THE UNIVERSITY OF CHICAGO

LIMITED ATTENTION: IMPLICATIONS FOR FINANCIAL REPORTING

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE UNIVERSITY OF CHICAGO
BOOTH SCHOOL OF BUSINESS
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
JINZHI LU

CHICAGO, ILLINOIS
JUNE 2019

For my wife Ruofan

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I am grateful to the members of my dissertation committee for their guidance and support: Haresh Sapra (chair), Philip Berger, Jonathan Bonham, and Pingyang Gao. I also appreciate helpful comments from Cyrus Aghamolla, John Barrios, Aysa Dordzhieva, Vivian Fang, Frank Gigler, Joao Granja, Zhaoyang Gu, Ilan Guttman, Eunhee Kim, Christian Laux, Volker Laux, Christian Leuz, Miao Liu, Yao Lu, Charles Mcclure, Xiaojing Meng, Michael Minnis, Jamse Ohlson, Douglas Skinner, Abbie Smith, Anastasia Zakolyukina, Guochang Zhang, Ruohuo Zheng, and workshop participants at the University of Chicago, New York University, University of Minnesota, University of Texas at Austin, University of Hong Kong, Hong Kong Polytechnic University, City University of Hong Kong, and Chinese University of Hong Kong. I gratefully acknowledge financial support from the University of Chicago Booth School of Business. All errors are my own.

ABSTRACT

I develop a theory to study the consequence of providing more detailed information to rationally inattentive investors. I first shed light on a fundamental trade-off between disclosing a summary versus disclosing details: although a summary contains less information about fundamentals than details, it is easier to process. Moreover, I find that when investors' decisions are complements, reporting details together with a summary does not always dominate reporting a summary alone. The main reason for this surprising result is that when investors care about the decisions of others, they are induced to process details, even if doing so is very costly. By uncovering a potential cost of reporting details, my paper contributes a novel insight into the consequence of providing detailed information, an issue that is currently being considered by the FASB in its performance disaggregation project.

CHAPTER 1

LIMITED ATTENTION: IMPLICATIONS FOR FINANCIAL REPORTING

1.1 Introduction

In financial statements, accounting information is usually presented in summarized forms with a lot of details suppressed. Based on the theory that finer information can improve investors' decision quality, disaggregating the summarized items by providing more details seems appealing (e.g., Blackwell (1950)). In fact, the FASB has similar thoughts in its performance disaggregation project (see FASB's performance disaggregation project for details). Specifically, they are considering the disaggregation of income statement items, under the objective that such disaggregation can "improve the decision-usefulness of income statements". In this paper, I provide a theory that focuses on the trade-off of providing more disaggregated information. My model highlights a novel cost of providing more details: investors with limited attention could be overloaded by too much detail.

I develop a model in which investors are rational but are subject to limited information processing capacity. I first shed light on a fundamental trade-off between disclosing a summary and disclosing details: although a summary is less precise than details, it is easier to process. More specifically, in an unlimited attention setting, because aggregation inevitably results in information loss, details should always be preferred to a summary. However, in a limited attention setting, investors may prefer a summary if their information processing capacity is low.

Given this trade-off between a summary and details, one might think reporting details along with a summary can ensure the best outcome. This intuition only holds when investors do not care about the decisions of others and can thus ignore the details they cannot process. However, I show that the intuition no longer holds when investors' decisions are complements. This result occurs because when investors care about the decisions of others, they are induced to process details, even if doing so is very costly. Put differently, suppressing details could be desirable in environments in which complementarity is the key feature such as: (1) the stock market where traders care about price movements that are dependent on the actions of many traders, (2) IPOs for which investors need to coordinate their investment decisions, and (3) banks when the bank's creditors need to coordinate their decisions on whether or not to run. This result rationalizes the notion of information overload, defined as a reduction in decision quality when too much information is present (see Speier et al. (1999)). It also helps the FASB identify a potential cost of reporting more details in financial statements.

To illustrate my results, I model rationally inattentive investors following the seminal work of Sims (2003). This approach allows me to analyze the information overload problem without relying on behavioral assumptions. Sims assumes that investors behave as if public signals are transmitted and observed through finite capacity channels. This means public signals will be confounded by private noise after being processed by investors. Sims uses mutual entropy, a concept in information theory, to link investors' information processing capacity with the property of investors' private signals. Specifically, the less noisy an investor's private signal, the more information is transmitted, and the higher the mutual entropy between the original public signal and the private signal. Limited information processing capacity can thus be captured by imposing the constraint that the mutual entropy is finite.

I then introduce investors' limited attention into the investment coordination game in Angeletos and Pavan (2004), which is a clean approach to model complementarity. In the model, a continuum of investors considers their investment decisions in a firm. The return of an investor's investment increases not only in the firm's fundamental, but also in the aggregate level of investment. To help investors make better decisions, a social planner can mandate that the firm provides public signals about its fundamental. Specifically, the firm can disclose a summary of the fundamental, or details of the fundamental, or both.

The firm's public disclosure can affect investors' welfare in two ways. First, investors process the public disclosure and then update their beliefs about the fundamental so that their investment decisions are fine-tuned to the fundamental. Second, limited attention implies that public signals will be confounded by private noise after being processed by the investors. Consequently, following public disclosure, investors will find that aligning their investment decisions is harder when the magnitude of the private noise is larger. The total value of a public signal can thus be intuitively divided into two components: information value, defined as the extent to which the signal brings investors' investment closer to the fundamental, and coordination value, defined as the extent to which investors can coordinate their investment decisions following the signal.

I start my analysis by showing that even though details are more precise measures of the fundamental, after being processed, the summary can have higher information value than details. The main intuition is as follows. First, details require more capacity to process than the summary because each detail needs to be separately processed. Furthermore, the ability of processing details increases as investors' information processing capacity increases. Second, because details are more precise measures of the fundamental, details contain more information than the summary. Therefore, the trade-off between the summary and details is in fact a trade-off between signal precision and processing difficulty. Hence, I show that the summary has higher information value than details if and only if investors' information processing capacity is sufficiently small.

I then show that the summary has higher coordination value than details. The key intuition is that when investors' information processing capacity is divided among details, each detail is processed less thoroughly, compared to the case where all the capacity is devoted to processing the summary. Consequently, signals of details are more private than signals of the summary. Therefore, investors achieve less synergy among their investment decisions if they focus on the details rather than the summary. As a result, details tend to impede coordination, regardless of investors' information processing capacity.

Finally, I explore whether reporting details together with the summary is optimal. I show that when the summary and details are reported together, investors may choose to focus on details in equilibrium, even if doing so is Pareto inefficient (i.e., investors would be better off if they instead focused on the summary). Thus, attaching details to the summary may reduce investors' welfare. The intuition is as follows. First, because details are precise measures of the fundamental, predicting the aggregate investment is equivalent to predicting the fundamental if almost every investor focuses on the details. Consequently, if investors' information capacity is sufficiently large, such that the details have a higher information value, an equilibrium in which all investors focus on the details exists. However, because details have lower coordination value and higher information value, this equilibrium is less efficient than the equilibrium where all investors focus on the summary, if and only if the strength of complementarity is higher than some threshold (i.e., coordination is sufficiently important). The striking aspect of this result is that information overload happens precisely when investors have relatively high capacity (i.e., when details have higher information value than the summary). It is the lower coordination value of details that drives the information overload result.

One might think the above results rely on the assumption that coordination is socially beneficial, as in Angeletos and Pavan (2004). Interestingly, I show that a similar result holds even though coordination is not socially beneficial. To illustrate this point, in one of the extensions of the model (Section 1.7), I consider an alternative setting in which investors still need to coordinate their decisions, but the social planner only cares about whether investors' decisions are fine-tuned to the fundamental, as in Morris and Shin (2002). I find a similar result: suppressing details is desirable when investors' capacity is low or when the strength of complementarity is sufficiently high. The intuition is as follows. First, for the same reason as before, when the firm reports both the summary and details, if investors' information capacity is sufficiently large, such that details have higher information value, an equilibrium in which all investors focus on the details exists. Second, in this equilibrium, investors put less weight on their signals and more weight on the prior of the fundamental, compared to the equilibrium in which all investors focus on the summary. This is because signals of details are more private than signals of the summary. As a result, to better coordinate their decisions, investors behave as if they "discount" signals of details more than signals of the summary. Hence, when the strength of complementarity is sufficiently high, such that the "discount" effect of details outweighs the information advantage of details, details should be suppressed.

In summary, I show that the limited attention approach generates interesting implications for financial reporting, especially in environments where investors' decisions are complements. Specifically, while the summary tends to be less precise than the details, I find the summary can dominate details because details are harder to process and impede coordination. Moreover, I find attaching details to the summary does not always improve the quality of investors' decisions.

1.1.1 Related literature

That investors have limited attention is consistent with much empirical and experimental evidence. Hirshleifer and Teoh (2003) provides a good survey of theories in psychology and experimental evidence that support this perspective. In fact, empirical evidence shows that even sophisticated investors, such as institutional investors, analysts, and market makers, have limited attention (see Kempf et al. (2016), Driskill et al. (2019), and Chakrabarty and Moulton (2012)). The limited attention perspective is also the key to understanding many policy debates, including the debate on whether accounting items should be disclosed or recognized (Bernard and Schipper (1994)). In addition, the recent literature on textual analysis implies that the complexity of financial statements may affect investors' assessment of firms' performance (Li (2008); Guay et al. (2016)). These studies thus indirectly support the idea that investors have limited information processing capacity.

My paper contributes to the concept of information overload by offering a theory of it. Information overload can be defined as a reduction in decision quality when too much information is present (e.g., Speier et al. (1999)). It is consistent with the results of many empirical and experimental studies. For example, Drake et al. (2017) shows that non-professional coverage of corporate news may decrease the price responsiveness to the news. In an experimental study, Elliott et al. (2015) designs a setting in which participants trade based on disclosure of earnings. The authors find that price is most efficient when the transitory components of earnings are excluded from the earnings metrics.

My model generates novel and testable empirical implications. Disclosure theory suggests that firms face the trade-off between the benefit and the cost of disclosure. In my model, the disclosure cost is information overload. My model thus predicts that firms are more likely to disclose details, the higher investors' information processing capacity is, and the lower the strength of complementarity among investors' decisions. I will discuss the details of these predictions in Section 6 of the paper.

In the economics literature, the mutual entropy approach to model limited attention is formalized by Sims (2003) and is applied by studies such as Mackowiak and Wiederholt (2009) and Myatt and Wallace (2011). My paper applies Sims' approach to the information overload problem. I also follow the approach of Mackowiak and Wiederholt (2009) in assuming that paying attention to details are independent processes. Hirshleifer and Teoh (2003) also adopts the limited attention perspective to study problems in financial reporting. Their study relies on exogenous assumptions of investors' attention allocation strategies. In contrast, my study endogenizes investors' attention allocation strategies.

My paper is related to the accounting literature on aggregation. Although aggregation typically leads to loss of information (Lev (1968)), studies in this area of the literature have established that aggregation may have benefits (see Arya and Glover (2013) for a survey). For example, aggregation may: convey information (Sunder (1997)), allow the cancellation of errors in individual items (e.g., Datar and Gupta (1994), Lim and Sunder (1991)), and discourage earnings management (Dye and Sridhar (2004)). My paper offers new insights on why a summary can dominate details. In addition, unlike my paper, these studies do not focus on why details are sometimes not reported together with a summary.

My paper is also related to the literature on beauty contests and contributes to the debate on whether more public disclosure is beneficial. In contrast to models such as Morris and Shin (2002), where investors are assumed to be endowed with private signals, I show that in environments with strategic complementarity, more public disclosure can reduce welfare even if investors have no access to private information sources. In my model, the key friction that makes more public disclosure undesirable is investors' limited attention, which also differs from Morris and Shin (2002), in which the main mechanism is that investors put too much weight on public signals due to their coordination roles.

Finally, my paper is closely related to the literature on information acquisition in the stock market and beauty contests. The seminal work of Grossman and Stiglitz (1980) illustrates that costly information acquisition prevents stock prices from fully reflecting all the information. The limited attention perspective, however, is different from the information acquisition cost view in Grossman and Stiglitz (1980). This is because, even if stock prices reflect all the information, investors with limited attention still need to process the stock price. In the context of my model, because investors have free access to financial statements, their processing capacity constraint is more relevant than other types of information acquisition costs. Moreover, unlike traditional information acquisition cost functions, the entropy measure of information processing cost provides a consistent way to model costly information processing across different structures of information, because mutual entropy is well defined for any dimension of signals. In this literature, three closely related papers are Hellwig and Veldkamp (2009), Myatt and Wallace (2011), and Yang (2015), which consider how rationally inattentive players acquire information in beauty contest games. Their studies focus on characterizing investors' information acquisition strategies, whereas my paper focuses on the welfare consequence of increased disclosure of details. Other related works in this area include: Froot et al. (1992), which explains the herding behavior of investors by showing that a short horizon can cause investors to study and trade on irrelevant information, and Banerjee et al. (2018), which shows that lowering the cost of acquiring fundamental information may induce investors to learn more about the beliefs of others. My paper also differs from these studies by revealing a unique insight: even when details are precise measures of the fundamental, disclosing details in addition to a summary may cause coordination problems and may reduce investors' welfare.

The rest of the paper is organized as follows. Section 2 describes the model set-up and defines the equilibrium concept. Section 3 derives equilibrium and characterizes investors' attention-allocation strategies. Section 4 conducts the main analysis on the trade-off between a summary and details and on whether reporting details together with a summary is optimal. Section 5 discusses the empirical implications of the model. Section 6 provides two examples of complementarity and shows that similar results hold in these two examples. Section 7 extends the model in various dimensions and shows that similar results hold. Section 8 concludes.

