i^*)\alpha_i^W(2-\alpha_i^W)(\theta_i^*-E[\theta_i|\theta_i<\theta_i^*])^2 \\
&+ \frac{\partial \theta_i^*}{\partial \theta_{\sim i}^*}f(\theta_i^*)(\alpha_i^D-\alpha_i^W)(2-\alpha_i^D-\alpha_i^W)(\theta_i^*-E[\theta_i])^2. \qquad (48)
\end{aligned}
$$

By inspection of equations (15)-(18), the derivatives of α_i^W and α_i^D with respect to $\theta^*_{\sim i}$ are

$$\frac{\partial \alpha_i^W}{\partial \theta^*_{\sim i}} = \frac{\partial \alpha_i^D}{\partial \theta^*_{\sim i}} = 0.5 s_{\sim i} f(\theta^*_{\sim i}) > 0. \tag{49}$$

Also,

$$\frac{\partial \theta_i^*}{\partial \theta^*_{\sim i}} = \frac{\partial \theta_i^*}{\partial \gamma_i}\frac{\partial \gamma_i}{\partial \theta^*_{\sim i}} = \frac{\partial \theta_i^*}{\partial \gamma_i}\left[-0.5 s_{\sim i} f(\theta^*_{\sim i})\frac{\alpha_i^D - \alpha_i^W}{(\alpha_i^D)^2}\right] < 0. \tag{50}$$

Therefore, in the neighborhood of a stable equilibrium ($\partial \theta_i^* / \partial \gamma_i > 0$), the mean squared error of arena i decreases with the threshold of the other arena $\sim i$,

$$\frac{\partial E[(\hat{\theta}_i - \theta_i)^2]}{\partial \theta^*_{\sim i}} < 0.$$

$\|$

Proof of Proposition 10, Part 2:

$$\begin{aligned}
\frac{\partial E[\hat{\theta}_i - \theta_i]}{\partial \theta^*_{\sim i}} &= \left(\frac{\partial \alpha_i^D}{\partial \theta^*_{\sim i}} - \frac{\partial \alpha_i^W}{\partial \theta^*_{\sim i}}\right) F(\theta_i^*)(E[\theta_i] - E[\theta_i | \theta_i < \theta_i^*]) \\
&\quad + (\alpha_i^D - \alpha_i^W) f(\theta_i^*)(E[\theta_i] - \theta_i^*)\frac{\partial \theta_i^*}{\partial \theta^*_{\sim i}} \\
&= (\alpha_i^D - \alpha_i^W) f(\theta_i^*)(E[\theta_i] - \theta_i^*)\frac{\partial \theta_i^*}{\partial \theta^*_{\sim i}} < 0 \qquad (51)
\end{aligned}$$

The last line follows from equations (49) and (50). $\|$

Proof of Proposition 10, Part 3: To prove Part 3, we differentiate the welfare function with respect to θ_A^* when θ_A and θ_B are distributed uniformly over $[0,1]$. The attention probabilities are given in equations (15)-(18).

$$\begin{aligned}
\frac{\partial W}{\partial \theta_A^*} &= \frac{1}{48}(-9\lambda(\theta_A^*)^2 \\
&- 2s\left\{-1 + (\theta_B^*)^3 + \lambda\left[4 - 12\theta_A^* - (\theta_B^*)^3 + 9(\theta_A^*)^2 + 3(\theta_A^*)^2\theta_B^*\right]\right\} \\
&+ s^2\{2(\theta_B^* - 1)[3(\theta_B^* - 1)\theta_B^* + \theta_A^* + \theta_A^*\theta_B^* + \theta_A^*(\theta_B^*)^2] \\
&+ \lambda[-3 + 12(\theta_B^*)^2 - 6(\theta_B^*)^3 - 9(\theta_A^*)^2 + 18(\theta_A^2)^2\theta_B^* \\
&+ 3(\theta_A^2)^2(\theta_B^*)^2 + 14\theta_A^* - 24\theta_A^*\theta_B^* - 2\theta_A^*(\theta_B^*)^3]\}). \qquad (52)
\end{aligned}$$

