Properties of Optimal Accounting Rules in a Signaling Game[*]

Xu Jiang

Fuqua School of Business, Duke University

xu.jiang@duke.edu

Ming Yang

Fuqua School of Business, Duke University

ming.yang@duke.edu

July, 2016

Abstract

We characterize the properties of optimal accounting rules in a signaling game. An impatient firm sells shares to competitive investors. The firm can signal its private information about the fundamental by retaining a fraction of the shares. In addition, the firm can commit to disclosing information according to a set of accounting rules chosen ex ante. Information disclosure reduces signaling cost so that perfect disclosure is optimal. When perfect disclosure is impossible, the optimal accounting rule features a lower bound and a summary statistic of the fundamental. The interpretation of the lower bound is consistent with accounting conservatism, and the statistic summarizes the information most relevant to the firm's valuation. The justification for accounting conservatism relies on the existence of information asymmetry and the infeasibility of perfect accounting disclosure. This is consistent with the conjecture of LaFond and Watts (2008) that information asymmetry calls for accounting conservatism.

1 Introduction

This paper studies the properties of optimal accounting rules that a firm would like to choose before selling shares to outside investors. The firm is impatient relative to the investors, so selling shares may benefit both parties. This transaction, however, may be impeded by adverse selection. The firm then signals its private information through retaining a fraction of the shares, and the signaling cost depends on the degree of information asymmetry. The firm may benefit from committing to disclosing information according to a set of accounting rules, which is specified before the arrival of the firm's private information. We investigate the optimal accounting rule and find its features consistent with some features of the financial reporting rules used in practice, in particular, accounting conservatism.

We study the optimal accounting rule in the context of an initial public offering (IPO). Costly signaling in the context of IPO is an important topic in the finance literature (see, for example, Leland and Pyle 1977; DeMarzo and Duffie 1999). Disclosure helps alleviate signaling cost by reducing information asymmetry. In general, accounting standards feature accounting rules that make part of the signals more informative than the rest. For example, accounting rules impose more stringent verification requirements for recognizing gains than losses. As suggested by, for example, Gigler et al. (2009), this differential verification requirement leads to reported gains being more informative than reported losses or, equivalently, good realization of signals being more informative than bad realization of signals, which we denote as "qualitative" properties of signals. In other words, "quantitative" properties of signals refer to the *overall* informativeness of signals, whereas "qualitative" properties of signals concern how the overall informativeness is *distributed*, for example, whether good realization of signals are more informative than bad ones or vice versa.[1] Previous studies, relying on additive information structures, can only address quantitative properties, that is, whether more or less disclosure is desirable.[2] This is because the

---

[1] This notion of informativeness relates to but is not exactly the same as the notion of timeliness. Gigler et al. (2009) argue that bad news is recognized more promptly but often prematurely because of less stringent recognition criteria. Good news, on the contrary, is recognized less promptly but more accurately because sufficiently convincing evidence is required for recognition of good news. This implies that good news is more informative than bad news, which is assumed by Gigler et al. (2009) but emerges endogenously as one of the properties of the optimal accounting rules in our model. Timeliness itself is an interesting topic, and we leave it for future research.

[2] An information structure $f(x|\theta)$ is a conditional distribution of signal x given fundamental θ. A widely used example in the accounting literature, which we call an "additive information

distribution of the noise is the same regardless of the underlying fundamentals, resulting in good news that is as informative as bad news. Our model captures the qualitative features by allowing more flexible information structure. [3]

Since the firm is less patient, selling shares to outside investors creates gains for both parties. Before deciding the fraction of shares to sell, the firm receives some private information. Hence the fraction of shares sold becomes a signal of its private information. As is standard in the literature (for example, DeMarzo and Duffie 1999), firms with more favorable private information retain more shares, which is costly as it results in inefficient allocations. In order to reduce such inefficiency, before the arrival of the private information, the firm can commit to disclosing information according to a pre-specified accounting rule. This benefits the firm because it reduces the information asymmetry between the two parties.

Our main results are as follows. Intuitively, it is optimal to commit to disclosing all private information of the firm since it eliminates the information asymmetry between the two parties. In most practical situations, however, full disclosure is impossible as it requires disclosing infinite amount of information, according to the information theory. We therefore study the optimal accounting rule subject to a constraint on the amount of information. In order to address the qualitative features, we do not impose any parametric form, such as Gaussian or Poisson, on feasible disclosures. With this flexibility in choices, we find that the optimal accounting rule is characterized by disclosing a lower bound and a summary statistic of the fundamental. While it may appear that the results in our paper apply only to an initial public offering (IPO) setting, what is crucial to our results is the presence of a signaling game. In our IPO setting, the firm uses the fraction of shares retained to signal its private information. In other settings, so long as the firm can signal its private information through other means and the purpose of disclosure is to reduce information asymmetry, the qualitative properties of our results will remain valid.

The interpretation of the lower bound is consistent with accounting conservatism. Accounting conservatism, as argued in Watts (2003), is defined as requiring stricter criteria for recognizing gains than losses. This results in two effects. First, good news is more informative than bad news, as argued by Gigler

---

structure," is $x = \theta + \eta$, where η is a Gaussian noise. Since each x generates a posterior distribution about θ, an information structure is a mapping from each realization of θ to a distribution of x, or, equivalently, distribution of distributions of θ. See section 4 for more details.