1.2 Model set-up

As explained in the introduction, my analysis relies on two key elements: limited attention and complementarity among investors' actions. To capture these two elements, I study a model of reporting in the coordination game of Angeletos and Pavan (2004), with investors' limited information capacity as the main additional friction.

A continuum of risk-neutral investors, who are indexed by i and uniformly distributed on the $[0,1]$ interval, are considering their investment decisions on a firm that has access to a project. For each investor i, investing k_i will cost him $\frac{1}{2}k_i^2$ and generate a payoff of Ak_i. Thus, the utility of investor i is

$$u_i = Ak_i - \frac{1}{2}k_i^2, \tag{1.1}$$

where A is the return on the project. Following Angeletos and Pavan (2004), I introduce complementarity among investors' actions by assuming

$$A = (1 - r)\theta + rK, \tag{1.2}$$

where θ is the fundamental of the firm, $K = \int_i k_i di$ is the aggregate level of investment, and $r \in (0, 1)$ captures the strength of complementarity. In fact, risk neutrality implies the optimal k_i should satisfy

$$k_i = E[A] = (1 - r)E[\theta] + rE[K], \tag{1.3}$$

where $E[.]$ denotes expectation. Hence, investors care not only about the fundamental, but also about the aggregate investment. The relative importance of these two objectives is captured by r.

To model disaggregated information (details), I assume the firm's fundamental θ is the sum of two independent components, θ_x and θ_y. That is

$$\theta = \theta_x + \theta_y, \tag{1.4}$$

where $\theta_x \sim N(0, \sigma_x^2)$ and $\theta_y \sim N(0, \sigma_y^2)$. For example, if θ is a firm's total revenue, θ_x and θ_y could be the firm's domestic revenue and foreign revenue. To simplify the analysis and without biasing the main intuitions, I assume symmetry of the two details: $\sigma_x = \sigma_y = 1$.

The firm's objective is to maximize the welfare (aggregate utility) of all investors, which is $W = \int_0^1 u_i di$. The firm can provide public disclosure to help investors make better decisions. The structure of firms' public signals will be specified as the key element of the model in section 1.2.1.

Since coordination is socially beneficial, as in Angeletos and Pavan (2004), the firm's public disclosure can affect investors' welfare in two ways. First, public disclosure updates investors' beliefs about the fundamental θ and thus bring their investment closer to the fundamental. Second, as will be explained in the next subsection, limited attention implies that public signals will be confounded by private noise after being processed by investors. Therefore, following public disclosure, investors will find aligning their investment decisions harder. The total value of a public signal can thus be intuitively decomposed into two components: **information value**, defined as the extent to which the signal brings investors' investment closer to the fundamental, and **coordination value**, defined as the extent to which investors can coordinate their investment following the signal. In the extension of the model, I show that even if the welfare criterion is how well investors actions are fine-tuned to the fundamental, as in Morris and Shin (2002), my results still hold.

1.2.1 Reporting Regimes

I assume the firm can choose one from the following three reporting regimes.

In Regime S (summary), the firm's reporting system only reports a noisy summary of the firm's fundamental: $I_S = \{\theta + \gamma\}$, where $\gamma \sim N(0, \sigma_\gamma^2)$.

In Regime D (details), the firm's reporting system only reports details: $I_D = \{\theta_x, \theta_y\}$.

In Regime SD (summary and details), the firm's reporting system reports both the summary and details: $I_{SD} = \{\theta_x, \theta_y, \theta + \gamma\}$.

Therefore, based on the above definitions, a comparison of investors' welfare in S and D should reveal why summarizing details is important. A comparison of investors' welfare in regimes S and SD should provide insight for why attaching details to the summary is not always desirable.

A crucial assumption I make is that the summary is a noisy measure of the fundamental, whereas details are precise. This assumption reflects the fact that loss of information is often unavoidable in the aggregation or summarization process. For example, details in a firm's 10-K usually contain more information than summary measures such as the firm's net income. Another example is that the abstract of a paper often contains less information than the introduction. Hence, the assumption that the summary is contaminated by some noise term seems a reasonable description of its nature.

However, as soon will be explained in the following sections, due to limited attention and the fact that the summary is easier to process than details, loss of information in the aggregation process does not necessarily imply that the summary has lower information value than details. For the rest of the paper, I assume $\sigma_\gamma^2 < 2$ to focus on the more interesting case where a trade-off between the summary and details exists.

Assumption 1. $\sigma_\gamma^2 < 2$.

1.2.2 Investors' limited attention

Investors make investment decisions after the release of the firm's report. However, due to limited attention, investors are not able to fully process the information available in the firm's report.

Following Sims (2003), I model limited attention by assuming that investors behave as if public signals are observed through finite capacity channels. Here, I briefly explain Sims' method. In information theory, the entropy of a continuous random variable X is a measure of the uncertainty associated with X, defined as $H(X) = -E[\log(f(X))]$, where $f(.)$ is the density of X. Suppose another random variable Y is correlated with X, so that we know more about X through observation of Y. The mutual entropy between X and Y, which is a measure of the reduction of uncertainty in X after observing Y, is defined as $I(X, Y) = H(X) - H(X|Y)$. Sims' central point is that limited attention can be captured by imposing the constraint that the mutual entropy $I(X, Y)$ must be less than investors' information processing capacity, denoted by κ. Such a constraint reflects the fact that higher processing capacity is required to learn more about X. Furthermore, Sims shows that when X is normal, limited information processing capacity implies investors behave as if X is observed with a normally distributed noise. When this is the case, the mutual entropy constraint $I(X,Y) \leq \kappa$ simplifies to

$$\frac{\text{Var}(X)}{\text{Var}(X|Y)} \leq e^{2\kappa}. \tag{1.5}$$

The inequality constraint in (5) will bind, implying that

$$\text{Var}(X|Y) = \frac{\text{Var}(X)}{e^{2\kappa}}. \tag{1.6}$$

The right hand side of (1.6) is convex in κ. Therefore, if our objective is to choose $Y|X$ to maximize $-\text{Var}(X|Y)$, or equivalently, choose $Y|X$ to maximize $-E[(Y-X)^2]$, then the marginal value of κ is **decreasing**. The following lemma summarizes the discussion so far.

Lemma 1. *Suppose X is normally distributed and we want to choose $Y|X$ to maximize $-\text{Var}(X|Y)$, subject to $I(X,Y) \leq \kappa$. We have $\frac{\partial^2 - \text{Var}(X|Y^*)}{\partial \kappa^2} < 0$, where Y^* is the optimal signal.*

In other words, there is decreasing marginal return from investing in information processing capacity. This Lemma will be useful later.

Similar to Myatt and Wallace (2011), in my model, I assume investors process the firm's public disclosure independently, which means that public signals will be confounded by private noise after being processed by the investors. Specifically, the smaller the variance of the private noise, the higher the reduction in uncertainty, and the higher the mutual entropy between the underlying public signal and the private signal. To simplify the algebra, I define $C \equiv e^{\kappa}$.

I now describe investors' optimization problem in the three reporting regimes when the constraint on mutual entropy is imposed.

Regime S

In regime S, investors can only pay attention to the summary, which means investor i will obtain a noisy signal z_i about $\theta + \gamma$. Specifically, $z_i = \theta + \gamma + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma^2)$ and is independent across i. As a result, before z_i is realized, investor i's attention-allocation problem is to choose σ^2 to maximize his expected utility, subject to a constraint on the mutual entropy between z_i and $\theta + \gamma$, which I denote as $I(z_i, \theta + \gamma)$:

$$\max_{\sigma^2} E[u_i]$$
$$\text{subject to } I(z_i, \theta + \gamma) \leq \kappa \Leftrightarrow \frac{\frac{1}{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2} + \frac{1}{\sigma^2}}{\frac{1}{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2}} \leq C^2.$$

Note that in regime S, an investor can at most process $\theta + \gamma$ perfectly. In other words, I assume that the noise γ generated in the firm's aggregation process cannot be eliminated through paying attention. Such an assumption echoes Myatt and Wallace (2011), which assumes a receiver cannot reduce the noise generated by a sender through learning.

To sum up, in regime S, investors can only pay attention to the summary and then make investment decisions based on the signals z_i.

Regime D

In regime D, investors have to process the two details independently. This means investor i will receive one signal x_i about detail θ_x, and another signal y_i about detail θ_y:

$$x_i = \theta_x + \epsilon_{1i},$$
$$y_i = \theta_y + \epsilon_{2i},$$

where $\epsilon_{1i} \sim N(0, \sigma_1^2)$ and $\epsilon_{2i} \sim N(0, \sigma_2^2)$ are independent across i.

As a result, before x_i and y_i are realized, investor i's attention-allocation problem is to choose σ_1^2 and σ_2^2 to maximize his expected utility, subject to a constraint on the mutual entropy between $\{x_i, y_i\}$ and $\{\theta_x, \theta_y\}$, which I denote as $I(\{x_i, y_i\}, \{\theta_x, \theta_y\})$. Investor i's attention-allocation problem is thus as follows:

$$\max_{\sigma_1^2, \sigma_2^2} E[u_i]$$

$$\text{subject to } I(\{x_i, y_i\}, \{\theta_x, \theta_y\}) \leq \kappa \Leftrightarrow \frac{(\frac{1}{\sigma_x^2} + \frac{1}{\sigma_1^2})(\frac{1}{\sigma_y^2} + \frac{1}{\sigma_2^2})}{\frac{1}{\sigma_x^2}\frac{1}{\sigma_y^2}} \leq C^2.$$

To sum up, I assume that in regime D, an investor i first allocates attention to each detail separately and then makes an investment decision based on the signals x_i and y_i. Such an assumption is similar to Mackowiak and Wiederholt (2009), which considers the optimal attention-allocation problem between aggregate and idiosyncratic conditions.

Regime SD

In regime SD, investors have more flexibility in acquiring information. For the purpose of tractability, and without biasing the main intuitions[1], I make the following assumption:

Assumption 2. *In regime SD, investors can choose one from the following two options:*

*The first option (**option SU**) is to pay attention to the summary, $\theta + \gamma$, which means investor i's attention-allocation problem is identical to the problem in regime S.*

*The second option (**option DE**) is to pay attention to the details, which means investor i's attention-allocation problem is identical to the problem in regime D.*

1.2.3 Timeline

The timeline of the model is as follows:

At $t = 0$, the firm chooses a reporting system (S, D, or SD) to maximize social welfare,

1. The reason will be explained in Section 4.2.

defined as $W = \int_0^1 u_i di$.

At $t = 1$, each investor chooses how much attention to allocate to the summary or the details. At $t = 2$, investors receive their private signals and make their investment decisions. The outcome of the project is then realized.

1.2.4 Definition of equilibrium

I focus on a Symmetric Perfect Bayesian Equilibrium, which is defined as follows:

(i) For any investor i, he first optimally chooses the precision of his private signals to maximize his expected utility, given other investors' investment strategies. He then optimally chooses his investment k_i to maximize his expected utility, given other investors' investment strategies and the private signal he receives.

(ii) All investors choose the same precision of signals.

1.3 Equilibrium

In this section, I derive investors' equilibrium attention-allocation and investment strategies in each of the three reporting regimes, to prepare for the study of the main research questions in the next section.

Because of symmetry and Assumption 2, only the following four strategy profiles can be equilibrium:

Definition 1. *SU: the strategy profile in which investors devote all attention to the summary,* $\theta + \gamma$.

DE_X: the strategy profile in which investors devote all attention to the first detail, θ_x.

DE_Y: the strategy profile in which investors devote all attention to the second detail, θ_y.

DE_BOTH: the strategy profile in which all investors divide attention evenly between the two details, θ_x *and* θ_y.

In addition, investors' actions are linear in their signals, mainly because their signals are normal. In fact, as shown by Morris and Shin (2002), linear strategy is the only possible strategy in such games.

Lemma 2. *Investors' investment strategies are linear in their signals. In particular, in SU, investor i chooses $k_i = b_z z_i$, where b_z is a constant. In DE_BOTH, investor i chooses $k_i = b_x x_i + b_y y_i$, where $b_x = b_y$ are constants. In DE_X, investor i chooses $k_i = b x_i$, where b is a constant.*

See the formal proof in Morris and Shin (2002).

1.3.1 Regime S

Lemma 3 summarizes investors' attention-allocation strategies in regime S.

Lemma 3. *In regime S, each investor will utilize his information capacity to minimize σ^2. Specifically, investors will all choose $\sigma^2 = \frac{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2}{C^2-1}$.*

The intuition for Lemma 3 is simple: investors will devote all their attention to the summary, because it is the only signal to pay attention to.

1.3.2 Regime D

To characterize investors' equilibrium strategy in D, I start with the simple case $r = 0$, in which investors only care about the information value of signals (i.e., investors only care about whether investment decisions are fine-tuned to the fundamental).

Lemma 4. *When $r = 0$, DE_BOTH is the unique equilibrium in regime D.*

The intuition is as follows. When $r = 0$, investors only care about predicting the fundamental θ. Since the two details are equally important, it is thus natural to expect that investors will divide their attention evenly between the two details in order to get a best estimation of the fundamental.

I now proceed to the general case where $r \in (0, 1)$.

Proposition 1. *There exists a unique $r_0 \in (0,1)$, such that:*

1) if $r \in (0, r_0)$, regime D has a unique equilibrium: DE_BOTH.

2) if $r \in (r_0, 1)$, regime D has three equilibria: DE_X, DE_Y, DE_BOTH.