Using the symmetric solution to equation (25) when salience is s, $\partial W / \partial \theta_A^* = 0.02244$ when $(s, \lambda) = (0.8, 0.2)$, and it becomes -0.01663 when $(s, \lambda) = (0.8, 0.9)$. $\|$

Calculations Underlying Figure 3: We differentiate the welfare W with respect to the common threshold θ^* where θ_A and θ_B are distributed uniformly over $[0,1]$

$$\frac{\partial W}{\partial \theta^*} = \frac{1}{48}\left\{-9\theta^{*2} - 2s[2-12\theta^* + 9(\theta^*)^2 + 4(\theta^*)^3] + s^2[-3+22\theta^* - 45(\theta^*)^2 + 24(\theta^*)^3 + 5(\theta^*)^4]\right\}.$$

The derivative $\partial W/\partial\theta^*$ is positive for high value and negative for low values of s. For example, $\partial W/\partial\theta^* = 0.0114$ when $s = 0.8$, and $\partial W/\partial\theta^* = -0.0079$ when $s = 0.2$. $\|$

Proof of Proposition 11: For notational simplicity, we henceforth suppress the A subscripts of α_A^W, α_A^D, and θ_A. By (32) and (33) we substitute out

$$\alpha_B = \frac{1 - s^W(\sigma^W\alpha^W)^p - s^D(\sigma^D\alpha^D)^p}{(1-s^W-s^D)},$$

and write the first order conditions of the optimization problem with respect to α^W and α^D as

$$\begin{aligned} 0 &= \int_{\underline{\theta}}^{\theta^*}(-\theta)f(\theta)d\theta - \frac{Kp}{(1-s^W-s^D)}s^W(\sigma^W)^p(\alpha^W)^{p-1} \\ 0 &= \int_{\theta^*}^{0}(-\theta)f(\theta)d\theta - \frac{Kp}{(1-s^W-s^D)}s^D(\sigma^D)^p(\alpha^D)^{p-1}. \end{aligned} \tag{53}$$

It follows that

$$\frac{\alpha^W}{\alpha^D} = \left(\frac{\sigma^D}{\sigma^W}\right)\left[\frac{\sigma^D s^D\int_{\underline{\theta}}^{\theta^*}(-\theta)f(\theta)d\theta}{\sigma^W s^W\int_{\theta^*}^{0}(-\theta)f(\theta)d\theta}\right]^{1/(p-1)}. \tag{54}$$

The disclosure threshold is determined by equations (5) and (6) of the basic model applied to arena A. By (54), the ratio α^W/α^D does not depend on K, the weight in the observers objective on payoffs derived from attending to arena B. Therefore, we can ensure that $\alpha_B > 0$ by selecting K sufficiently large without affecting the equilibrium disclosure level. Also, for appropriate values of parameters σ^W and σ^D, $\alpha^D > \alpha^W$, ensuring that the equilibrium involves only partial disclosure. By equation (54) and since $E[\theta]=0$, equation (5) can be rearranged to yield

$$\theta^* = (1-\gamma)E[\theta|\theta<\theta^*] = (1-\gamma)\left(\frac{\int_{\underline{\theta}}^{\theta^*}\theta f(\theta)d\theta}{\int_{\underline{\theta}}^{\theta^*}f(\theta)d\theta}\right) \tag{55}$$

$$\begin{aligned} &= \frac{\alpha^W}{\alpha^D}\left(\frac{\int_{\underline{\theta}}^{\theta^*}\theta f(\theta)d\theta}{\int_{\underline{\theta}}^{\theta^*}f(\theta)d\theta}\right) = \left(\frac{\sigma^D}{\sigma^W}\right)^{\frac{p}{p-1}}\left(\frac{s^D}{s^W}\right)^{\frac{1}{p-1}}\left[\frac{\int_{\underline{\theta}}^{\theta^*}(-\theta)f(\theta)d\theta}{\int_{\theta^*}^{0}(-\theta)f(\theta)d\theta}\right]^{\frac{1}{p-1}}\left(\frac{\int_{\underline{\theta}}^{\theta^*}\theta f(\theta)d\theta}{\int_{\underline{\theta}}^{\theta^*}f(\theta)d\theta}\right) \\ &= -\left(\frac{\sigma^D}{\sigma^W}\right)^{\frac{p}{p-1}}\left(\frac{s^D}{s^W}\right)^{\frac{1}{p-1}}\frac{\left[\int_{\underline{\theta}}^{\theta^*}(-\theta)f(\theta)d\theta\right]^{\frac{p}{p-1}}}{\left[\int_{\theta^*}^{0}(-\theta)f(\theta)d\theta\right]^{\frac{1}{p-1}}\left[\int_{\underline{\theta}}^{\theta^*}f(\theta)d\theta\right]}. \end{aligned} \tag{56}$$