[3] See below for more detailed discussions of "qualitative" versus "quantitative" properties.

et al. (2009). Second, the book values on the balance sheets are on average smaller than the true values of the (net) assets of a firm, or equivalently, the balance sheets represent lower bounds of the firm value, as evidenced by various accounting rules. One example is impairment accounting rule or the so-called "lower of cost or the market" rule. Under such rule, quite a few categories of assets (e.g., inventory, tangible assets, goodwill) must be tested periodically to see whether the fair value is lower than recorded acquisition cost (subject to depreciation or amortization). If so, the asset must be recorded on the balance sheet using the lower fair value, that is, the asset has to be written down. In contrast, if the fair value is higher than the recorded acquisition cost, the firm cannot record the higher value on its balance sheet and must record the asset at its book value, that is, the firm cannot write-up the value of the asset.[4] This "lower of cost or market" rule results in the balance sheet values being lower bounds of firm values.[5] In this sense, our result of disclosing a lower bound is consistent with the conservatism principle embedded in many accounting rules and is derived in a fairly general setting with information asymmetry being the only friction.

The summary statistic can be interpreted as a financial report that provides a summary of firm value conditional on it verifiably exceeding the lower bound. In practice, most accounting rules require that auditors issue an opinion regarding whether the firm can continue as a going concern, and most, if not all, firms' financial reports indicate that the presented numbers are based on the assumption that those firms will continue as going concerns. The lower bound in our setting can be interpreted as providing a lower bound of the expected firm value if the firm is liquidated today (i.e., the liquidation value cannot be lower than this lower bound). The summary statistic then reflects the expected firm value based on the firm continuing as a going concern and its value exceeding this lower bound. We caution against inferring too much from the exact form of the summary statistic, as the exact form varies in different signaling games.

The intuition of the main results is as follows. In the presence of information

---

[4] Under International Financial Reporting Standards (IFRS), in certain circumstances, a write-up after a write-down is allowed. However, the asset's recorded value cannot exceed its original acquisition cost.

[5] The accounting treatments for research expenditures and inventory provide different lower bounds, with zero for research expenditures and the lower of cost or market for inventory. While our results do not speak directly to why different lower bounds exist for different classes of assets, complete expensing of research expenditures can be interpreted as a special case of our more general results when disclosure is completely uninformative. We discuss this issue and compare our results to the "lower of cost or market" principle in Section 4.4.

asymmetry, outside investors are concerned that firms with low fundamentals mislead them by disclosing overly optimistic information. This creates incentives for firms with more favorable information to signal by retaining a larger fraction of shares, which is inefficient. The best way to disclose is thus to reassure outside investors that the firm's expected value cannot be lower than a verifiable lower bound. In contrast, disclosing an upper bound does not help since it does not give outsiders such reassurance. Besides the lower bound, the summary statistic is disclosed in order to economize the information resource. It summarizes the relevant information that is sufficient to determine the firm's expected value under any posterior belief.

We make two contributions to the accounting literature. First, to the best of our knowledge, our paper is the first to study the optimal qualitative properties of accounting rules. DeMarzo and Duffie (1999) study the problem of security design in a similar signaling game. Their focus is security design whereas our focus is the accounting rule design. Gigler et al. (2009) vary accounting conservatism by changing the relative information content of good earnings versus bad earnings. They show that accounting conservatism could result in inefficient debt contracting. Our way of modelling information disclosure is more flexible than theirs, and we show that accounting conservatism can be welfare-enhancing.

It is worth elaborating more on our terminology of "*qualitative*" and "*quantitative*" properties of information systems because those two terms are widely used in the accounting literature to refer to different concepts.[6] The two terms used in our paper are defined in a growing, new economics literature following the seminal work of Kamenica and Gentzkow (2011). Consider a widely used example in the accounting literature where accounting report x is a noisy signal about the unknown firm value θ, i.e., $x = \theta + \eta$, where η is a Gaussian noise with mean 0 and variance σ^2. A typical way of modelling the cost of improving the reporting quality is to associate the variance σ^2 with a cost $C(\sigma^2)$, which is decreasing in σ^2. This implies that the firm sprinkles the resources equally on every part of the distribution. When σ^2 is reduced, *every* realization of x becomes equally more accurate. The shape of the distribution of η does not vary with θ. We call this "quantitative" because it focuses on the overall informativeness of the accounting rule, which is captured by the quantity of the parameters. In particular, it is the variance of the noise in the aforementioned

---

[6] We thank the referee for suggesting the discussion on this aspect.

example. Our way of modelling information disclosure does not impose any structure on the distribution of η. It is "qualitative" in the sense that we allow the distribution of η to vary with θ. To give a loose analogy, consider a city with a limited budget for road maintenance deciding how to spend the money. The "quantitative" approach assumes that the city divides the money equally on each section of the road. In contrast, the "qualitative" approach allows the city to optimally spend more money repairing the worst section of the road.

Secondly, by interpreting the optimal properties of accounting rules in accounting standard setting, we provide results that are consistent with arguably one of the most important properties of financial reporting rules: accounting conservatism. Conservatism, as reflected in the "lower of cost or market" rule or more strict recognition criteria for gains versus losses, often serves as a starting point in studying other accounting issues. For example, Beyer (2013) examines cost of capital and debt contract efficiency, Burkhardt and Strausz (2009) and Caskey and Hughes (2012) studies asset substitution. We adopt the definition of accounting conservatism as requiring more stringent verification for gains than for losses. We show that the optimal accounting rules in our setting generate outcomes that are consistent with this definition.

The rest of the paper is organized as follows. The next section further discusses the related literature. Section 3 introduces our model and briefly reviews the standard results of signaling games. Section 4 discusses how to model disclosure and how to measure information. In particular, we compare our measure to variance reduction, one of the most commonly used measures of uncertainty in the finance and accounting literature. The main results are then presented. Section 5 concludes with discussions on future research. Appendix A contains more background about the concepts of entropy and information, which will be used in our setting. Appendix B contains the proofs.