The intuition for Proposition 1 is as follows.

First, DE_BOTH is an equilibrium for any $r \in (0,1)$. The reason is that within this strategy profile, the linearity of investors' strategy implies that $K = \int_i b_x x_i + b_y y_i di = b_x\theta_x + b_y\theta_y = b_x\theta$. Hence, predicting K is equivalent to predicting the fundamental θ. As a result, based on Lemma 4, dividing attention evenly between the details is investor i's best response.

Second, consider the strategy profile DE_X. If investor i deviates and pays slightly more attention to θ_y, he can more easily predict the fundamental. However, he finds predicting K harder because $K = \int_i bx_i di = b\theta_x$. Therefore, DE_X and DE_Y are equilibria, when the strength of complementarity is sufficiently high.

1.3.3 Regime SD

The central message of this section is that multiple equilibria exist in regime SD: DE_X, DE_Y, DE_BOTH, SU can all be equilibria. To understand the intuition for this result, it is useful to start with the benchmark case in which $r = 0$, when investors only care about the information value of signals.

Proposition 2. *Under the benchmark case in which $r = 0$, if $C < C^* \equiv \frac{2}{\sigma_\gamma^2}$, then all investors choose option SU. Otherwise, if $C > C^* \equiv \frac{2}{\sigma_\gamma^2}$, then all investors choose option DE and will divide attention evenly between the two details.*

To understand Proposition 2, note that we can compare the information value of signals by comparing the magnitude of noise in signals. The larger the noise, the lower the information value. Hence, we can compare the information value of the summary and details by comparing the magnitude of noise in z_i and $x_i + y_i$. This is because, according to Lemma

2, investor i bases his action on z_i if he focuses on summary, and on $x_i + y_i$ if he focuses on details. We know that $z_i = \theta + \gamma + \epsilon_i$ and $x_i + y_i = \theta + \epsilon_{1i} + \epsilon_{2i}$. Therefore, the summary has higher information value than details, or equivalently, investors will prefer option SU over option DE, if and only if

$$\begin{aligned} & \mathrm{Var}(\gamma + \epsilon_i) \le \mathrm{Var}(\epsilon_{1i} + \epsilon_{2i}) \\ \Leftrightarrow & \sigma_\gamma^2 + \frac{2 + \sigma_\gamma^2}{C^2 - 1} \le \frac{2}{C - 1} \\ \Leftrightarrow & C \le \frac{2}{\sigma_\gamma^2}. \end{aligned} \tag{1.7}$$

From (1.7), we can clearly see the fundamental trade-off between summary and details. First, details are harder to process than the summary, because each detail needs to be processed separately. This follows from the fact that the noise in z_i is discounted by $\frac{1}{C^2-1}$, whereas the noise in $x_i + y_i$ is only discounted by $\frac{1}{C-1}$. Furthermore, because the marginal value of C is decreasing (Lemma 1), details should be less subject to this problem when investors' information processing capacity is larger. On the other hand, z_i has an extra noise γ, because aggregation inevitably leads to a loss of information. Therefore, the trade-off between the summary and details is in fact a trade-off between signal precision and processing difficulty. Hence, the summary has higher information value than details, or equivalently, investors will prefer option SU over option DE, if and only if investors' information processing capacity is sufficiently small.

Proposition 3 describes investors' attention allocation strategy in the general case in which $r \in (0, 1)$.

Proposition 3. *If $C < C^*$, there exists a unique $r_1 \in (0, 1)$ such that*

1) if $r \in (0, r_1)$, the game has a unique equilibrium: SU.

2) if $r \in (r_1, 1)$, the game has three equilibria: SU, DE_X, DE_Y.

If $C > C^*$, there exists a unique $r_2 \in (0, r_0)$, such that:

1) if $r \in (0, r_2)$, the game has a unique equilibrium: DE_BOTH.

2) if $r \in (r_2, r_0)$, the game has two equilibria: SU and DE_BOTH.

3) if $r \in (r_0, 1)$, the game has four equilibria: SU, DE_X, DE_Y, DE_BOTH.

r_0 is define in Proposition 1.

The intuition for Proposition 3 is as follows.

Consider the case in which $C < C^*$, when the summary has higher information value than details.

First, it is obvious that SU is still an equilibrium for any $r \in (0, 1)$, because within this strategy profile, $K = b(\theta + \gamma)$. Thus, any investor i will not deviate, because sticking to option SU allows him to better predict both the fundamental θ and the aggregate investment K.

Second, DE_BOTH will never be an equilibrium, because within this strategy profile, the aggregate investment takes the form $K = b_x\theta_x + b_y\theta_y$, where $b_x = b_y$ by symmetry. Hence, $K = b_x\theta$. As a result, deviating to option SU will strictly increase player i's utility because option SU allows him to better predict both the fundamental and the aggregate investment.

Furthermore, consider the strategy profile DE_X. If investor i pays slightly more attention to θ_y or devotes all of his attention to the summary, he can more easily predict the fundamental. However, he finds it harder to predict K because $K = \int_i bx_i di = b\theta_x$. Therefore, DE_X and DE_Y are equilibria, when the strength of complementarity is sufficiently high.

Consider then the case in which $C > C^*$, when details have higher information value than the summary.

First, DE_BOTH is still obviously an equilibrium for any $r \in (0, 1)$, because within this strategy profile, $K = b_x\theta_x + b_y\theta_y = b_x\theta$. As a result, deviating to option SU will strictly decrease investor i's utility because option SU is worse at predicting both the fundamental and the aggregate investment.

Second, consider the strategy profile SU. If investor i deviates to option DE and divides attention evenly between the two details, he can more easily predict the fundamental. However, he may find it harder to predict K because $K = b(\theta + \gamma)$, which is better predicted if he devotes all attention to the summary. Hence, when the strength of complementarity is sufficiently high, another equilibrium exists in which all investors choose option SU.

Furthermore, DE_X and DE_Y could be equilibria for the same reason as in the previous case.

1.4 Main results

In this section, I study the main research questions: First, can the limited attention perspective rationalize the need for summarizing details? Second, under the limited attention perspective, should details always be attached to the summary? The first question can be investigated through comparing investors' welfare in regime S and regime D. The second question can be investigated through comparing investors' welfare in regime S and regime SD.

1.4.1 Comparison of regime S and regime D

In this subsection, I first compare the information value and coordination value of the summary and details. I then analyze the trade-off between the summary and details by comparing investors' welfare in regime S and regime D.

Corollary 1. *Consider the case $r = 0$. S is preferred to D if and only if $C \leq C^* \equiv \frac{2}{\sigma_\gamma^2}$.*

When $r = 0$, investors only care about the information value of signals. Proposition 2 shows that the summary has higher information value than details if and only if $C \leq \frac{2}{\sigma_\gamma^2}$. Hence, Corollary 1 is directly implied by Proposition 2.

Perhaps a good example to illustrate the trade-off between summary and details is Yelp. What Yelp does is providing a score for each restaurant. For example, restaurant XYZ gets a score of 4.5, because customer A gives it a score of 5, and customer B gives it a score of 4. The interesting fact is that these customers also leave comments in Yelp. Customers' comments usually represent their true thoughts but can NOT be aggregated without loss of information. In fact, to obtain an aggregated score of 4.5, each customer must first convert his preference to a number, which is a noisy process. The trade-off implied by Corollary 1 is then very natural. On one hand, looking at the average score 4.5 allows Yelp's users to process information more efficiently, but the score 4.5 itself is noisy. On the other hand, users can also look at the comments, which require more capacity to process. However, if the users have unlimited capacity, they can understand the quality of the restaurant better by looking at the comments.

Similar logic applies to financial reporting: Summary measures such as net income or EPS are easier to process. However, to obtain these summary measures, loss of information must occur in the aggregation or summarization process.

In addition, as the summary becomes noisier, it is likely that the summary would have higher information value than details. This result follows from the Corollary below:

Corollary 2. $\frac{\partial C^*}{\partial \sigma_\gamma} < 0$.

How does the coordination value of the summary and compare with the coordination value of details? Intuitively, private noise in investors' signals prevent investors from perfectly aligning their investment decisions. In other words, the smaller the variance of private noise (normalized by the variance of the public part of investors' signals), the higher the coordination value. We can thus compare the coordination value of the summary and details by comparing $\frac{\mathrm{Var}(\epsilon_i)}{\mathrm{Var}(\theta+\gamma)}$ with $\frac{\mathrm{Var}(\epsilon_{1i}+\epsilon_{2i})}{\mathrm{Var}(\theta)}$. The following lemma shows that the summary always has higher coordination value than details.

Lemma 5. $\frac{\mathrm{Var}(\epsilon_{1i}+\epsilon_{2i})}{\mathrm{Var}(\theta)} = \frac{1}{C-1} > \frac{\mathrm{Var}(\epsilon_i)}{\mathrm{Var}(\theta+\gamma)} = \frac{1}{C^2-1}$.

Lemma 5 holds, because when attention is divided among details, each detail is processed less thoroughly, compared to the case where all attention is devoted to the summary. Therefore, investors achieve less synergy among their investment decisions if they focus on details rather than the summary. As a result, details have lower coordination value than the summary, regardless of investors' information processing capacity.

Based on the results so far, when $C \leq C^*$, details have both lower coordination value and lower information value, and therefore S should always be preferred to D. When $C > C^*$, details have lower coordination value but higher information value, and therefore S should be preferred to D if and only if coordination is sufficiently important. These conjectures are formally verified in Proposition 4.

Proposition 4. *Consider the general case $r \in (0,1)$. If $C \leq C^* \equiv \frac{2}{\sigma_\gamma^2}$, S is always preferred to D. If $C > C^* \equiv \frac{2}{\sigma_\gamma^2}$, there exists a unique $r^* \in (r_2, r_0)$, such that S is preferred to D if and only if $r > r^*$. r_0 and r_2 are defined in Proposition 1 and Proposition 3.*

1.4.2 Comparison of regime S and regime SD

In this subsection, I study whether attaching details to the summary is always optimal by comparing investors' welfare in regime S and regime SD. I start with a useful lemma to simplify the analysis.

Lemma 6. *Assumption 1 implies DE_X (or DE_Y) is always less efficient than SU.*

The intuition for this lemma can be understood in two steps. First, in both SU and DE_X, investors devote all their attention to one single signal, which means investors achieve the same precision in predicting the aggregate investment K. The two strategy profiles are thus equivalent in facilitating coordination among investors. Second, both $\theta + \gamma$ and θ_x are noisy measures of fundamental θ. However, because $\sigma_\gamma^2 < 2$, the former is less noisy. Hence, investors' prediction of θ is better in SU. It is then natural to expect that social welfare should be higher in SU.

In addition, because multiple equilibria exist in regime SD, I adopt the following definition when comparing the two regimes.

Definition 2. *$S \succeq SD$ if the social welfare in S is greater than or equal to the social welfare in every possible equilibrium outcome in SD. Conversely, $S \preceq SD$ if the social welfare in S is smaller than or equal to the social welfare in every possible equilibrium outcome in SD.*

Now consider the comparison between the two regimes.

When $C < C^*$ and $r < r_1$, based on Proposition 3, because SU is the only possible equilibrium in regime SD, the two regimes are equivalent.

When $C < C^*$ and $r > r_1$, DE_X and DE_Y could also be equilibria. However, according to Lemma 6, these two equilibria are always less efficient than SU. Hence, SD is preferred to S in this case.

Comparison of the two regimes in the case $C > C^*$ boils down to the comparison between DE_BOTH and SU, because equilibrium DE_X and DE_Y, which only exist when $r > r_0$, are always less efficient than equilibrium SU, which also exists when $r > r_0$.

The trade-off between DE_BOTH and SU is the same as in Proposition 4. Since details have lower coordination value and higher information value, SU should be preferred to DE_BOTH if and only if coordination is sufficiently important.

Proposition 5 is thus a natural consequence of the discussion above.

Proposition 5. *If $C < C^*$ and $r < r_1$, the two regimes are equivalent.*
If $C < C^*$ and $r > r_1$, then $S \succeq SD$.
If $C > C^*$, then $S \succeq SD$ if and only if $r > r^*$.
r_1 and r^* are defined in Proposition 3 and Proposition 4.

Proposition 5 presents a perhaps surprising result: even when investors' information processing capacity is sufficiently large such that details have higher information value than the summary, attaching details to the summary may still reduce investors' welfare. Intuitively, this result can be understood in the following way. First, as explained in the previous section, if details are reported together with the summary, investors may jointly focus on the details when their information processing capacity is large. As a result, because details have lower coordination value and higher information value, attaching details to the summary reduces investors' welfare if and only if the strength of complementarity is sufficiently high (i.e., coordination is sufficiently important). Furthermore, this intuition likely carries over to the case in which investors can choose to pay attention to both the summary and the details in regime SD. This is because even if there exists an equilibrium in which investors pay attention to both the summary and the details, such an equilibrium still has higher coordination value and lower information value than the equilibrium in which investors devote all their attention to the summary.

Proposition 5 provides a rationale for why accounting information is highly aggregated. It also rationalizes the notion of information overload, defined as a reduction in decision quality when too much information is present (e.g., Speier et al. (1999)).

How does the firm's willingness to disclose details change as investors' information capacity becomes larger? Intuitively, as investors' information processing capacity becomes larger, investors can better process the details, and providing details in addition to the summary will become less likely to reduce investors' decision quality. The following comparative statics exercise verifies this intuition.

Corollary 3. $\frac{\partial r^*}{\partial C} > 0$.