Multiplying both sides by -1 and taking natural logs gives

$$\begin{aligned} Ln(-\theta^*) &= \frac{p}{p-1} Ln\left(\frac{\sigma^D}{\sigma^W}\right) + \frac{1}{p-1} Ln\left(\frac{s^D}{s^W}\right) + \frac{p}{p-1} Ln\left[\int_{\underline{\theta}}^{\theta^*} (-\theta) f(\theta) d\theta\right] \\ & -\frac{1}{p-1} Ln\left[\int_{\theta^*}^{0} (-\theta) f(\theta) d\theta\right] - Ln\left[\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta\right]. \end{aligned} \tag{57}$$

Differentiating both sides of equation (57) with respect to σ^D,

$$\frac{1}{\theta^*}\frac{\partial \theta^*}{\partial \sigma^D} = \frac{p}{p-1}\frac{1}{\sigma^D} + \frac{p}{p-1}\frac{-\theta^* f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*}(-\theta) f(\theta) d\theta}\frac{\partial \theta^*}{\partial \sigma^D} - \frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta) f(\theta) d\theta}\frac{\partial \theta^*}{\partial \sigma^D} - \frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta}\frac{\partial \theta^*}{\partial \sigma^D},$$

$$\frac{\partial \theta^*}{\partial \sigma^D}\left[\frac{1}{\theta^*} + \frac{p}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*}(-\theta) f(\theta) d\theta} + \frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta) f(\theta) d\theta} + \frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta}\right] = \frac{p}{p-1}\frac{1}{\sigma^D} \tag{58}$$

Similarly, differentiating both sides of equation (57) with respect to σ^W gives

$$\frac{\partial \theta^*}{\partial \sigma^W}\left[\frac{1}{\theta^*} + \frac{p}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*}(-\theta) f(\theta) d\theta} + \frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta) f(\theta) d\theta} + \frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta}\right] = -\frac{p}{p-1}\frac{1}{\sigma^W}. \tag{59}$$

Multiplying equation (55) by $\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta$ and differentiating both sides with respect to γ gives $\partial\theta^*/\partial\gamma$, which was shown to be positive in (10), so

$$\frac{\partial \theta^*}{\partial \gamma} = \frac{-E[\theta|\theta < \theta^*]}{1 + \gamma\frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta}\theta^*} > 0 \Rightarrow 1 + \gamma\frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta}\theta^* > 0. \tag{60}$$

The second term of the expression inside the brackets in equations (58) and (59) can be rewritten using equation (55) as

$$\frac{p}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*}(-\theta) f(\theta) d\theta} = -\frac{p}{p-1}\frac{(1-\gamma) f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta}. \tag{61}$$

Substituting the right hand side of equation (61) for the second term inside the brackets in equations (58) and (59), the full bracketed term becomes

$$\begin{aligned} & \frac{1}{\theta^*} + \frac{p}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*}(-\theta) f(\theta) d\theta} + \frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta) f(\theta) d\theta} + \frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta} \\ = & \frac{1}{\theta^*} - \frac{p}{p-1}\frac{(1-\gamma) f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta} + \frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta) f(\theta) d\theta} + \frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta} \\ = & \frac{1}{\theta^*} + \frac{\gamma f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta} - \frac{(1-\gamma)}{p-1}\frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta) d\theta} + \frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta) f(\theta) d\theta}. \end{aligned} \tag{62}$$