2 Related literature

First, our paper relates to the disclosure literature. We study how ex ante commitment to ex post disclosure can help reduce the inefficiency caused by signaling. Bhattacharya (1979), Miller and Rock (1985), and Leland and Pyle (1977) document how firms can send (endogenously) costly signals revealing their private information to outside investors in the form of dividends or the fraction of shares retained. The papers acknowledge, but do not consider, whether credible

6

disclosure can help alleviate the signaling cost. Kanodia and Lee (1998) explicitly consider the role of mandatory disclosure in an investment setting and show that endogenously imperfect mandatory disclosure is essential in supporting a signaling-by-overinvesting equilibrium. The information structure of Kanodia and Lee (1998) is cast in a CARA-normal framework. In such a framework, all information signals, including the disclosed ones, are modelled as some true value plus a normally distributed noise. Variations in disclosure are thus equivalent to variation of the precision or overall informativeness in a Gaussian distributed world. While CARA-normal framework is widely used in the literature to address how much information should be disclosed in strategic settings, as reviewed in, for example, Dye (2001), Verrecchia (2001), Kanodia (2006), Beyer et al. (2010), and Stocken (2013), it does not fit our purpose here. The distribution of noise is constant across the states and thus is too rigid to capture the qualitative properties of accounting rules such as conservatism. This has been noticed by some critics who argue that the theoretical constructs in accounting theory models are devoid of accounting content (e.g., Dye 2001, Demski and Sappington 1990).[7] In this paper, we allow for more general forms of information structures and are able to focus on the qualitative properties of optimal accounting rules with accounting implications.

Secondly, our paper relates to the vast literature on conservatism. In addition to the aforementioned papers, quite a few recent studies justify conservatism in various settings. Watts (2003), Guay and Verrecchia (2006), Chen et al. (2007), Gox and Wagenhofer (2009), and Gao (2013a) document the desirability of conservative accounting in different debt-contracting settings, while Caskey and Laux (2013) justify conservatism as strengthening the governance role of the corporate board, and Bertomeu and Magee (2015) rationalize conservatism in a model of political influence. More recently, Armstrong et al. (2014) argue that asymmetric reporting is optimal, as risk-averse investors in a CAPM setting benefit more from reducing uncertainty in bad states. On the other hand, Gigler et al. (2009) cast doubt on the benefits of conservatism on debt-contracting efficiency, while Bertomeu et al. (2013), Gao (2013a), and Li (2013) incorporate factors such as managerial compensation, earnings management or renegotiation cost and suggest the possibility of an interior degree of conservatism. LaFond and Watts (2008) and Zhang (2008) provide empirical evidence that conservative accounting benefits equityholders and debtholders,

---

[7] We thank the referee for making this point.

respectively. Analytically, our paper demonstrates the desirability of accounting conservatism for equityholders and Gao (2013a) shows similar results for debtholders. The key ingredient of our framework is the signaling problem, which also appears in other circumstances such as seasoned equity offerings and debt financing. Our approach still goes through in those circumstances when the signaling problem is the key issue.

Thirdly, our paper relates to the literature of rational inattention. Sims (2003, 2005) introduce the concept of entropy and mutual information into macroeconomic models in order to capture the informational constraints of the economic agents. This constraint results in price stickiness and other rigidities that are consistent with the empirical findings in the macroeconomic literature. Yang (2015) applies rational inattention in coordination games, focusing on the equilibrium implications of the flexibility of information acquisition rather than the informational constraint. We adopt a similar informational constraint to highlight the qualitative properties in the design of the accounting rules.

Finally, our paper also relates to the economics literature of Bayesian persuasion. For examples of other accounting applications of this approach, see, e.g., Friedman et al. (2015) and Huang (2015). Kamenica and Gentzkow (2011) study a persuasion game which can be viewed as a cheap talk game with commitment. Our paper differs from Kamenica and Gentzkow (2011) in two aspects. First, we study optimal disclosure in a signaling game while Kamenica and Gentzkow (2011) studies a persuasion game. Second, we design the accounting rules subject to an informational constraint which is absent in their model.

3 The model

3.1 Model setup

The game involves a risk-neutral firm and a continuum of competitive risk-neutral investors.[8] The firm is less patient and hence has an incentive to raise cash by selling some percentage of its shares to outside investors.[9] Specifically, the firm discounts future cash flows at rate $\rho \in (0, 1)$, and the outside investors

---

[8]These are standard assumptions in the literature, when the focus is the informational issues rather than risk-sharing issues.

[9]This assumption, while uncommon in the accounting literature, is as innocuous as the assumption in quite a few accounting papers that old investors have to sell the whole firm to new investors before the final cash flow is realized (e.g., Kanodia and Lee 1998, Kanodia, Singh, and Spero 2005). If there is no such exogenous reason for selling shares to outside investors, disclosure will play no role.

do not discount. This could be justified by the situation that the firm may raise capital to invest in a new project with gross return ρ^{-1} or the firm simply has to raise cash to meet some regulatory capital constraint.

The firm's future cash flow is given by $v = \theta + \varepsilon$, where θ represents the firm's private information about the future cash flow and ε represents the residual noise that is independent of θ and has a zero mean.[10] We assume that $\theta \in [\underline{\theta}, \overline{\theta}] \in \mathbb{R}_{++}$, and the investors share a common prior belief $\Pi \in \Delta\left([\underline{\theta}, \overline{\theta}]\right)$. We assume Π is absolutely continuous with respect to the Lebesgue measure and fully supported on $[\underline{\theta}, \overline{\theta}]$. The outside investors know neither θ nor ε. The firm has an informational advantage over the outside investors in that the firm privately learns θ before issuing shares.

The timeline of the model is as follows:

$t = 0$: The firm chooses an accounting rule and commits to disclosing information according to it. Specifically, the accounting rule is modelled as an information structure that specifies the conditional distribution of the accounting report at $t = 1$, given each value of the true state θ. There is a cost associated with disclosing more information, which will be discussed in detail later.