To briefly summarize the main results, I find that although the summary is less precise than details, the firm may prefer the summary to details because details are harder to process and impede coordination. Moreover, I find that reporting details in addition to the summary may reduce the quality of investors' decisions, when C is small, or when C is large and r is large. My results thus shed light on a potential cost of increased disclosure of details: it may exacerbate the coordination problem among investors when their information-processing capacity is small.

1.5 Empirical implications

Following Proposition 5 in the main model, the information overloaded problem caused by disclosure of details should be less severe when investors have high capacity, or when the strength of complementarity among investors' decisions is low. If we view information overload as the cost of more disclosure, we can generalize the main results to settings of voluntary disclosure, in which firms weigh the benefit of disclosure against the cost of disclosure. We therefore have the following two main implications.

Prediction 1: Firms will increase their disclosure of details when investors' information processing capacity (cost) become higher (lower).

The findings of Blankespoor (2019) support this prediction. In the paper, the author looks at firms' response to the adoption of XBRL, which lowers investors' information processing cost. The author finds that firms increased their footnote disclosure upon the adoption. This is therefore direct evidence showing that firms do take into account investors' processing cost and are more willing to disclose details when the cost of processing details goes down.

This prediction can also be tested by finding other proxies for investors' information processing capacity. For example, if institutional investors have higher information processing capacity than retail investors, the prediction is then firms with higher fraction of institutional investors will more details about their performances.

Prediction 2: Firms whose investors exhibit higher complementarity in decisions will disclose less details.

This cross-sectional prediction can be tested on financial institutions such as banks, in which coordination problem exists among banks' creditors/depositors. Following Chen et al. (2010) and Goldstein and Sapra (2014), the strength of complementarity among a bank's creditors can be measured by the illiquidity of the bank's asset is, as characterized by the fraction of level 2 and level 3 assets. Following Goldstein and Sapra (2014), the higher strength of complementarity can also be captured by a less concentrated base of creditors. The prediction is therefore banks with more illiquid assets or with less concentrated creditors tend to disclose less details about their performance.

1.6 Examples for strategic complementarity

In this section, I study two examples (an IPO problem and a bank run problem) to provide justification for why investors' actions can be complements. I show that in these two examples, when firms and banks are deciding what information to disclose, they may indeed want to suppress details, for the same reasons as in Proposition 5.

1.6.1 An IPO game

Here I consider an IPO game similar to the one in Frankel et al. (2019).

A firm is raising capital from a continuum of investors (uniformly distributed on the $[0, 1]$ interval). Specifically, the firm sells s units of its shares in return for t units of capital. IPO price is thus $p = \frac{t}{s}$ per share.

The IPO works in the following way. If fraction K of investors subscribe to the IPO, firm value will be $\theta + rK$, where $r > 0$ captures the strength of complementarity. Each subscribing investor contributes her capital, in return for $\frac{1}{p}$ share of the firm. Thus, investor i's net gain from subscribing is

$$u_i = \frac{1}{p}(\theta + rK) - 1 \tag{1.8}$$

How will investors allocate their attention and make their investment decisions? Based on (1.8), an investor i will subscribe (choose $k_i = 1$) if and only if $E[u_i] \geq 0$, or equivalently:

$$k_i = \begin{cases} 1, \text{ if } E[\theta] + rE[K] \geq p \\ 0, \text{ otherwise} \end{cases} \tag{1.9}$$

From (1.9), we can clearly see how higher order beliefs play a role in this IPO game, because each investor will not only form expectations about θ, the fundamental, but also will form expectations about K, the aggregate subscribing fraction. Furthermore, fixing the fundamental θ, the higher the aggregate subscription K is, the more likely investor i will subscribe. Investors' decisions are therefore complements, with r capturing the strength of complementarity.

Investors' strategy in regime S

In regime S, due to limited attention, investor i will receive $z_i = \theta + \gamma + \epsilon_i$, which is a private version of the public signal $\theta + \gamma$. Standard results in the global game literature imply that investors will follow a threshold strategy: investor i will subscribe if and only if his signal z_i is above some threshold q_S. Then, using the law of large numbers, we know that $K = \Pr(z_j \geq q_S|z_i)$. In other words, conditional on his private signal z_i, investor i understands that the aggregate subscription will depend on whether other investors' signals are above the threshold.

We now solve for the threshold q_S. A creditor receiving $z_i = q_S$ should be indifferent between investing and not investing. This implies that $E(\theta|z_i = q_S) + r\Pr(z_j \geq q_S|z_i = q_S) = 1$. By symmetry, we know that $\Pr(z_j \geq q_S|z_i = q_S) = \frac{1}{2}$. Hence, we must have $E(\theta|z_i = q_S) = 1 - \frac{r}{2}$, or equivalently, the following proposition:

Proposition 6. *In regime S, each creditor withdraws if and only if $z_i \leq q_S$, where* $q_S \equiv \frac{1-\frac{r}{2}}{\frac{2(C^2-1)}{C^2(\sigma_\gamma^2+2)}}$.

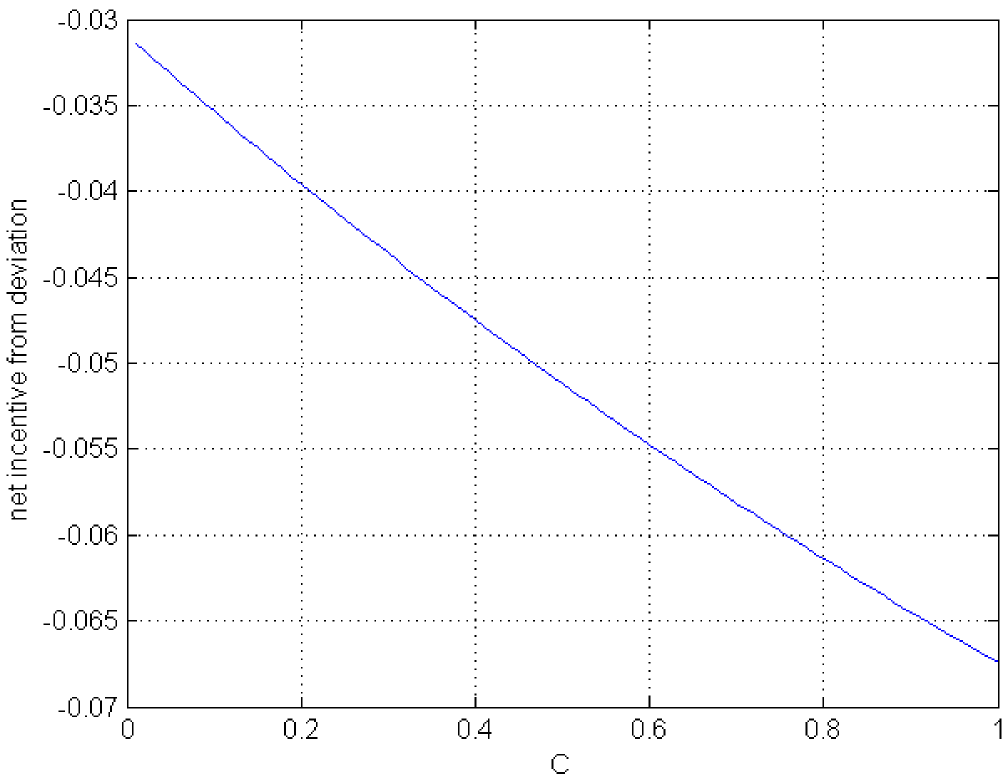
Investors' strategy in regime SD

How will investors allocate their attention in regime SD? The following conjecture echoes the results in the basic model.

Conjecture: If C is sufficiently large, there exists an equilibrium where all investors choose option DE and will divide attention evenly between the two details. Furthermore, in this equilibrium, investor i chooses $k_i = 0$ if $x_i + y_i < q_{SD}$, and $k_i = 1$ if $x_i + y_i \geq_{SD}$, where $q_{SD} \equiv \frac{1-\frac{r}{2}}{1-\frac{1}{C}}$.

The conjecture is verified numerically. In Figure 1.1 below, I plot investor i's incentive of deviating to option SU as a function of C, when $\sigma_\gamma^2 = 1.5$, $r = 1$, and $p = 1$. The plot indeed shows that the larger C is, the lower the incentive to deviate. Hence, when C is sufficiently large, we should expect there exists an equilibrium where all investors choose option DE.

Figure 1.1: Deviation incentive as a function of C, when $\sigma_\gamma^2 = 1.5$, $r = 1$, and $p = 1$.



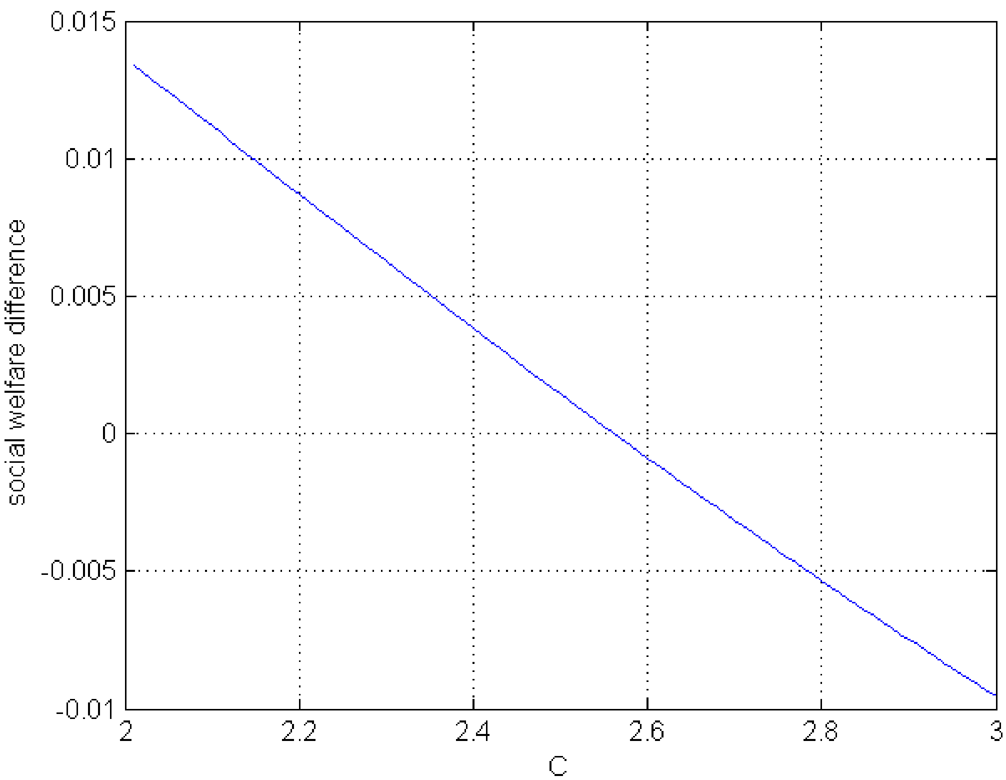
Comparison between the two regimes

The trade-off between regime S and SD is similar to the basic model. Specifically, details allow investors to better predict θ, the fundamental, while summary allows investors to better coordinate their investment decisions. Furthermore, as C increases, details can be more easily processed, and it should become more likely that regime SD dominates regime S. Using W_{SU} and W_{DE} to denote investors' welfare in equilibrium SU and DE, we thus have the following conjecture.

Conjecture 2: $W_{SU} - W_{DE}$ decreases in C.

In Figure 1.2 below, I plot the welfare difference between $W_{DE} - W_{SU}$ as a function of C. the plot indeed shows that the larger C, the smaller $W_{SU} - W_{DE}$. Furthermore, $W_{SU} - W_{DE}$ changes sign around $C = 2.35$, which shows that summary dominates summary plus details when C is sufficiently small.

Figure 1.2: $W_{SU} - W_{DE}$ as a function of C, when $\sigma_\gamma^2 = 1.5$, $r = 1$, and $p = 1$.



1.6.2 Bank run

I study a bank run problem similar to Morris and Shin (2000) to further illustrate the idea that investors' actions can be complements. I show that if banks worry about runs, they may suppress details when disclosing information about fundamentals to their creditors.

Consider a risk-neutral economy with no discounting, three dates ($t = 0, 1, 2$), a bank, and a continuum $[0, 1]$ of creditors. The bank has access to a long-term project that transforms 1 unit of deposit to θ units on $t = 2$.

However, the project is illiquid in the sense that if fraction l of the creditors withdraw on $t = 1$, the return of the project on $t = 2$ will decrease to $\theta - \delta l$, where $\delta > 0$ captures the intensity of strategic complementarity among creditors. If a creditor withdraw on $t = 1$, he is guaranteed to get 1 unit back. Hence, using u_i to denote creditor i's utility, we have

$$u_i = \begin{cases} 1, \text{ if early withdraw} \\ \theta - \delta l, \text{ otherwise} \end{cases} \tag{1.10}$$

The information structure and reporting regimes are identical to the basic model.

How will creditors allocate their attention and make their withdrawing decisions? Based on (1.10), creditor i will stay if and only if $E[\theta - \delta l] \geq 1$, or equivalently:

$$l_i = \begin{cases} 0, \text{ if } E[\theta] - \delta E[l] \geq 1 \\ 1, \text{ otherwise} \end{cases} \tag{1.11}$$

From (1.11), we can clearly see how higher order beliefs play a role in this bank run setting, because each creditor will not only form expectations about θ, the fundamental, but also will form expectations about l, the aggregate withdrawing fraction. Furthermore, the higher l is, the more likely investor i will withdraw (choosing $l_i = 1$). This shows that creditors' withdrawing decisions are complements.