From (60) and since $\theta^* < 0$,

$$\frac{1}{\theta^*} + \frac{\gamma f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta)d\theta} = \frac{1}{\theta^*}\left(1 + \frac{\gamma f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta)d\theta}\theta^*\right) < 0. \tag{63}$$

$p > 1$ and $\theta^* < 0$ also imply that

$$\frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta)f(\theta)d\theta} < 0. \tag{64}$$

Therefore,

$$\frac{1}{\theta^*} + \frac{\gamma f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta)d\theta} - \frac{(1-\gamma)}{p-1}\frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta)d\theta} + \frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta)f(\theta)d\theta} < 0. \tag{65}$$

From (58), (59), and (65), Parts 1 and 2 of Proposition 11 follow immediately.

To prove Proposition 11 Part 3, we differentiate both sides of equation (57) with respect to p:

$$\begin{aligned}
\frac{1}{\theta^*}\frac{\partial\theta^*}{\partial p} = & -\frac{1}{(p-1)^2}Ln\left(\frac{\sigma^D}{\sigma^W}\right) - \frac{1}{(p-1)^2}Ln\left(\frac{s^D}{s^W}\right) - \frac{1}{(p-1)^2}Ln\left[\int_{\underline{\theta}}^{\theta^*}(-\theta)f(\theta)d\theta\right] \\
& + \frac{p}{p-1}\frac{-\theta^* f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*}(-\theta)f(\theta)d\theta}\frac{\partial\theta^*}{\partial p} + \frac{1}{(p-1)^2}Ln\left[\int_{\theta^*}^{0}(-\theta)f(\theta)d\theta\right] \\
& - \frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta)f(\theta)d\theta}\frac{\partial\theta^*}{\partial p} - \frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta)d\theta}\frac{\partial\theta^*}{\partial p},
\end{aligned} \tag{66}$$

so

$$\begin{aligned}
& \frac{\partial\theta^*}{\partial p}\left[\frac{1}{\theta^*} + \frac{p}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*}(-\theta)f(\theta)d\theta} + \frac{1}{p-1}\frac{\theta^* f(\theta^*)}{\int_{\theta^*}^{0}(-\theta)f(\theta)d\theta} + \frac{f(\theta^*)}{\int_{\underline{\theta}}^{\theta^*} f(\theta)d\theta}\right] \\
= & -\frac{1}{(p-1)^2}Ln\left[\left(\frac{\sigma^D s^D}{\sigma^W s^W}\right)\left(\frac{\int_{\underline{\theta}}^{\theta^*}(-\theta)f(\theta)d\theta}{\int_{\theta^*}^{0}(-\theta)f(\theta)d\theta}\right)\right].
\end{aligned} \tag{67}$$

By (54), the right hand side of equation (67) can be rewritten as

$$-\frac{1}{(p-1)^2}Ln\left[\left(\frac{\sigma^D s^D}{\sigma^W s^W}\right)\left(\frac{\int_{\underline{\theta}}^{\theta^*}-\theta f(\theta)d\theta}{\int_{\theta^*}^{0}-\theta f(\theta)d\theta}\right)\right] = -\frac{1}{(p-1)^2}Ln\left[\left(\frac{\sigma^W\alpha^W}{\sigma^D\alpha^D}\right)^{p-1}\right]. \tag{68}$$

By (65), the left hand side of equation (67) is negative, and the sign of the right hand side depends on whether $\sigma^W\alpha^W$ is greater or less than $\sigma^D\alpha^D$, $\partial\theta^*/\partial p < 0$ when $\sigma^W\alpha^W < \sigma^D\alpha^D$ and $\partial\theta^*/\partial p > 0$ $\sigma^W\alpha^W > \sigma^D\alpha^D$. This confirms Proposition 11 Part 3.

References

Baker, A. G., P. Mercier, F. Valleettourangeau, R. Frank, and M. Pan, 1993, Selective associations and causality judgments – Presence of a strong causal factor may reduce judgements of a weaker one, *Journal of Experimental Psychology – Learning, Memory and Cognition* **19**, 414–432.

Bernard, V. L. and J. K. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* **13**, 305–340.