$t = 1$: Nature draws θ according to the common prior Π, and the firm learns θ. A report is published to investors. The investors then form their posterior belief based on Bayes' rule, denoted by G. Since G is determined by the realized report and is itself the information content of the report, G is used to denote both the posterior and the realized report interchangeably in the rest of the paper. After the report is published, the firm chooses to sell a fraction $q \in [0, 1]$ of the shares to outside investors. Thus, q becomes a signal of θ in addition to G and the investors price the firm at

$$p = \mathbf{E}\left[\theta + \varepsilon | q, G\right],$$

which is the price under perfect competition.

$t = 2$: Residual noise ε is realized. The firm and the investors consume their shares of the firm value accordingly.

3.2   Solution to the signalling game

We solve the model by backward induction. First, we study the signaling game after the report G is published.

---

[10]The exact distribution of ε does not matter because of risk-neutrality.

At the end of date 1, if the firm knows θ and sells fraction q at price p, its expected payoff is

$$E[\rho (1 - q) (\theta + \varepsilon) + qp|\theta]$$
$$= q (p - \rho\theta) + \rho\theta .$$

Since the second term has no strategic effect, just let

$$u_f (\theta, q, p) = q (p - \rho\theta)$$

denote the firm's expected payoff at the end of date 1 from selling q fraction of the shares at price p, conditional on θ. Since perfect competition drives investors' expected payoff to zero, the accounting rule that maximizes the firm's welfare maximizes the social welfare. Consequently, the rule that should be chosen by the regulator is the same as the optimal rule adopted by the firm in this setting, similar to Kanodia and Lee (1998).[11]

Given G, the firm's strategy is $q_G : \text{supp}(G) \to [0, 1]$, and the investors' strategy is $p_G : [0, 1] \to [\underline{\theta}, \overline{\theta}]$. The equilibrium of this signaling game is defined as follows.

Definition 1 *A Bayesian-Nash equilibrium of the signaling game is a pair* $\{q_G, p_G\}$ *such that:*

(i) $q_G(\theta, G) \in \arg\max_q u_f (\theta, q, p_G(q, G))$ almost surely.

(ii) $p_G(q_G, G) = E[\theta + \varepsilon|q_G(\theta, G), G]$ almost surely.

Part (i) of definition 1 states that the firm chooses the fraction of shares sold to maximize its expected payoff. Part (ii) says that investors will price the shares at the expected firm value based on q and the disclosure.

The signaling game is essentially the same as that of DeMarzo and Duffie (1999). We summarize their results in our notation in the following lemma.

Lemma 1 *Given posterior G, the signaling game has a unique equilibrium,*

---

which is fully separating[12] and characterized by

$$q_G(\theta, G) = \left[\frac{\widehat{\theta}_G}{\theta}\right]^{\frac{1}{1-\rho}} \tag{1}$$

and

$$p_G(q, G) = \frac{\widehat{\theta}_G}{q^{1-\rho}} \ , \tag{2}$$

where

$$\widehat{\theta}_G = \inf\left(supp\left(G\right)\right) \ . \tag{3}$$

The firm's expected payoff under any given posterior belief G is thus given by

$$
\begin{aligned}
U_f(G) &= \mathbf{E}\left[u_f\left(\theta, q_G(\theta, G), p_G(q, G)\right) | G\right] \\
&= (1-\rho) \cdot \left[\widehat{\theta}_G\right]^{\frac{1}{1-\rho}} \int \theta^{\frac{-\rho}{1-\rho}} dG(\theta) \ .
\end{aligned} \tag{4}
$$

Equation (4) shows that firm's expected payoff in the signaling game is a function of the report G. Given the common prior, choosing an information structure amounts to choosing an ex ante distribution of report G, denoted by Λ. Therefore the firm's ex ante expected payoff is $\int U_f(G)\Lambda(dG)$.

4 Modelling of accounting rules and the main results

In this section, we first argue that we can maximize the firm's expected payoff $\int U_f(G)\Lambda(dG)$ by directly choosing the distribution of posteriors Λ, instead of working on the information structure. Secondly we show a benchmark result that the firm always ranks accounting rules according to Blackwell's ordering, absent informational constraint. We then introduce a measure of information and characterize the optimal accounting rule subject to some informational constraint based on the measure.

---

[12] While it is true that, in our setting, the percentage of shares retained by the firm fully reveals the private information, we can use a setting similar to that of Kanodia and Lee (1998) in the sense that all shares are sold to the investors and firms signal their private information through other channels, e.g., investment. So long as the signaling equilibrium is fully revealing, our results will not change qualitatively.

4.1 A general approach of modelling accounting rules

We consider a general approach of modelling disclosure. In the literature, an accounting rule is often modelled as an information structure, $\Pr(G|\theta)$, where θ is the uncertain fundamental and G denotes a possible report. As discussed before, since a report is characterized by its information content — the induced posterior, we also use G to denote the investors' posterior belief induced by this report. Note that the information structure $\Pr(G|\theta)$, together with the prior Π, determines the ex ante distribution of posteriors Λ, which, according to Bayes' Rule, satisfies that

$$\Pi = \int G\Lambda(dG). \tag{5}$$

In the literature of Bayesian persuasion, a distribution of posteriors, Λ, is said to be Bayesian plausible if equation (5) holds (see, e.g., Kamenica and Genzkow 2011). Moreover, Kamenica and Genzkow (2011) show that any Λ that satisfies Bayesian plausibility condition can be generated by an information structure. Therefore, without loss of generality, we can directly design the ex-ante distribution of posteriors subject to the Bayesian plausibility condition constraints, equation (5).

Let $S(\Pi)$ denote the set of distribution of posteriors that are Bayes plausible with respect to the prior Π. Designing an accounting rule amounts to choosing a member from $S(\Pi)$ where $S(\Pi) = \left\{\Lambda \in \Delta\left(\Delta\left(\left[\underline{\theta},\overline{\theta}\right]\right)\right) : \Pi = \int G \cdot \Lambda(dG)\right\}$. Before proceeding to the main results, we illustrate our approach by relating it to the examples often employed in the accounting literature.