Creditors' strategy and payoff in regime S

In regime S, due to limited attention, creditor i will receive $z_i = \theta + \gamma + \epsilon_i$, which is a private version of the public signal $\theta + \gamma$. Standard results in the global game literature imply that creditors will follow a threshold strategy: creditor i will stay if and only if his signal z_i is above some threshold q_S. Then, using the law of large numbers, we know that $l = \Pr(z_j < q_S | z_i)$. In other words, conditional on his private signal z_i, investor i understands that the aggregate withdrawal will depend on whether other investors' signals are below the threshold.

We now solve for the threshold q_S. A creditor receiving $z_i = q_S$ should be indifferent between investing and not investing. This implies that $E(\theta | z_i = q_S) - \delta \Pr(z_j < q_S | z_i =$

$q_S) = 1$. By symmetry, we know that $\Pr(z_j < q_S | z_i = q_S) = \frac{1}{2}$. Hence, we must have $E(\theta | z_i = q_S) = 1 + \frac{\delta}{2}$, or equivalently, the following proposition:

Proposition 7. *In regime S, each creditor withdraws if and only if $z_i \leq q_S$, where $q_S \equiv \frac{1+\frac{\delta}{2}}{\frac{2(C^2-1)}{C^2(\sigma_\gamma^2+2)}}$.*

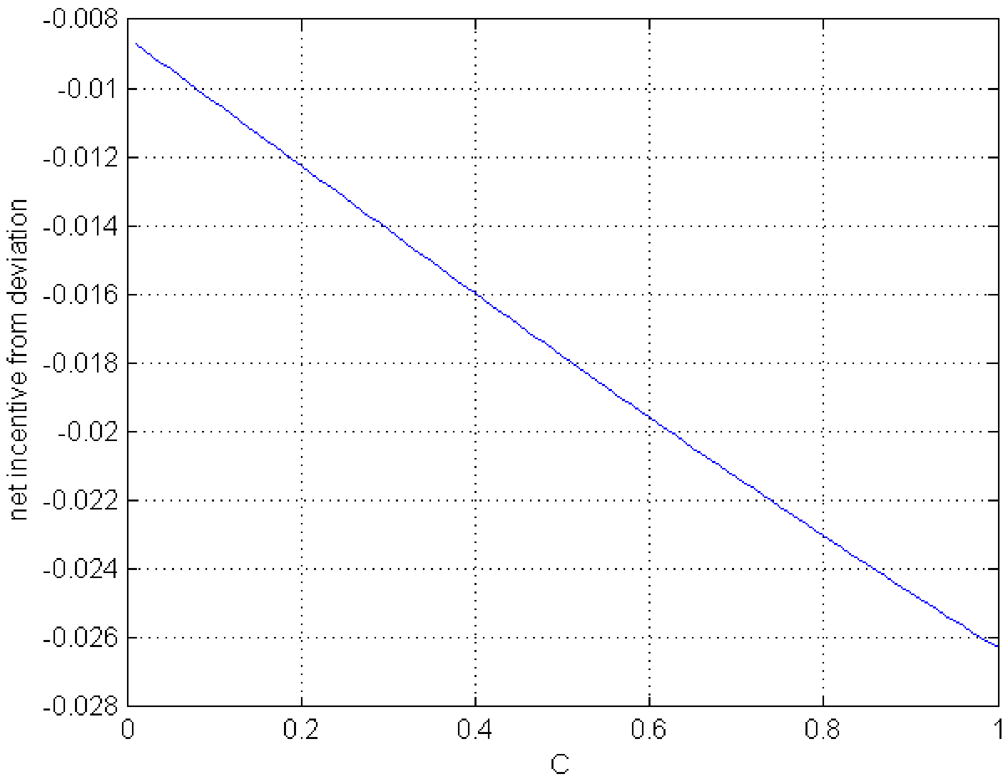
Creditors' strategy and payoff in regime SD

In regime SD, for the same reason as in the basic model, when C is sufficiently large, there should exist an equilibrium in which all creditors choose option DE. This is formally stated in the following conjecture.

Conjecture: If C is sufficiently large, there exists an equilibrium where all investors choose option DE and will divide attention evenly between the two details. Furthermore, in this equilibrium, creditor i withdraws if and only if $x_i + y_i \leq q_{SD}$, where $q_{SD} \equiv \frac{1+\frac{\delta}{2}}{1-\frac{1}{C}}$.

The conjecture is verified numerically. In Figure 1.3 below, I plot investor i's incentive of deviating to option SU as a function of C, when $\sigma_\gamma^2 = 1.5$ and $\delta = 1$. The plot indeed shows that the larger C is, the lower the incentive to deviate. Hence, when C is sufficiently large, we should expect there exists an equilibrium where all investors choose option DE.

Figure 1.3: Deviation incentive as a function of C, when $\sigma_\gamma^2 = 1.5$ and $\delta = 1$.



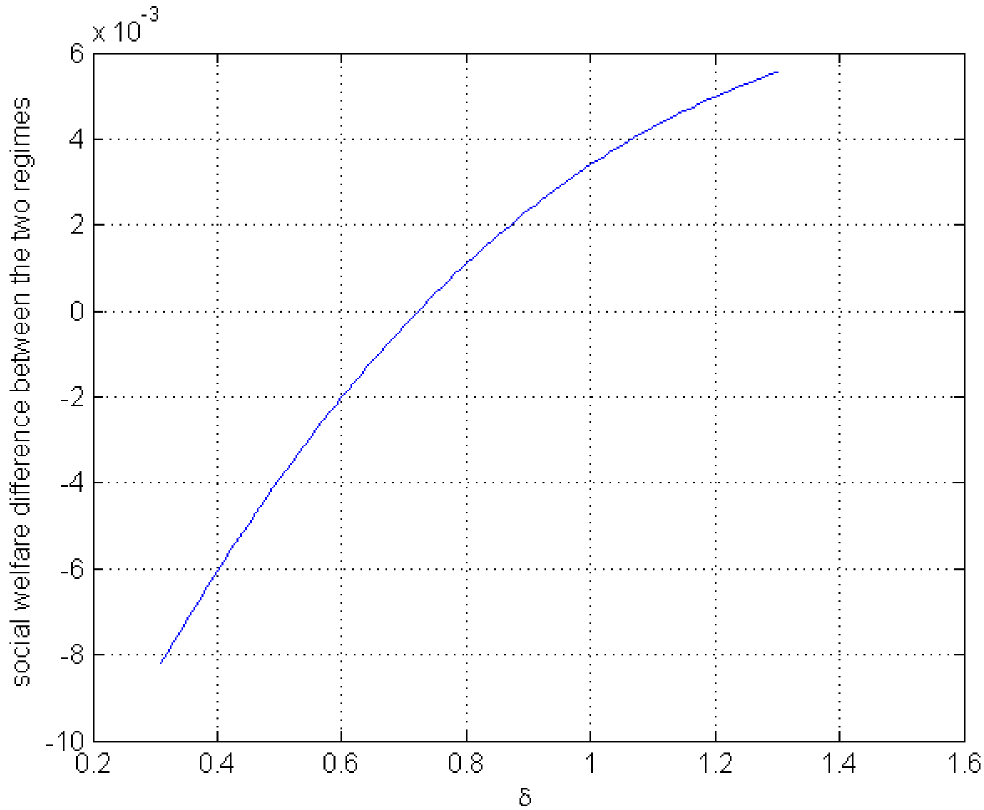
Comparison between the two regimes

The trade-off between regime S and SD is similar to the basic model. Specifically, choosing option DE allows a creditor to better predict θ, the fundamental. On the other hand, choosing option SU allows investors to better coordinate their actions and thus reduces panic runs. The trade-off depends on how likely panic runs are, which in turns is determined by the magnitude of complementarity.

Conjecture: $W_{SU} - W_{DE}$ increases in δ.

In Figure 1.4 below, I plot the welfare difference $W_{SU} - W_{DE}$ as a function of δ. the plot indeed shows that the larger δ, the larger $W_{SU} - W_{DE}$. Furthermore, $W_{SU} - W_{DE}$ changes sign around $\delta = 0.7$.

Figure 1.4: $W_{SU} - W_{DE}$ as a function of δ, when $\sigma_\gamma^2 = 1.5$ and $C = 2$.



Furthermore, as C further increases, details should become easier to process, and $W_S - W_{SD}$ should become smaller.

Conjecture: $W_{SU} - W_{DE}$ decreases in C.

In Figure 1.5 below, I plot the welfare difference $W_{SU} - W_{DE}$ as a function of C. The plot indeed shows that the larger C, the smaller $W_{SU} - W_{DE}$. Furthermore, $W_{SU} - W_{DE}$ changes sign around $C = 2.4$.

Figure 1.5: $W_{SU} - W_{DE}$ as a function of C, when $\sigma_\gamma^2 = 1.5$ and $\delta = 1$.

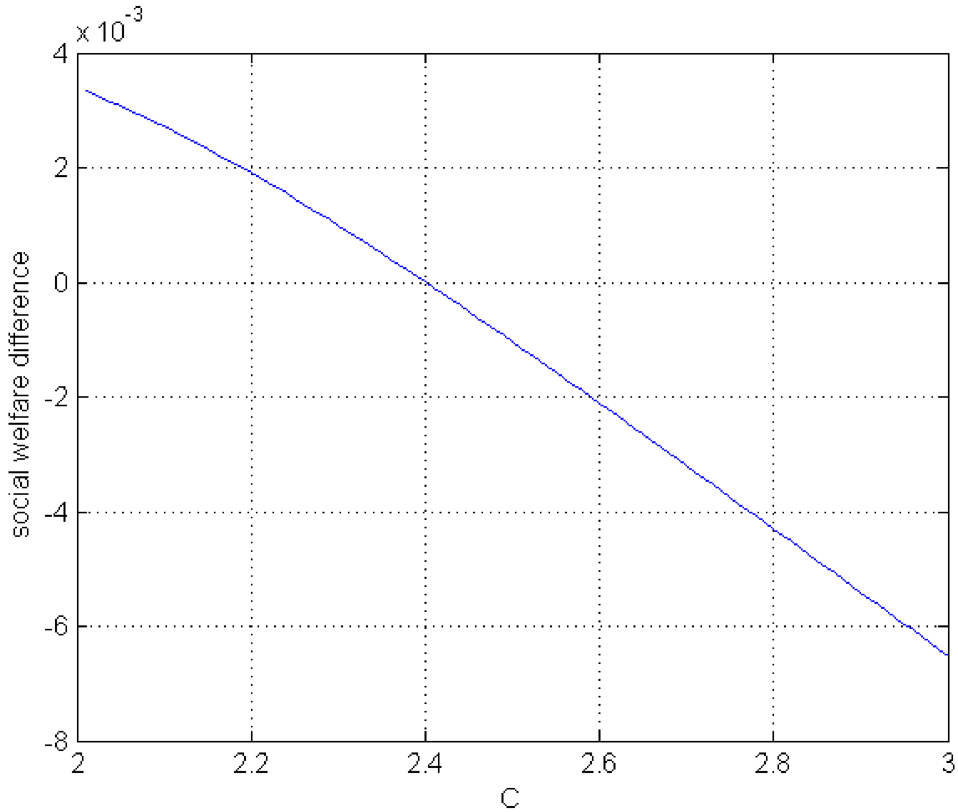


Figure 1.4 and 1.5 together imply that summary dominates summary plus details when C is sufficiently small or δ is large, consistent with the main results in the basic model.

1.7 Extensions

In this section, I study four extensions of the basic model. In the first one, I show that suppressing details might be desirable even if coordination is not socially beneficial. In the second one, I provide foundation for why aggregation leads to loss of information by studying a simple example of aggregating different components of revenues. In the third one, I show that even if the summary is a perfect measure of the fundamental, investors may still inefficiently focus on the details. In the fourth one, I show that even though investors can pay attention to both the summary and the details in regime SD, there still exists an equilibrium in which investors all focus on the details, as long as their capacity is large. This shows that providing details in addition to the summary may still overload investors even if we relax Assumption 2.

1.7.1 When coordination is not socially beneficial

One might think all the results so far rely on the assumption that coordination is socially beneficial, as in Angeletos and Pavan (2004). In this section, I extend my analysis by considering an alternative setting in which investors still need to coordinate their decisions, but social welfare is defined as how well investors' decisions are fine-tuned to the fundamental, as in Morris and Shin (2002). Interestingly, I find a similar result: suppressing details is desirable when investors' capacity is low or when the strength of complementarity is high.

To start, I assume investor i's utility is

$$u_i = -(1-r)(k_i - \theta)^2 - r(k_i - K)^2, \tag{1.12}$$

where k_i is investor i's action and $K = \int k_i di$ is the average action. The social welfare is defined as $W = -\int (k_i - \theta)^2 di$. In other words, whereas investors still care about both the fundamental and the average action, social welfare is solely determined by the mean-squared distance of investors' decisions to the fundamental. All other assumptions remain the same.

What are investors' attention allocation strategies when their utility function is described by (1.12)? In fact, their strategies should remain the same as before, because the optimal investment choice satisfies the same first order condition as before, which is given by

$$k_i = (1-r)E[\theta] + rE[K].$$

We thus have the following Lemma.

Lemma 7. *If investors' utility functions are defined in* (1.12)*, their attention-allocation* *strategies in regime S, D, and SD are the same as in Lemma 3, Proposition 1, and Proposition 3.*

However, the welfare comparisons between the regimes should be obviously different. The following Proposition shows that a similar result to Proposition 5 holds.