Boot, A. W. A. and A. V. Thakor, 2001, The many faces of information disclosure, *Review of Financial Studies* **14**, 1021–1057.

Busemeyer, J. R., I. J. Myung, and M. McDaniel, 1993, Cue competition effects – empirical tests of adaptive network leaninrg models, *Psychological Science* **4**, 190–195.

Daniel, K. D., D. Hirshleifer, and A. Subrahmanyam, 2001, Overconfidence, arbitrage, and equilibrium asset pricing, *Journal of Finance* **56**, 921–965.

Daniel, K. D., D. Hirshleifer, and S. H. Teoh, 2002, Investor psychology in capital markets: Evidence and policy implications, *Journal of Monetary Economics* **49** (1)/*Carnegie Rochester Series in Public Policy* **56** **49**, 139–209.

Darrough, M. and N. Stoughton, 1990, Financial disclosure policy in an entry game, *Journal of Accounting and Economics* **12**, 219–244.

DellaVigna, S. and J. Pollet, 2003, Attention, demographics, and the stock market, UC Berkeley and Harvard University.

DellaVigna, S. and J. Pollet, 2004, Strategic release of information on fridays: Evidence from earnings announcements, UC Berkeley and UIUC, working paper.

Feltovich, N., R. Harbaugh, and T. To, 2002, Too cool for school? Signalling and countersignalling, *Rand Journal of Economics* **33**, 630–649.

Fischoff, B., P. Slovic, and S. Lichtenstein, 1978, Fault trees: Sensitivity of estimated failure probabilities to problem representation, *Journal of Experimental Psychology: Human Perception and Performance* **4**, 330–334.

Fishman, M. and K. Hagerty, 1989, Disclosure decisions by firms and the competition for price efficiency, *Journal of Finance* **44**, 633–646.

Fishman, M. J. and K. M. Hagerty, 2003, Mandatory vs. voluntary disclosure in markets with informed and uninformed customers, *Journal of Law, Economics & Organization* **19**, 45–63.

Fiske, S. T., 1995,. Social cognition, . In A. Tesser, ed., *Advanced Social Psychology*, pp. 149–194. (McGraw-Hill, New York).

Fiske, S. T. and S. Taylor, 1991,. Social encoding: Attention and consciousness, . In *Social Cognition* (2nd ed.)., Chapter 7, pp. 243–294. (McGraw-Hill, New York, NY).

Francis, J., D. Pagach, and J. Stephan, 1992, The stock market response to earnings announcements released during trading versus nontrading periods, *Journal of Accounting Research* **30**, 165–184.

Gabaix, X. and D. Laibson, 2002,. The 6D bias and the equity premium puzzle, . In B. Bernanke and K. Rogoff, eds., *NBER Macroeconomics Annual*, pp. 257–312. (MIT Press, Cambridge, Mass).

Gabaix, X. and D. Laibson, 2004, Bounded rationality and directed cognition, MIT and Harvard University, working paper.

Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg, 2003, The allocation of attention: Theory and evidence, MIT, Harvard University and NBER, working paper.

Gifford, S., 1992, Allocation of entrepreneurial attention, *Journal of Economic Behavior and Organization* **19**, 265–284.

Gifford, S., 1992, Innovation, firm size and growth in a centralized organization, *RAND Journal of Economics* **23**, 284–298.

Gifford, S. and C. Wilson, 1995, A model of project evaluation with limited attention, *Economic Theory* **5**, 67–78.

Grossman, S. J., 1981, The informational role of warranties and private disclosure about product quality, *Journal of Law and Economics* **24**, 461–483.

Hanson, J. D. and D. A. Kysar, 1999, Taking behavioralism seriously: Some evidence of market manipulation, *Harvard Law Review* **112**, 1420–1572.

Hensley, S. and P. Landers, 2004, If drug firm's star falls, does anyone hear it? Companies can be tight-lipped when hot prospects fail, leaving investors to deduce their fate, *Wall Street Journal*. February 24.

Hirshleifer, D. and S. H. Teoh, 2003, Limited attention, information disclosure, and financial reporting, *Journal of Accounting & Economics* **36**, 337–386.