Example 1 *(Binary information structure) Let the prior of θ be $\Pi(\theta)$ on $[1,2]$ with density $\pi(\theta)$. The realization of an accounting signal s can be either high (denoted as H) or low (denoted as L) with conditional probabilities $\Pr(s = H|\theta)$ and $\Pr(s = L|\theta) = 1 - \Pr(s = H|\theta)$. The ex ante probabilities of receiving $s = H$, and $s = L$ are $\lambda_H \equiv \int_1^2 \Pr(s = H|\theta)\pi(\theta)d\theta$ and $\lambda_L \equiv 1 - \lambda_H$, respectively. By Bayes' rule, the investors' posterior belief of θ upon receiving $s = H$ and $s = L$ is given by $G_H \equiv \frac{\Pr(s=H|\theta)\pi(\theta)}{\lambda_H}$ and $G_L \equiv \frac{\Pr(s=L|\theta)\pi(\theta)}{\lambda_L}$, respectively. Our formulation corresponds to a distribution of posteriors. By definition Λ is Bayes plausible, i.e., equation (5) holds.*

Example 2 *(Partition)* [13] *Let the prior of θ be uniform on $[1,2]$. Consider partitioning the state space into $\{[1, 1.3], (1.3, 1.7], (1.7, 2]\}$, which can be interpreted*

---

[13] We discuss this example because in standard cheap talk games (e.g., Crawford and Sobel 1982), the equilibrium information structure is typically a partition.

as an information structure with three signal realizations, H, M, and L, such that

$$\Pr(s = L|\theta) = \begin{cases} 1 & \text{if } \theta \in [1, 1.3] \\ 0 & \text{otherwise,} \end{cases} \quad \Pr(s = M|\theta) = \begin{cases} 1 & \text{if } \theta \in (1.3, 1.7] \\ 0 & \text{otherwise} \end{cases},$$

and $\Pr(s = H|\theta) = \begin{cases} 1 & \text{if } \theta \in (1.7, 2] \\ 0 & \text{otherwise} \end{cases}$ *. Then the ex-ante probabilities of receiving $s = H$, M and L are $\lambda_H = 0.3$, $\lambda_M = 0.4$, and $\lambda_L = 0.3$, respectively. The corresponding beliefs are $G_H = Uniform_{(1.7,2]}$, $G_M = Uniform_{(1.3,1.7]}$, and $G_L = Uniform_{[1,1.3]}$ where $Uniform_T$ denotes a uniform distribution on T. In our approach, this corresponds to a distribution of posteriors:*

$$\Lambda(G) = \begin{cases} \lambda_H & \text{if } G = G_H \\ \lambda_M & \text{if } G = G_M \\ \lambda_L & \text{if } G = G_L \\ 0 & \text{otherwise} \end{cases}.$$

Example 3 *Following the notation of Gigler et al. (2009), $\widetilde{x}$ denotes the uncertain cash flow, $\varphi(\cdot|y, \delta)$ is a posterior belief about $\widetilde{x}$ upon signal realization y, and $\delta > 0$ is a parameter representing accounting conservatism. Then $h(\cdot, \delta)$ represents the distribution of posterior beliefs $\varphi(\cdot|y, \delta)$. In our approach, for each δ, the distribution of posteriors is given by $\Lambda(dG) = \begin{cases} h(y, \delta)dy & \text{if } G = \varphi(\cdot|y, \delta) \text{ for some } y \\ 0 & \text{otherwise} \end{cases}$. In Gigler et al. (2009), the distribution of posterior beliefs can vary only within the one-dimensional family parametrized by δ. In contrast, our approach does not have such restriction.*

4.2 Blackwell's ordering as a benchmark

This subsection establishes a benchmark result showing that the desirability of information structures is consistent with their Blackwell's ordering.

Proposition 1 *Let Λ_1, $\Lambda_2 \in S(\Pi)$ be two information structures such that Λ_2 dominates Λ_1 in the sense of Blackwell's ordering. Then*

$$\mathbf{E}_{\Lambda_2}[U_f(G)] \geq \mathbf{E}_{\Lambda_1}[U_f(G)],$$

where

$$\mathbf{E}_{\Lambda}[U_f(G)] = \int U_f(G) \cdot \Lambda(dG). \tag{6}$$

and the inequality is strict if the Blackwell dominance is strict.

The intuition underlying Proposition 1 is as follows. The friction of our IPO framework stems from the information asymmetry between the firm and the outside investors, which inefficiently limits the scope of the transaction. Since disclosing more information reduces the signaling cost, an information structure that Blackwell dominates the other cannot do worse and will be strictly better if the Blackwell dominance is strict.[14] An implication of this proposition is that the firm always wants to disclose more information to the outside investors. Absent any constraints on disclosure, full disclosure will be the optimal solution. We summarize this implication in the following corollary. The proof is omitted as it directly follows Proposition 1.

Corollary 1 *Absent any information constraint, the firm's optimal information structure Λ^* is characterized by $\Lambda^*(dG) = \begin{cases} \pi(\theta) & \text{if } G = \delta_\theta \text{ for some } \theta \in [\underline{\theta}, \overline{\theta}] \\ 0 & \text{otherwise} \end{cases}$ where δ_θ is the Dirac distribution at point θ and $\pi(\theta)$ is the density function of the prior.*

Corollary 1 provides a potential justification of fair value accounting where managers disclose their best estimates to outside investors to reduce information asymmetry as much as possible. In practice, however, perfect disclosure may be infeasible. In the rest of the paper, we impose some constraints on information structures to capture the infeasibility of perfect disclosure. Our main results concern comparing information structures subject to such information constraints. In addition, such a comparison also makes sense since, according to Proposition 1, it is trivial to compare two information structures when one Blackwell dominates the other. In other words, we are comparing information structures that are of the same level of overall informativeness but may be informative in different aspects. To this end, we introduce a measure of information in the next subsection.