Proposition 8. *Consider here the case where $C > C^* \equiv \frac{2}{\sigma_\gamma^2}$, so that DE_BOTH is always an equilibrium in regime SD. There exists a unique C_0 such that $S \succeq SD$, if and only if $C \in (\frac{2}{\sigma_\gamma^2}, C_0)$. Furthermore $\frac{\partial C_0}{\partial r} > 0$.*

In other words, details should be suppressed when investors' information processing capacity is not too high. Furthermore, the higher the strength of complementarity among investors' actions, the more likely it is that suppressing details increases welfare.

Why should details be suppressed even when they have higher information value (i.e., when $C > \frac{2}{\sigma_\gamma^2}$)? This is because the coordination value of signals still affects social welfare indirectly, even if it is not directly included in the social welfare function. Intuitively, compared to SU, in DE_BOTH, investors put less weight on their signals and more weight on the prior of the fundamental. This is because investors' signals of details are more private than their signals of the summary. As a result, to better coordinate their decisions, investors behave as if they "discount" signals of details more than signals of the summary. Hence, when the strength of complementarity is sufficiently high, such that the "discount" effect of details outweighs the information advantage of details, details should be suppressed.

1.7.2 When details have different persistence

In this extension, I study a simple example in which the two details (θ_x and θ_y) have different importance, to endogeneize the assumption that aggregation (summarization) leads to loss of information (i.e., the assumption that the summary is contaminated by γ). I show that in this example, the trade-off between summary and details still holds.

Suppose the firm's total revenue θ consist of two independent components, domestic revenue θ_x and foreign revenue θ_y, where $\theta_x \sim N(0,1)$ and $\theta_y \sim N(0,1)$. Investors want to take actions k_i that are fine tuned to $t\theta_x + \theta_y$, that is:

$$u_i = -(k_i - (t\theta_x + \theta_y))^2, \tag{1.13}$$

where $t > 1$ is a parameter that captures the relative importance of the two components. For example, if domestic revenue is more persistent than foreign revenue, investors will attach higher weight to domestic revenue when evaluating the performance of the firm. However, the firm cannot directly report $t\theta_x + \theta_y$ in an aggregated form. Instead, because of exogenous reasons such as the accounting rules, the firm can only report the total revenue $\theta = \theta_x + \theta_y$. Formally, we consider the following two reporting regimes:

In Regime S (summary), the reporting system reports the total revenue: $I_S = \{\theta_x + \theta_y\}$.

In Regime D (details), the reporting system reports both foreign and domestic revenues: $I_D = \{\theta_x, \theta_y\}$.

How will investors allocate their attention? In regime S, investors only have the total revenue to attend to and therefore will devote all capacity to it. In regime D, investor will allocate their attention between the two components. Domestic revenue will attract more attention since it is more persistent. This is summarized in the following proposition.

Proposition 9. *In regime S, investors will pay all the attention to the summary. In regime D, investors will pay more attention to domestic revenue than foreign revenue.*

In this example, what is the trade-off between summary and details? First, if C is very large, details are preferred to the summary because the summary is an imperfect measure of the fundamental, while details are precise. On the other hand, if C is very small, summary dominates because it is easier to process even though it is imprecise. Specifically, we know from the basic model that in regime D, investors have to allocate their capacity between the two details, which results in an less efficient use of their capacity, compared to the case in which all the capacity is used to process one signal. The following proposition is thus a natural representation of the trade-off between summary and details.

Proposition 10. *Consider the case in which investors' attention allocation strategy has an interior solution in regime D, so that investors will pay positive amount of attention to both details. There exists C^* such that regime S is preferred to regime D if and only if $C < C^*$.*

Proposition 10 shows a similar result as in Proposition 1 in the basic model. Therefore, this extension shows that the trade-off between summary and details is robust to how the information loss in the aggregation process is specified.

1.7.3 When the summary is a perfect measure of the fundamental

One might wonder that Proposition 5 relies on the assumption that the summary is an imperfect measure of the fundamental. In this extension, I show that even if the summary perfectly measures the fundamental, as long as the two details are not perfectly symmetric, attaching details to the summary may still overload investors.

Specifically, suppose everything is identical to the basic model, except that we are relaxing the following two assumptions. First, the two details are no longer symmetric: $\sigma_x^2 \neq \sigma_y^2$. Second, the summary is a perfect measure of the fundamental: $\sigma_\gamma^2 = 0$. We will show that reporting details together with the summary may still overload investors.

We start with investors' equilibrium attention allocation strategy in regime SD.

Proposition 11. *In regime SD, SU is always an equilibrium.*

The intuition for Proposition 11 is as follows.

First, because investors are ultimately interested in $\theta = \theta_x + \theta_y$, choosing the summary report is more efficient in predicting the fundamental θ. Namely, when $r = 0$, choosing option SU is a dominant strategy for all investors.

When $r > 0$, investor i gets less payoff if his investment deviates from the aggregate investment. However, investor i's best response to every other investor choosing option

SU is also option SU, not option DE, because $K \propto theta$. In other words, predicting the aggregate investment K is equivalent to predicting the fundamental θ. Hence, investor i will not deviate from choosing option SU if other investors are also choosing option SU.

However, is the equilibrium in Proposition 11 the only possible equilibrium? It turns out that a "bad" equilibrium can occur as well.

Proposition 12. *In regime SD, if* $\frac{\sigma_y}{\sigma_x} > \sqrt{3}$ *and if* $r \in (\frac{C^2\sigma_y - \sigma_x}{(C^2-1)\sigma_y}, 1)$, *then DE_BOTH is also an equilibrium.*

Proposition 12 provides an interesting case in which investors may end up choosing option DE in the equilibrium. Now I dig into the intuition for this finding.

In the strategy profile in which every investor chooses option DE, the aggregate investment should take the form $K = \int_i (b_x x_i + b_y y_i) di = b_x \theta_x + b_y \theta_y$.

Consider investor i's deviation to option SU (summary). This deviation will make predicting θ easier. However, if b_x and b_y differ a lot, deviating to option SU will make predicting the aggregate investment level K harder. Thus, investor i faces a trade-off between these two effects when considering the deviation to option SU. As a result, if b_x and b_y differ a lot (which happens when σ_x and σ_y differ a lot, similar to Mackowiak and Wiederholt (2009)) and if coordination plays an important role (when r is sufficiently large), investor i will NOT deviate and will stick to option DE.

The following proposition explains why the equilibrium in Proposition 12 is a bad one.

Proposition 13. *Investors' welfare in DE_BOTH is always lower than in SU.*

In other words, if the "bad" equilibrium exists, social welfare must be lower in the "bad" equilibrium. As a result, providing an option of details in addition to the summary may overload investors and reduce social welfare.

The intuition for Proposition 13 is clear. First, choosing the summary over the details can help investors better predict the fundamental. Second, choosing the summary also helps investors better coordinate their investment decisions, because devoting all the attention to the summary report results in a lower variance of investors' private noise, compared to the case in which attention is allocated between the details. Consequently, because private noise impedes coordination, investors can achieve better coordination by focusing on the summary.

Similar to Proposition 5 in the basic model, Proposition 13 describes a situation in which investors are trapped in a bad equilibrium in which they all pay attention to the details. This equilibrium exists because any investor i may find it optimal to choose details over the summary if he believes other investors are doing the same, despite the fact that the equilibrium outcome is inefficient.

1.7.4 When investors can pay attention to both the summary and details in regime SD

In this subsection, we allow investors to focus on both the summary and the details in regime SD. While the complete set of equilibria cannot be characterized, I will show that an equilibrium exists in which investors will devote all the attention to details when their capacity is sufficiently large (i.e., $C > C^*$). In other words, even though given the option to pay attention to both the summary and details, investors may still focus only on the details when their capacity is large. Hence, this result implies that the information overload problem described in Proposition 5 may still happen even if investors can attend to both the summary and details. Formally, we have the following result.

Proposition 14. *When $C > C^* \equiv \frac{2}{\sigma_\gamma^2}$, DE_BOTH is always an equilibrium in regime SD. Furthermore, in this case, social welfare in DE_BOTH is larger than social welfare in regime S if and only if $r > r^*$.*

The proof of Proposition 14 is in the Appendix. The basic idea is to show that when $C > C^*$, investors do not have incentives to deviate from the equilibrium. In fact, in DE_BOTH, $K \propto \theta$. Therefore, investors will not deviate if devoting all their attention to the details is their optimal strategy to estimate θ, which is true under the entropy technology.

1.8 Conclusions

In this paper, I develop a theory that focuses on the consequence of providing detailed information to rationally inattentive investors. I first provide a sensible way to model the natural trade-off between summary and details. I show that investors with limited information capacity may prefer a summary to details even if less information is contained in the summary than details. Furthermore, I show that, in environments where investors' decisions are complements, attaching details to a summary may not be desirable, mainly because details are hard to process and may exacerbate the coordination problem among investors.

My results have important implications for accounting standard setting. In its performance disaggregation project, the FASB is currently considering the disaggregation of income statement items. Whereas such disaggregation may improve transparency, my results shed light on an important potential cost from the perspective of investors: providing detailed information may overload investors and lead to a reduction in investors' decision quality.

Although my model aims to capture the "beauty contest" feature of the secondary market, I do not micro-found the formation of stock price and do not directly show the "beauty contest" effect. Therefore, the implications of the model may not extend to the case where stock price efficiency is the criterion for welfare. Future research can explore the implications of limited attention on the formation and the assimilation of stock prices.

REFERENCES

Angeletos, George-Marios and Alessandro Pavan (2004), "Transparency of information and coordination in economies with investment complementarities." *American Economic Review*, 94, 91–98.

Arya, Anil and Jonathan C Glover (2013), "On the upsides of aggregation." *Journal of Management Accounting Research*, 26, 151–166.

Banerjee, Snehal, Jesse Davis, and Naveen Gondhi (2018), "When transparency improves, must prices reflect fundamentals better?" *The Review of Financial Studies*, 31, 2377–2414.

Bernard, Victor and Katherine Schipper (1994), "Recognition and disclosure in financial reporting." *Unpublished paper, University of Michigan*.

Blackwell, David (1950), "Comparison of experiments." Technical report, HOWARD UNIVERSITY Washington United States.

Blankespoor, Elizabeth (2019), "The impact of information processing costs on firm disclosure choice: Evidence from the xbrl mandate." *Journal of Accounting Research*.

Chakrabarty, Bidisha and Pamela C Moulton (2012), "Earnings announcements and attention constraints: The role of market design." *Journal of Accounting and Economics*, 53, 612–634.

Chen, Qi, Itay Goldstein, and Wei Jiang (2010), "Payoff complementarities and financial fragility: Evidence from mutual fund outflows." *Journal of Financial Economics*, 97, 239–262.

Datar, Srikant and Mahendra Gupta (1994), "Aggregation, specification and measurement errors in product costing." *The Accounting Review*, 567–591.

Drake, Michael S, Jacob R Thornock, and Brady J Twedt (2017), "The internet as an information intermediary." *Review of Accounting Studies*, 22, 543–576.

Driskill, Matthew, Marcus Kirk, and Jenny Wu Tucker (2019), "Concurrent earnings announcements and analysts information production." *SSRN*.

Dye, Ronald A and Sri S Sridhar (2004), "Reliability-relevance trade-offs and the efficiency of aggregation." *Journal of Accounting Research*, 42, 51–88.

Elliott, W Brooke, Jessen L Hobson, and Brian J White (2015), "Earnings metrics, information processing, and price efficiency in laboratory markets." *Journal of Accounting Research*, 53, 555–592.

Frankel, David, Xiangou Deng, and zhongzheng Zhou (2019), "Underpricing and underreaction in initial public offerings: The role of strategic risk." *Working Paper*.

Froot, Kenneth A, David S Scharfstein, and Jeremy C Stein (1992), "Herd on the street: Informational inefficiencies in a market with short-term speculation." *The Journal of Finance*, 47, 1461–1484.

Goldstein, Itay and Haresh Sapra (2014), "Should banks' stress test results be disclosed? an analysis of the costs and benefits." *Foundations and Trends in Finance*, 8, 1–54.

Grossman, Sanford J and Joseph E Stiglitz (1980), "On the impossibility of informationally efficient markets." *American Economic Review*, 70, 393–408.

Guay, Wayne, Delphine Samuels, and Daniel Taylor (2016), "Guiding through the fog: Financial statement complexity and voluntary disclosure." *Journal of Accounting and Economics*, 62, 234–269.

Hellwig, Christian and Laura Veldkamp (2009), "Knowing what others know: Coordination motives in information acquisition." *The Review of Economic Studies*, 76, 223–251.

Hirshleifer, David and Siew Hong Teoh (2003), "Limited attention, information disclosure, and financial reporting." *Journal of Accounting and Economics*, 36, 337–386.

Kempf, Elisabeth, Alberto Manconi, and Oliver Spalt (2016), "Distracted shareholders and corporate actions." *The Review of Financial Studies*, 30, 1660–1695.

Lev, Baruch (1968), "The aggregation problem in financial statements: An informational approach." *Journal of Accounting Research*, 247–261.

Li, Feng (2008), "Annual report readability, current earnings, and earnings persistence." *Journal of Accounting and economics*, 45, 221–247.

Lim, Suk S and Shyam Sunder (1991), "Efficiency of asset valuation rules under price movement and measurement errors." *The Accounting Review*, 669–693.

Mackowiak, Bartosz and Mirko Wiederholt (2009), "Optimal sticky prices under rational inattention." *American Economic Review*, 99, 769–803.

Morris, Stephen and Hyun Song Shin (2000), "Rethinking multiple equilibria in macroeconomic modeling." *NBER macroeconomics Annual*, 15, 139–161.

Morris, Stephen and Hyun Song Shin (2002), "Social value of public information." *American Economic Review*, 92, 1521–1534.