Hirst, D. and P. E. Hopkins, 1998, Comprehensive income reporting and analysts' valuation judgments, *Journal of Accounting Research* **36**, 47–75. Supplement.

Ho, T. and R. Michaely, 1988, Information quality and market efficiency, *Journal of Financial and Quantitative Analysis* **5**, 357–386.

Hof, R. D., 2004, Earth to Silicon Valley: You've lost this battle; Why tech should end its fight against options expensing, *Business Week*, 75–76. July 12.

Huberman, G. and T. Regev, 2001, Contagious speculation and a cure for cancer, *Journal of Finance* **56**, 387–396.

Ikenberry, D., J. Lakonishok, and T. Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* **39**, 181–208.

Jovanovic, B., 1982, Truthful disclosure of information, *Bell Journal of Economics* **13**, 36–44.

Kahneman, D., 1973, *Attention and Effort* (Prentice-Hall, Englewood Cliffs, New Jersey).

Kahneman, D. and A. Tversky, 1973, On the psychology of prediction, *Psychological Review* **80**, 237–25l.

Kennedy, S. R. and K. M. Dimick, 1987, Career maturity and professional sports expectations of college football and basketball player, *Journal of College Student Development* **28**, 293–297.

Klibanoff, P., O. Lamont, and T. A. Wizman, 1998, Investor reaction to salient news in closed-end country funds, *Journal of Finance* **53**, 673–699.

Kruschke, J. K. and M. K. Johansen, 1999, A model of probabilistic category learning, *Journal of Experimental Psychology: Learning, Memory, and Cognition* **25**, 1083–1119.

Loughran, T. and J. Ritter, 1995, The new issues puzzle, *Journal of Finance* **50**, 23–52.

March, J. and H. Simon, 1958, *Organizations* (Wiley, New York, NY).

Mathios, A. D., 2000, The impact of mandatory disclosure laws on product choices: An analysis of the salad dressing market, *Journal of Law and Economics* **43**, 651–677.

Mayer, J., 2002, The accountants' war: Arthur Levitt wanted to rein in the auditors. but he didn't reckon on congress, *The New Yorker,* **April 22 & 29**.

Milgrom, P. and J. Roberts, 1986, Relying on the information of interested parties, *RAND Journal of Economics* **17**, 18–32.

Milgrom, P. R., 1981, Good news and bad news: Representation theorems and applications, *Bell Journal of Economics* **12**, 380–391.

Newman, J., W. Wolff, and E. Hearst, 1980, The feature-positive effect in adult human subjects, *Journal of Experimental Psychology: Human Learning and Memory* **6**, 630–650.

Nisbett, R. and L. Ross, 1980, *Human Inference: Strategies and Shortcomings of Social Judgment* (Prentice-Hall, Englewood Cliffs, NJ).

Paredes, T., 2004, Blinded by the light: Information overload and its consequences for securities regulation, *Washington University Law Quarterly, forthcoming*.

Pashler, H. E., 1998, *The Psychology of Attention* (MIT Press, Cambridge, MA).

Patell, J. and M. Wolfson, 1982, Good news, bad news, and the intraday timing of corporate disclosures, *The Accounting Review*, 509–527.

Peng, L., 2004, Learning with information capacity constraints, *Journal of Financial and Quantitative Analysis, forthcoming*.

Peng, L. and W. Xiong, 2002, Capacity constrained learning and asset price comovement, working paper, Duke University and Princeton University.

Radner, R., 1975, A behavioral model of cost reduction, *Rand Journal of Economics* **6**, 196–215.

Radner, R. and M. Rothschild, 1975, On the allocation of effort, *Journal of Economic Theory* **10**, 358–376.

Rashes, M. S., 2001, Massively confused investors making conspicuously ignorant choices (MCI - MCIC), *Journal of Finance* **56**, 1911–1928.

Read, D., G. Loewenstein, and M. Rabin, 1999, Choice bracketing, *Journal of Risk and Uncertainty* **19**, 171–197.

Riley, D. and H. L. Roitblat, 1978,. Selective attention and related cognitive processes in pigeons, . In S. H. Hulse, H. Fowler, and W. K. Honig, eds., *Cognitive Processes in Animal Behavior*, pp. 249–276. (Erlbaum, Hillsdale, NJ).