4.3 The measure of uncertainty

Information, by its nature, is the reduction of uncertainty. Measuring information thus calls for a good measure of uncertainty in the first place.

---

[14]Proposition 1 differs from the seminal result of Blackwell's Theorem, which concerns an individual's decision problems, whereas Proposition 1 examines information structures in a signaling game. In principle, it is not obvious that more information in the Blackwell sense is always more desirable in the signaling game.

Uncertainty is measured by entropy in the information theory. Hence the information conveyed by an information structure is measured by the reduction of entropy through using the information structure. Specifically, let $\Lambda \in S(\Pi)$ be an information structure and $G \in \mathrm{supp}(\Lambda)$ be a generic posterior belief resulting from Λ. The posterior uncertainty is measured by the entropy associated with G,

$$h(G) = -\mathbf{E}_G \log g(\theta),$$

where g is the density of G[15] and $\mathbf{E}_G$ denotes the expectation under the posterior belief G. The expected posterior uncertainty is then given by $\mathbf{E}_\Lambda[h(G)]$. Similarly, the prior uncertainty is measured by the entropy associated with the prior belief Π, which is given by

$$h(\Pi) = -\mathbf{E}_\Pi \log \pi(\theta),$$

where $\pi(\theta)$ is the density of Π and $\mathbf{E}_\Pi$ denotes the expectation under prior belief Π. Then the reduction of entropy is given by[16]

$$I(\Lambda) = h(\Pi) - \mathbf{E}_\Lambda[h(G)], \tag{7}$$

which measures the amount of information conveyed by Λ. We illustrate the intuition behind this definition by the following example.

Consider a special case of example 1, where the prior belief, Π, is the uniform distribution on $[1, 2]$. To fix the idea, we consider a family of binary information structures that help us distinguish the event $[\frac{3}{2}, 2]$ from the event $[1, \frac{3}{2})$. Specifically, there are two signal realizations, H and L, resulting in two posterior beliefs, G_H and G_L, respectively. The probability density function of G_H is $g_H^k(\theta) = 2k\theta + (1 - 3k)$, and the probability density function of G_L is $g_L^k(\theta) = 2 - g_H^k(\theta)$, where $k \in [0, 1]$ is the slope of g_H, as shown in Figure 1. In general, when $k > 1$, let $g_H^k(\theta) = \begin{cases} 0 & \text{if } \theta \in [1, \frac{3k-1}{2k}) \\ 2k\theta + (1 - 3k) & \text{if } \theta \in [\frac{3k-1}{2k}, \frac{3k+1}{2k}] \\ 1 & \text{if } \theta \in (\frac{3k+1}{2k}, 2] \end{cases}$,

and $g_L^k(\theta) = 2 - g_H^k(\theta)$. Then we obtain a family of information structures

---

[15] More rigorously, g is the Radon-Nikodym derivative of G with respect to the Lebesgue measure.

[16] Note that $h(\cdot)$ is a convex functional on $\Delta([\underline{\theta}, \overline{\theta}])$. Therefore $I(\Lambda)$ is always nonnegative by equation (7).

parametrized by the slope k, denoted by $\Lambda^k(G) = \begin{cases} \lambda_H & \text{if } G = G_H \\ \lambda_L & \text{if } G = G_L \\ 0 & \text{otherwise} \end{cases}$. It is

straightforward to see that when $k = 0$ both posteriors are exactly the same as the prior and the information structure Λ are completely uninformative, a result consistent with expression (7) where $I^k(\Lambda) = 0$. The two posterior beliefs g_H^k and g_L^k deviate from each other as k increases, which enables a sharper differentiation between the two events $\{\theta : \theta \in [\frac{3}{2}, 2]\}$ and $\{\theta : \theta \in [1, \frac{3}{2})\}$. This suggests a larger amount of information conveyed by information structure Λ^k. We demonstrate how the amount of information varies with k in Figure 2. Finally, when $k \to \infty$, g_H^k and g_L^k become the indicator functions $1_{[\frac{3}{2},2]}$ and $1_{[1,\frac{3}{2})}$, respectively. In this case, the amount of information is $\log 2$, as suggested by the asymptotic line in Figure 2.
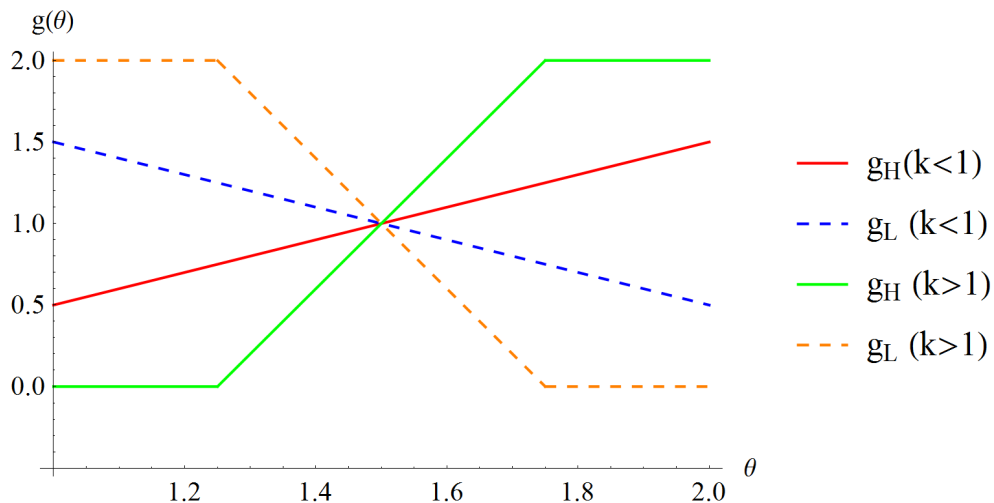