Myatt, David P and Chris Wallace (2011), "Endogenous information acquisition in coordination games." *The Review of Economic Studies*, 79, 340–374.

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

Speier, Cheri, Joseph S Valacich, and Iris Vessey (1999), "The influence of task interruption on individual decision making: An information overload perspective." *Decision Sciences*, 30, 337–360.

Sunder, Shyam (1997), *Theory of accounting and control*. South-Western College Pub.

Yang, Ming (2015), "Coordination with flexible information acquisition." *Journal of Economic Theory*, 158, 721–738.

APPENDIX A

PROOFS

A.1 Proof of Proposition 1

Consider investor i's decision in regime D. Suppose his conjecture of other investors' actions is $k_j = \hat{b}_x x_j + \hat{b}_y y_j$, where $\hat{b}_x$ and $\hat{b}_y$ are the conjectured coefficients. As a result, investor i's conjecture of the aggregate investment is $K = \int_j k_j dj = \hat{b}_x \theta_x + \hat{b}_y \theta_y$.

Hence, upon receiving the signals x_i and y_i, investor i will choose the investment level as

$$k_i = (1-r)E[\theta|x_i, y_i] + rE[K|x_i, y_i] = b_x x_i + b_y y_i,$$

where

$$b_x = (1-r)\beta_x + r\beta_x \hat{b}_x,$$
$$b_y = (1-r)\beta_y + r\beta_y \hat{b}_y,$$

and β_x and β_y are the precisions of x and y:

$$\beta_x = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_1^2},$$
$$\beta_y = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_2^2}.$$

Therefore, investor i's optimal attention-allocation problem is choosing σ_1 and σ_2 to maximize his expected utility, subject to the information capacity constraint:

$$\max_{\sigma_1,\sigma_2} E[u_i] = E[Ak_i - \frac{1}{2}k_i^2] \tag{A.1}$$
$$= b_x(1-r+r\hat{b}_x)\sigma_x^2 + b_y(1-r+r\hat{b}_y)\sigma_y^2 - \frac{1}{2}(b_x^2(\sigma_x^2+\sigma_1^2) + b_y^2(\sigma_y^2+\sigma_2^2)),$$
$$s.t.\ \frac{(\frac{1}{\sigma_x^2}+\frac{1}{\sigma_1^2})(\frac{1}{\sigma_y^2}+\frac{1}{\sigma_2^2})}{\frac{1}{\sigma_x^2}\frac{1}{\sigma_y^2}} \leq C^2.$$

Interior solution: DE_BOTH

The interior solution of (A.1) is

$$\sigma_1^2 = \frac{(1-r+r\hat{b}_x)\sigma_x\sigma_y^2}{C\sigma_y(r\hat{b}_y+1-r) - \sigma_x(r\hat{b}_x+1-r)},$$
$$\sigma_2^2 = \frac{(1-r+r\hat{b}_y)\sigma_y\sigma_x^2}{C\sigma_x(r\hat{b}_x+1-r) - \sigma_y(r\hat{b}_y+1-r)}.$$

In the equilibrium, we should have $b_x = \hat{b}_x$ and $b_y = \hat{b}_y$. Plugging in the above expression of σ_1 and σ_2 yields the following results:

$$\hat{b}_x = \frac{(C-1)(1-r)}{r+(1-r)C},$$
$$\hat{b}_y = \frac{(C-1)(1-r)}{r+(1-r)C}.$$

Now we verify whether the interior solution (DE_BOTH) is indeed an equilibrium.

Plugging in the formula for $\hat{b}_x$ and $\hat{b}_y$ and resolve (A.1), we get $\sigma_1 = \sigma_2 = \frac{1}{C-1}$. Therefore, DE_BOTH is always an equilibrium in regime D. Social welfare in this equilibrium is equal to the expected utility of any investor i:

$$W_{DE\_BOTH} = u_{DE\_BOTH} = \frac{(C-1)C(r-1)^2}{(C(1-r)+r)^2}.$$

Corner solution: DE_X

The corner solution of (A.1) is

$$\sigma_1^2 = +\infty,$$
$$\sigma_2^2 = \frac{1}{C^2 - 1}.$$

In the equilibrium, we should have $b_x = \hat{b}_x$ and $b_y = \hat{b}_y$. Plugging in the above expression of σ_1 and σ_2 yields the following results:

$$\hat{b}_x = \frac{(C^2 - 1)(1 - r)}{C^2(1 - r) + r},$$
$$\hat{b}_y = 0.$$

Now we verify whether the corner solution (DE_X) is indeed an equilibrium.

Plugging in the formula for $\hat{b}_x$ and $\hat{b}_y$ and resolve (A.1), I find that the optimization problem above generates an interior solution of σ_1 and σ_2 if and only if $r \in (0, \frac{C}{C+1})$. When $r \in (\frac{C}{C+1}, 1)$, the unique corner solution is $\sigma_1 < +\infty$ and $\sigma_2 = +\infty$. Let $r_0 \equiv \frac{C}{C+1}$. Therefore, DE_X is an equilibrium if and only if $r > r_0$. By symmetry, DE_Y is an equilibrium under the same condition. Social welfare in these two equilibria is equal to the expected utility of any investor i:

$$W_{DE\_X} = u_{DE\_X} = \frac{C^2\left(C^2 - 1\right)(r - 1)^2}{2\left(r - C^2(r - 1)\right)^2}.$$

Q.E.D.

A.2 Proof of Lemma 4

When $r = 0$, based on the proof of proposition 1, we know that DE_X is not an equilibrium. Therefore, DE_BOTH is the unique equilibrium in regime D. Q.E.D.

A.3 Proof of Proposition 2

In the proof of Proposition 3 below, I show that under the general case $r \in [0,1]$, DE_BOTH exists if and only if $C > \frac{2}{\sigma_\gamma^2}$. Furthermore, SU exists if and only if one of the following two conditions holds:

$$\text{①}\ C < \frac{2}{\sigma_\gamma^2}.$$
$$\text{②}\ C > \frac{2}{\sigma_\gamma^2} \text{ and } r \in (r_2, 1).$$

Since $r_2 > 0$. we know that in the special case $r = 0$, SU exists if and only if $C < \frac{2}{\sigma_\gamma^2}$. Q.E.D.

A.4 Proof of Proposition 3

I divide the proof into three steps. In the first step, I derive the conditions for the existence of DE_BOTH. In the second step, I derive the conditions for the existence of SU. In the third step, I derive the conditions for the existence of DE_X.

Existence of DE_BOTH

Consider investor i's deviation to option SU. It is obvious that following this deviation, he will devote all his capacity to minimize σ^2, implying the mutual-entropy constraint will bind. Hence, $\sigma^2 = \frac{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2}{C^2-1}$.

Conditional on receiving the signal z_i, investor i will choose the investment level as

$$k_i = (1-r)E[\theta|z_i] + rE[K|z_i] = b_z z_i,$$

where $b_z = (1-r)\frac{\sigma_x^2+\sigma_y^2}{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2+\sigma^2} + r\frac{b_x\sigma_x^2+b_y\sigma_y^2}{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2+\sigma^2}$.His utility from deviation is thus

$$u_{\text{deviation}} = E[Ak_i - \frac{1}{2}k_i^2] = \frac{2\left(C^2-1\right)(r-1)^2}{(\sigma_\gamma^2+2)(C(1-r)+r)^2}.$$

It can be easily derived that $u_{DE\_BOTH} > u_{\text{deviation}}$ if and only if $C > \frac{2}{\sigma_\gamma^2}$. Hence, equilibrium DE_BOTH exists if and only if $C > \frac{2}{\sigma_\gamma^2}$. This completes the first step. Q.E.D.

Existence of SU

Consider investor i's decision in SU. Suppose his conjecture of any other investor's actions is $k_j = \hat{b}_z z_j$. As a result, investor i's conjecture of the aggregate investment is $K = \int_j k_j dj = \hat{b}_z\theta$.

It is obvious that following his choice of option SU, investor i will devote all his capacity to minimize σ^2. Hence, $\sigma^2 = \frac{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2}{C^2-1}$.

Conditional on receiving the signal z_i, investor i will choose the investment level as

$$k_i = (1-r)E[\theta|z_i] + rE[K|z_i] = b_z z_i = ((1-r)\beta + r\beta'\hat{b}_z)z_i,$$

where βz_i is investor i's estimation of θ, and $\beta' z_i$ is investor i's estimation of $\theta+\gamma$. Therefore,

$$\beta = \frac{\sigma_x^2+\sigma_y^2}{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2+\sigma^2},$$
$$\beta' = \frac{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2}{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2+\sigma^2} = 1-\frac{1}{C^2}.$$

Solving $b_z = \hat{b}_z$ yields $b_z = \frac{2(C^2-1)(r-1)}{(\sigma_\gamma^2+2)(C^2(r-1)-r)}$. Plugging b_z into $E[u_i]$, I get

$$W_{SU} = u_{SU} = \frac{2C^2\left(C^2-1\right)(r-1)^2}{(\sigma_\gamma^2+2)\left(r-C^2(r-1)\right)^2}.$$

Consider investor i's deviation to option DE. Conditional on receiving the signals x_i and

y_i, investor i will choose the investment level as

$$k_i = (1-r)E[\theta|x_i, y_i] + rE[K|x_i, y_i] = b_x x_i + b_y y_i$$
$$= ((1-r)\beta_x + r\beta_x b_z)x_i + ((1-r)\beta_y + r\beta_y b_z)y_i,$$

where $\beta_x x_i$ is investor i's estimation of θ_x, and $\beta_y y_i$ is investor i's estimation of θ_y. Therefore,

$$\beta_x = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_1^2},$$
$$\beta_y = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_2^2}.$$

Investor i's optimal attention-allocation problem is thus

$$\max_{\sigma_1,\sigma_2} E[u_i] = E[Ak_i - \frac{1}{2}k_i^2]$$
$$= b_x(1-r+rb_z)\sigma_x^2 + b_y(1-r+rb_z)\sigma_y^2 - \frac{1}{2}(b_x^2(\sigma_x^2+\sigma_1^2) + b_y^2(\sigma_y^2+\sigma_2^2)),$$
$$s.t.\ \frac{(\frac{1}{\sigma_x^2}+\frac{1}{\sigma_1^2})(\frac{1}{\sigma_y^2}+\frac{1}{\sigma_2^2})}{\frac{1}{\sigma_x^2}\frac{1}{\sigma_y^2}} \leq C^2.$$

Solving this problem gives $\sigma_1 = \sigma_2 = \frac{1}{C-1}$.

Therefore, $b_x = b_y = (1-r)\beta_x + r\beta_x b_z = \frac{(C-1)(1-r)\left(C^2((r-1)\sigma_\gamma^2-2)-r\sigma_\gamma^2\right)}{C(\sigma_\gamma^2+2)(C^2(r-1)-r)}$. Furthermore, investor i's utility from deviation is $u_{\text{deviation}} = \frac{(C-1)(r-1)^2\left(C^2(-r\sigma_\gamma^2+\sigma_\gamma^2+2)+r\sigma_\gamma^2\right)^2}{C(\sigma_\gamma^2+2)^2(r-C^2(r-1))^2}$.

It can be easily checked that $u_{SU} > u_{\text{deviation}}$ if and only if $C < \frac{2}{\sigma_\gamma^2}$, or $C > \frac{2}{\sigma_\gamma^2}$ and $r \in (\frac{C^2\sigma_\gamma^2+2C^2}{(C^2-1)\sigma_\gamma^2} - \sqrt{2}\sqrt{\frac{C^3\sigma_\gamma^2+2C^3}{(C-1)^2(C+1)\sigma_\gamma^4}}, 1)$. Let $r_2 \equiv \frac{C^2\sigma_\gamma^2+2C^2}{(C^2-1)\sigma_\gamma^2} - \sqrt{2}\sqrt{\frac{C^3\sigma_\gamma^2+2C^3}{(C-1)^2(C+1)\sigma_\gamma^4}}$. Hence, equilibrium SU exists if and only if one of the following two conditions hold:

① $C < \frac{2}{\sigma_\gamma^2}$.

② $C > \frac{2}{\sigma_\gamma^2}$ and $r \in (r_2, 1)$.

Q.E.D.

Existence of DE_X and DE_Y

Consider investor i's decision in DE_X. Investor i has two potential deviations. The first is to stick to option DE but pay slightly less attention to θ_x and pay slightly more attention to θ_y. The second is to switch to option SU.

In the proof of Proposition 1, I already showed that investor i will not make the first deviation if and only if $r > r_0$.

Consider investor i's deviation to option SU. It is obvious that following this deviation, he will devote all his capacity to minimize σ^2, implying that the mutual entropy constraint will bind. Hence, $\sigma^2 = \frac{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2}{C^2-1}$.

Conditional on receiving the signal z_i, investor i will choose the investment level as

$$k_i = (1-r)E[\theta|z_i] + rE[K|z_i] = b_z z_i,$$

where $b_z = (1-r)\frac{\sigma_x^2+\sigma_y^2}{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2+\sigma^2} + r\frac{\hat{b}_x\sigma_x^2}{\sigma_x^2+\sigma_y^2+\sigma_\gamma^2+\sigma^2}$. His utility from deviation is thus

$$u_{\text{deviation}} = E[Ak_i - \frac{1}{2}k_i^2] = \frac{\left(C^2-1\right)(r-1)^2\left(r - C^2(r-2)\right)^2}{2(\sigma_\gamma^2+2)\left(C^3(r-1)-Cr\right)^2}$$

It can be easily checked that $u_{\text{deviation}} < u_{\text{option DE\_X}}$ if and only if $r \in (r_3, 1)$, where $r_3 \equiv \frac{2C^2}{C^2-1} - \sqrt{\frac{C^4\sigma_\gamma^2+2C^4}{(C^2-1)^2}}$. Hence, investor i will not deviate to option SU when $r \in (r_3, 1)$.