Simon, H., 1976, *Administrative Behavior* (3rd ed.). (Free Press, New York, NY).

Skinner, D., 1994, Why firms voluntarily disclose bad news, *Journal of Accounting Research* **32**, 38–60.

Sloan, R., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting Review* **71**, 289–315.

Smith, A., 1776, *An Inquiry into the Nature and Causes of the Wealth of Nations* (W. Strahan and T. Cadell, London).

Taylor, S. and S. Fiske, 1978, Salience, attention and attribution, *Advances in Experimental Social Psychology* **11**, 249–288.

Teoh, S. H., 1997, Information disclosure and voluntary contributions to public goods, *Rand Journal of Economics* **29**, 385–406.

Teoh, S. H. and C.-Y. Hwang, 1991, Non-disclosure and adverse disclosure as signals of firm value, *Review of Financial Studies* **4**, 283–313.

Teoh, S. H., I. Welch, and T. J. Wong, 1998, Earnings management and the long-term market performance of initial public offerings, *Journal of Finance* **53**, 1935–1974.

Verrecchia, R. E., 1983, Discretionary disclosure, *Journal of Accounting and Economics* **5**, 179–194.

Weinstein, N. D., 1980, Unrealistic optimism about future life events, *Journal of Personality and Social Psychology* **39**, 806–820.

Weinstein, N. D., 1982, Unrealistic optimism about susceptibility to health-problems, *Journal of Behavioral Medicine* **5**, 441–460.

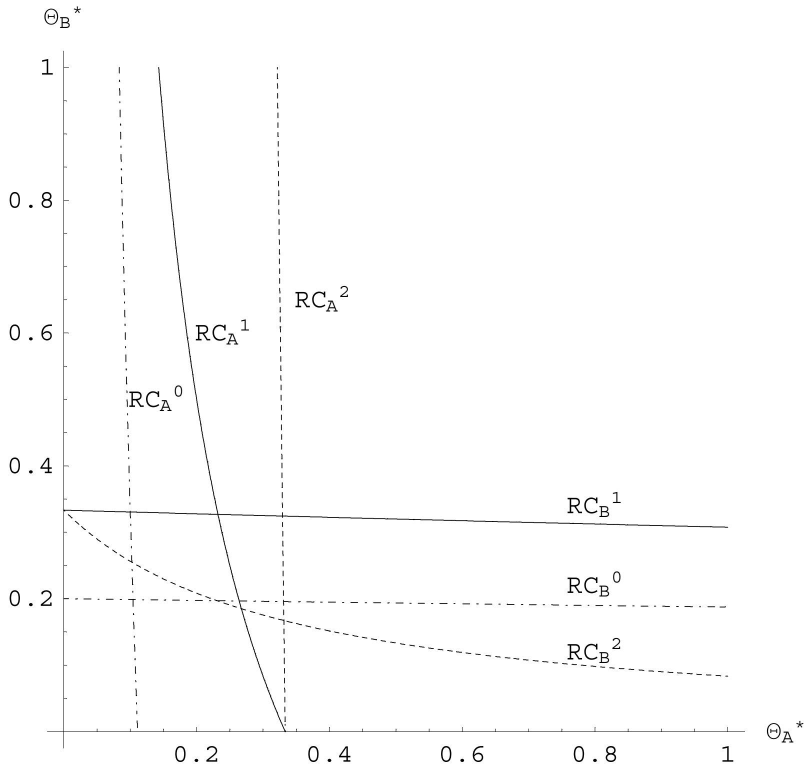

Figure 1: Reaction Curves for Disclosure in Arenas A and B

RC_A^0 and RC_B^0 are reaction curves for disclosure in arenas A and B when $(s_A, s_B) = (0.1, 0.3)$. RC_A^1 and RC_B^1 are the reaction curves when $(s_A, s_B) = (0.2, 0.8)$, and RC_A^2 and RC_B^2 are the reaction curves when $(s_A, s_B) = (0.9, 0.1)$.