Figure 1: A graphical illustration of posteriors g_H and g_L with different levels of k

In the literature, variance is often used to measure uncertainty. However, it may fail to capture some key ingredients of uncertainty in our context, which are well captured by entropy. The following example highlights this point. Suppose there are $N > 2$ states indexed by $n \in \{1, 2, ..., N\}$ and consider two probability distributions: (i) states 1 and N occur with probability $\frac{1}{2}$, and the rest $N - 2$ states never occur; (ii) all states occur with equal probability, which is $\frac{1}{N}$. The variance associated with probability distribution (i) is
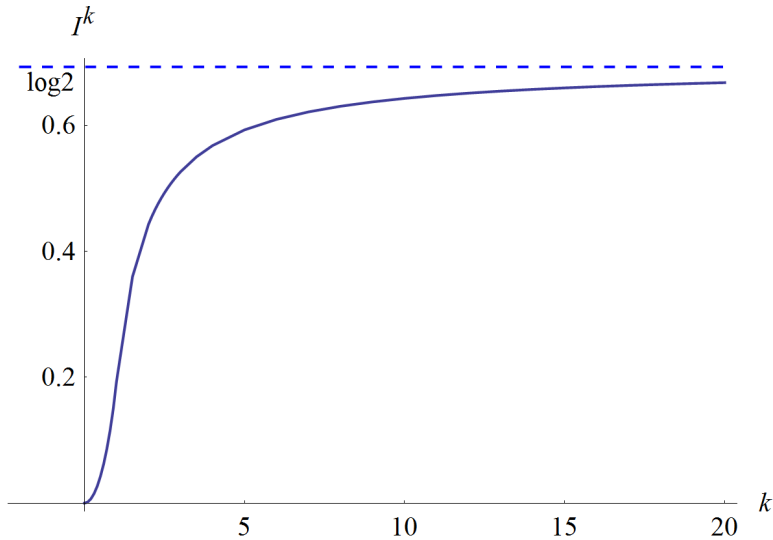
Figure 2: The amount of entropy reduction as a function of k

$var(i) = \frac{1}{2}(1 - \frac{N+1}{2})^2 + \frac{1}{2}(N - \frac{N+1}{2})^2 = \frac{(N-1)^2}{4}$, whereas the variance associated with probability distribution (ii) is $var(ii) = \sum_{m=1}^{N} \frac{1}{N}(m - \frac{N+1}{2})^2 = \frac{(N-1)(N+1)}{12}$. It is clear that $var(i) > var(ii)$. Thus, based on the variance measure, the state should be more uncertain under probability distribution (i) than under probability distribution (ii). This is counterintuitive. To resolve the uncertainty, a decision-maker facing distribution (i) only needs to tell state $n = 1$ from $n = N$; while facing distribution (ii), she needs to tell $\frac{N(N-1)}{2}$ pairs of states apart. In contrast, the entropies associated with probability distribution (i) and (ii) are given by $h(i) = -\frac{1}{2}\log(\frac{1}{2}) - \frac{1}{2}\log(\frac{1}{2}) = \log 2$ and $h(ii) = -\sum_{i=1}^{N} \frac{1}{N}\log(\frac{1}{N}) = \log N$, respectively. It is worth noting that the associated entropy increases in the number of states N, which captures the intuition that, the larger the number of equally probable states is, the more uncertainty we end up with.

4.4 Main results

As discussed in Section 4.2, we impose a constraint,

$$I(\Lambda) \leq \kappa \tag{8}$$

for some $\kappa \geq 0$ to capture the infeasibility of perfect disclosure. The firm's problem is to choose a $\Lambda \in S(\Pi)$ to maximize $E_\Lambda[U_f(G)]$ given by equation (6) defined on page 14 subject to constraint (8). The following proposition illustrates the solution to this optimization problem.

Proposition 2 *For any $\kappa \geq 0$, the firm has a unique optimal accounting rule, in the form of a distribution of posteriors $\Lambda_\kappa^* \in \Delta\left(\Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)\right)$, such that i) $\forall G_1, G_2 \in supp(\Lambda_\kappa^*)$, $\widehat{\theta}(G_1) \neq \widehat{\theta}(G_2)$; ii) $\forall G \in supp(\Lambda_\kappa^*)$, G is the maximum entropy distribution on $\left[\widehat{\theta}(G), \overline{\theta}\right]$, subject to a moment constraint that $\int_{\left[\widehat{\theta}(G), \overline{\theta}\right]} \theta^{\frac{-\rho}{1-\rho}} dG(\theta) = M$, where M is a constant determined by $\widehat{\theta}(G)$ and κ.*

Proposition 2 can be understood as follows. Any posterior belief that results from the optimal accounting rule Λ_κ^* is characterized by its lower bound and a summary statistic. First, any two posterior beliefs G_1, G_2 must differ in their lower bound. Otherwise, by replacing G_1 and G_2 with $\frac{\Lambda(G_1)}{\Lambda(G_1)+\Lambda(G_2)}G_1 + \frac{\Lambda(G_1)}{\Lambda(G_1)+\Lambda(G_2)}G_2$, we create a new accounting rule such that $U_f(G)$ remains the same, but the information constraint becomes strictly slack. The importance of disclosing the lower bound stems from the nature of the signaling game, since the inefficiency (signaling cost) depends on how far away the type is from the lower bound. In practice, the outside investors' main concern is of low types mimicking the high types. Disclosing the lower bound can assure them that the expected firm value cannot be any lower. In our setting, disclosing an upper bound does not help since outside investors are concerned about low types overreporting, and upper bound disclosures will not help alleviate such concerns.

The summary statistic of a posterior G corresponds to the firm's equilibrium expected payoff $U_f(G)$ to minimize the information cost. Given any level of the expected payoff, the firm should choose a posterior G that maximizes the posterior uncertainty, which makes G the maximum entropy distribution that delivers the given level of expected payoff. It is worth noting that, if the specific form of the signaling game changes, the exact form of the summary statistic will change accordingly but the lower bound result remains valid.