Furthermore, it can be easily shown that when $C > \frac{2}{\sigma_\gamma^2}$, $r_0 > r_3$. Hence, when $C > \frac{2}{\sigma_\gamma^2}$, the necessary and sufficient condition for the existence of equilibrium DE_X is $r \in (r_0, 1)$. When $C < \frac{2}{\sigma_\gamma^2}$, the necessary and sufficient condition for the existence of equilibrium DE_X is $r \in (r_1, 1)$, where $r_1 = r_3$ or $r_1 = r_0$, depending on the parameter values. Q.E.D.

A.5 Proof of Lemma 6

From the proof of Proposition 3, we already know

$$W_{SU} = \frac{2C^2\left(C^2-1\right)(r-1)^2}{(\sigma_\gamma^2+2)\left(r-C^2(r-1)\right)^2},$$
$$W_{DE\_X} = \frac{C^2\left(C^2-1\right)(r-1)^2}{2\left(r-C^2(r-1)\right)^2}.$$

Since $1+\frac{\sigma_\gamma^2}{2}<2$, it is straightforward that $W_{SU}>W_{DE\_X}$. Q.E.D.

A.6 Proof of Proposition 4 and Corollary 3

From the proof of Proposition 1 and Proposition 3, we already know

$$W_{SU} = \frac{2C^2\left(C^2-1\right)(r-1)^2}{(\sigma_\gamma^2+2)\left(r-C^2(r-1)\right)^2}.$$
$$W_{DE\_BOTH} = \frac{(C-1)C(r-1)^2}{(C(1-r)+r)^2}.$$

It can be easily derived that $W_{SU}>W_{DE\_BOTH}$ if and only if $r>r^*\equiv\frac{C^2\sigma_\gamma^2}{(C-1)(C\sigma_\gamma^2+\sigma_\gamma^2+2)}-\sqrt{2}\sqrt{\frac{C^3\sigma_\gamma^2+2C^3}{(C-1)^2(C+1)(C\sigma_\gamma^2+\sigma_\gamma^2+2)^2}}$. Furthermore, it can be easily checked that $r^*\in(r_2,r_0)$ and that $\frac{dr^*}{dC}>0$. Q.E.D.

A.7 Proof of Proposition 8

The proof can be started by directly computing $L=E[(k_i-\theta)^2]$ in SU and DE_BOTH:

$$L_{SU}=E[(b_x x_i+b_y y_i-\theta)^2]=\frac{2\left(C^4(r-1)^2\sigma_\gamma^2-2C^2(r-1)(r(\sigma_\gamma^2-1)+1)+r(r(\sigma_\gamma^2-2)+4)\right)}{(\sigma_\gamma^2+2)\left(r-C^2(r-1)\right)^2},$$
$$L_{DE\_BOTH}=E[(b_z z_i-\theta)^2]=\frac{2\left(C(r-1)^2-(r-2)r\right)}{(C(-r)+C+r)^2}.$$

Comparing L_{SU} and L_{DE_BOTH} is complicated because they include non-cancelable terms such as C^4. However, some programming in Mathematica shows that for any $r \in (0,1)$ and $\sigma_\gamma^2 \in (0,2)$, $C_0 = C_0(r, \sigma_\gamma^2)$ exists such that $L_{SU} < L_{DE\_BOTH}$ if and only if $C < C_0$. Furthermore, C_0 increases in r. The Mathematica code is available upon request.

A.8 Proof of Proposition 12

The proof is divided into two steps. I first derive investors' equilibrium strategy, assuming they all choose option DE. I then show that DE is indeed an equilibrium.

A.8.1 The first step

Consider the decision of an arbitrary investor i after he has chosen option DE. Suppose his conjecture of other investors' actions is $k_j = \hat{b}_x x_j + \hat{b}_y y_j$, where $\hat{b}_x$ and $\hat{b}_y$ are the conjectured coefficients. As a result, investor i's conjecture of the aggregate investment is $K = \int_j k_j dj = \hat{b}_x \theta_x + \hat{b}_y \theta_y$.

Hence, upon receiving the signals x_i and y_i, investor i will choose the investment level as

$$k_i = (1-r)E[\theta|x_i, y_i] + rE[K|x_i, y_i] = b_x x_i + b_y y_i,$$

where

$$b_x = (1-r)\beta_x + r\beta_x \hat{b}_x,$$

$$b_y = (1-r)\beta_y + r\beta_y \hat{b}_y,$$

and β_x and β_y are the precisions of x and y:

$$\beta_x = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_1^2},$$
$$\beta_y = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_2^2}.$$

Therefore, investor i's optimal attention-allocation problem is choosing σ_1 and σ_2 to maximize his expected utility, subject to the information capacity constraint:

$$\begin{aligned}\max_{\sigma_1,\sigma_2} E[u_i] &= E[Ak_i - \frac{1}{2}k_i^2] \\ &= b_x(1 - r + r\hat{b}_x)\sigma_x^2 + b_y(1 - r + r\hat{b}_y)\sigma_y^2 - \frac{1}{2}(b_x^2(\sigma_x^2 + \sigma_1^2) + b_y^2(\sigma_y^2 + \sigma_2^2)), \\ s.t.\ & \frac{(\frac{1}{\sigma_x^2} + \frac{1}{\sigma_1^2})(\frac{1}{\sigma_y^2} + \frac{1}{\sigma_2^2})}{\frac{1}{\sigma_x^2}\frac{1}{\sigma_y^2}} \leq C^2.\end{aligned}$$

The interior solution of the above problem is

$$\sigma_1^2 = \frac{(1 - r + r\hat{b}_x)\sigma_x\sigma_y^2}{C\sigma_y(r\hat{b}_y + 1 - r) - \sigma_x(r\hat{b}_x + 1 - r)},$$
$$\sigma_2^2 = \frac{(1 - r + r\hat{b}_y)\sigma_y\sigma_x^2}{C\sigma_x(r\hat{b}_x + 1 - r) - \sigma_y(r\hat{b}_y + 1 - r)}.$$

In the equilibrium, we should have $b_x = \hat{b}_x$ and $b_y = \hat{b}_y$. Plugging in the above expression of σ_1 and σ_2 yields the following results:

$$b_x = \frac{(r-1)\left(C^2(r-1)\sigma_x + C\sigma_y - r\sigma_x\right)}{\sigma_x\left(C^2(r-1)^2 - r^2\right)},$$
$$b_y = \frac{(r-1)\left(C^2(r-1)\sigma_y + C\sigma_x - r\sigma_y\right)}{\sigma_y\left(C^2(r-1)^2 - r^2\right)}.$$

The ex-ante expected utility of investor i, which is also the social welfare, can be calculated by plugging b_x and b_y into $E[u_i]$. We will get

$$u_{DE} = \frac{C(r-1)^2\left(C^3(r-1)^2\left(\sigma_x^2+\sigma_y)^2\right)+2C^2\left(r^2-1\right)\sigma_x\sigma_y - C(r-2)r\left(\sigma_x^2+\sigma_y^2\right)-2r^2\sigma_x\sigma_y\right)}{2\left(r^2-C^2(r-1)^2\right)^2}. \tag{A.2}$$

A.8.2 The second step

I now show that no investor will deviate from choosing option DE.

Consider investor i's deviation to option SU. It is obvious that following this deviation, he will devote all his capacity to minimize σ^2. Hence, $\sigma^2 = \frac{\sigma_x^2+\sigma_y^2}{C^2-1}$.

Conditional on receiving the signal z_i, investor i will choose the investment level as

$$k_i = (1-r)E[\theta|z_i] + rE[K|z_i] = b_z z_i,$$

where $b_z = (1-r)\frac{\sigma_x^2+\sigma_y^2}{\sigma_x^2+\sigma_y^2+\sigma^2} + r\frac{b_x\sigma_x^2+b_y\sigma_y^2}{\sigma_x^2+\sigma_y^2+\sigma^2}$. His utility from deviation is thus

$$u_{\text{deviation}} = E[Ak_i - \frac{1}{2}k_i^2] = \frac{\left(C^2-1\right)(r-1)^2\left(C(r-1)\left(\sigma_x^2+\sigma_y^2\right)+2r\sigma_x\sigma_y\right)^2}{2\left(r^2-C^2(r-1)^2\right)^2\left(\sigma_x^2+\sigma_y^2\right)}.$$

I now show that $u_{DE} > u_{\text{deviation}}$ under the conditions of Proposition 12.

First, it can be easily verified that when $\frac{\sigma_y}{\sigma_x} > \sqrt{3}$ and $r \in (\frac{C}{C+1}, 1)$, $u_{\text{deviation}} - u_D$ is increasing in r.

Second, because when $\frac{\sigma_y}{\sigma_x} > \sqrt{3}$, $\frac{C^2\sigma_y - \sigma_x}{(C^2-1)\sigma_y} > \frac{C}{C+1}$, we know that when $r \in (\frac{C^2\sigma_y-\sigma_x}{(C^2-1)\sigma_y}, 1)$, $u_{\text{deviation}} - u_D$ is increasing in r. Hence, $u_{\text{deviation}} - u_D$ takes its maximum value at $r = 1$. Because $u_{\text{deviation}} - u_D|_{r=1} = 0$, we know $u_{\text{deviation}} < u_D$, $\forall\ r \in (\frac{C^2\sigma_y-\sigma_x}{(C^2-1)\sigma_y}, 1)$, suggesting that deviating to option SU yields a strictly lower utility.

Furthermore, it can be checked that when $r \in (\frac{C^2\sigma_y-\sigma_x}{(C^2-1)\sigma_y}, 1)$ and when $\frac{\sigma_y}{\sigma_x} > \sqrt{3}$, the interior solution of σ_1 and σ_2 is guaranteed in the first step. This completes the second step. Q.E.D.

A.9 Proof of Proposition 13

I will prove the claim that $W_{SU} > W_{DE}$ holds whenever we have interior solutions $b_x > 0$ and $b_y > 0$. That is, $W_{SU} > W_{DE}$ holds whenever the parameters are such that investors will pay attention to both details following their choices of option DE. The proof for the corner-solution case is similar and is therefore omitted.

In fact, taking their difference, we have

$$W_{SU} - W_{DE} = \frac{1}{2}C(r-1)^2(\frac{C\left(C^2-1\right)\left(\sigma_x^2+\sigma_y^2\right)}{\left(r-C^2(r-1)\right)^2}$$
$$+\frac{C^3\left(-(r-1)^2\right)\left(\sigma_x^2+\sigma_y^2\right)-2C^2\left(r^2-1\right)\sigma_x\sigma_y+C(r-2)r\left(\sigma_x^2+\sigma_y^2\right)+2r^2\sigma_x\sigma_y}{\left(r^2-C^2(r-1)^2\right)^2}).$$

In the above expression, the first term, $\frac{C\left(C^2-1\right)\left(\sigma_x^2+\sigma_y^2\right)}{\left(r-C^2(r-1)\right)^2}$, is obviously always positive. It can be easily checked that the second term is also always positive as long as we have interior solutions ($b_x > 0$ and $b_y > 0$, or equivalently, $\sigma_1^2 < +\infty$ and $\sigma_2^2 < +\infty$). Hence, $W_{SU} - W_{DE} > 0$. Q.E.D

A.10 Proof of Proposition 14

We know that in DE_BOTH, $K \propto \theta$. Therefore, showing that DE_BOTH is an equilibrium is equivalent to showing that dividing attention evenly between the two details is investors' optimal attention-allocation strategy to predict θ.

In fact, if investors choose to pay attention to both the summary and details, they will make decisions based on the following three signals:

$$z_i = \theta + \gamma + \epsilon_i$$

$$x_i = \theta_x + \epsilon_{1i},$$

$$y_i = \theta_y + \epsilon_{2i},$$

where $\epsilon_i \sim N(0, \sigma^2)$, $\epsilon_{1i} \sim N(0, \sigma_1^2)$, and $\epsilon_{2i} \sim N(0, \sigma_2^2)$. By symmetry, we know that $\sigma_1 = \sigma_2$ must hold in equilibrium.

For any investor i, his utility maximization problem is thus

$$\max_{\sigma_1^2, \sigma^2} E[u_i] = E[\theta k_i - \frac{1}{2} k_i^2]$$

$$\text{subject to } I(\{x_i, y_i, z_i\}, \{\theta_x, \theta_y, \theta + \gamma\}) \leq \kappa$$

Hence, proving that investors will choose $\frac{dE[u_i]}{d\sigma^2 s} > 0$, or equivalently, $\frac{dE[u_i]}{d\sigma_1^2} < 0$ (meaning no attention will be paid to the summary) will suffice for the proposition.

Substituting the entropy constraint $I(\{x_i, y_i, z_i\}, \{\theta_x, \theta_y, \theta + \gamma\}) \leq \kappa$ into investors' objective, we can write down investors' expected utility as a function of σ_1^2:

$$E[u_i] = \frac{2\left(C^2 - 1\right)\sigma_1^2 + C^2 \sigma_\gamma^2 - 2}{C^2(\sigma_1^2(\sigma_\gamma^2 + 2) + \sigma_\gamma^2)}$$

It can be easily checked that $\frac{dE[u_i]}{d\sigma_1^2} < 0$ when $C > C^*$. Therefore, when $C > C^*$, any investor will choose to ignore the summary. Q.E.D.