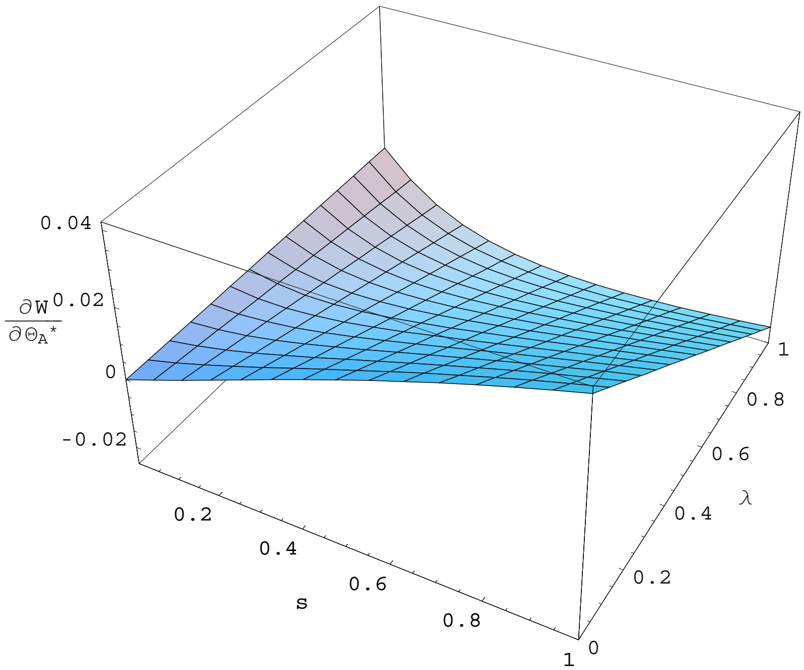

Figure 2: Welfare effect of exogenous increase in the disclosure threshold of arena A

The graph shows the derivative of welfare as defined by equation (30) with respect to an exogenous increase in the disclosure threshold in arena A, θ_A^*, as a function of the common salience of disclosure s, and the weight λ on arena A in the social welfare function.

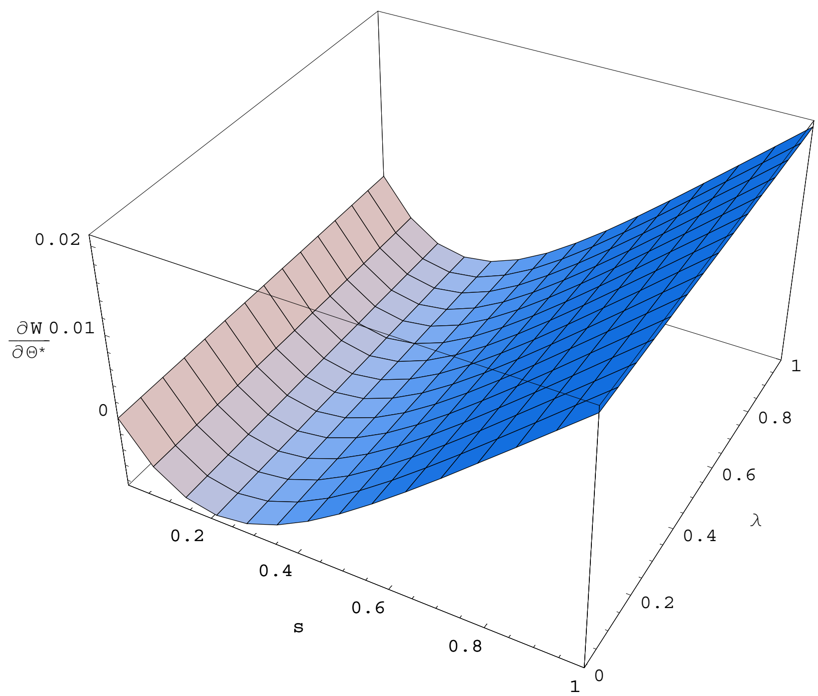

Figure 3: Welfare effect of exogenous increase in the common disclosure threshold of arenas A and B

The graph shows the derivative of welfare as defined by equation (30) with respect to an exogenous increase in the common disclosure threshold for arenas A and B, θ^*, as a function of the common salience of disclosure s, and the weight λ on arena A in the social welfare function.