We apply the results of Proposition 2 to two illustrative examples.

Example 4 *The accounting rule in example 2, i.e.,*

$$
\Lambda(G) = \left\{
\begin{array}{ll}
\lambda_H & \text{if } G = G_H \\
\lambda_M & \text{if } G = G_M \\
\lambda_L & \text{if } G = G_L \\
0 & \text{otherwise}
\end{array}
\right. ,
$$

can never be optimal because the upper bounds of the support of posteriors G_M and G_L are strictly smaller than $\overline{\theta} = 2$. As a result, G_M and G_L cannot be maximum entropy distributions over $(\widehat{\theta}(G_M), \overline{\theta}]$ and $[\widehat{\theta}(G_L), \overline{\theta}]$, respectively. This contradicts Proposition 2.

Example 5 *Suppose the density of the prior is $\pi(\theta) = \left\{ \begin{array}{ll} \frac{1}{2} & \text{if } \theta \in [1, 1.5) \\ \frac{3}{2} & \text{if } \theta \in [1.5, 2] \\ 0 & \text{otherwise} \end{array} \right.$.*

An accounting rule that is consistent with Proposition 2 is $\Lambda(G) = \left\{ \begin{array}{ll} \frac{1}{2} & \text{if } G = G_H \\ \frac{1}{2} & \text{if } G = G_L \\ 0 & \text{otherwise} \end{array} \right.$

where $G_H = Uniform_{[1.5,2]}$ and $G_L = Uniform_{[1,2]}$. This accounting rule can be operationalized by disclosing an accounting signal s that is either high (denoted as H) or low (denoted as L). Disclosing $s = H$ results in the posterior belief G_H, and disclosing $s = L$ results in the posterior belief G_L. Conditional upon θ, the probability of disclosing $s = H$ is given by $\Pr(s = H|\theta) = \left\{ \begin{array}{ll} 0 & \text{if } \theta \in [1, 1.5) \\ \frac{2}{3} & \text{if } \theta \in [1.5, 2] \end{array} \right.$, and the probability of disclosing $s = L$ is given by $\Pr(s = L|\theta) = 1 - \Pr(s = H|\theta) = \left\{ \begin{array}{ll} 1 & \text{if } \theta \in [1, 1.5) \\ \frac{1}{3} & \text{if } \theta \in [1.5, 2] \end{array} \right.$. The ex ante probabilities of receiving $s = H$ and $s = L$ are $\lambda_H \equiv \int_1^2 \Pr(s = H|\theta)\pi(\theta)d\theta = \frac{1}{2}$ and $\lambda_L \equiv 1 - \lambda_H = \frac{1}{2}$, respectively.

The accounting rule in example 5 exhibits conservatism. Note that disclosing $s = H$ implies that θ is at least 1.5, whereas disclosing $s = L$ implies that θ can be anywhere between 1 and 2. In other words, good news disclosures are highly informative, but bad news disclosures are not as informative. This is consistent with the general interpretation of conservatism as "more stringent criteria for recognizing gains than recognizing losses" (Watts 2003). To see this, note that under this interpretation firms would recognize losses more frequently but also more prematurely, implying that it is quite likely that the true losses would be smaller than the recognized losses. Thus bad news will not be very

informative. In contrast, firms will only recognize gains when there is sufficient evidence, implying that it is very likely that the true gains will equal the recognized gain, making the good news very informative. In contrast, aggressive reporting would imply that the bad news is more informative than the good news, whereas neutral reporting would imply that they are equally informative. The lower bound result can thus only be reconciled with accounting reports generated from conservative reporting rules.

Proposition 2 also sheds light on the accounting rule for research expenditures as the optimal accounting rule is consistent with complete expensing of research expenditures. One of the common reasons for completely expensing research expenditures is that measuring the productive part of the expenditure will be uninformative due to the extremely large amount of uncertainty when the expenditure is incurred. This can be interpreted as $\kappa = 0$ in our model, i.e., no information can be conveyed by the disclosure. The optimal accounting rule would dictate the firm to disclose $\underline{\theta}$. The lower bound in this rule is the absolute lower bound of θ, which is equivalent to putting a zero lower bound on productive assets that can be generated from research expenditures.

The optimal accounting rule also relates to the "lower of cost or market" principle applicable to assets that are easier to measure than, e.g., internally developed intangibles. To the extent that "lower of cost or market" principle implies that the book value of assets on average cannot exceed their true value, lower bound disclosures are consistent with this principle. However, part of the reason for using historical cost to value assets is that it represents the market value for many types of assets at the time of acquiring them. Subsequently, the market value and the true value of those assets may change, generating a dispersion among the true value, the market value, and the historical cost. Thus a complete link between the optimal accounting rule and the "lower of cost or market" principle will require a dynamic model of time-varying fundamentals, which is beyond the scope of this paper. Nevertheless we discuss the similarities and differences between our lower bound result and the "lower of cost or market" principle using the accounting rule based on example 5 in a static framework.

Suppose the historical cost is the acquisition cost of the asset, which has a future value $\theta + \varepsilon$. If we assume the acquisition cost of the asset $I = 1.5$, then the lower bound result can always be interpreted to be consistent with the "lower of cost or market" principle in the following sense. Since the book value is now 1.5, if there is any bad news (e.g., the market value of the asset decreases) indicating that θ will be smaller than 1.5, then the firm will disclose a lower bound of θ that

is smaller than 1.5 (i.e., 1). Using accounting terminology, we can say that the firm writes down the book value of its asset to reflect the adverse news regarding the value of the asset. When there is any good news indicating that θ will be higher than 1.5, the firm can at most disclose a lower bound of θ that is 1.5. This implies that the firm can at most disclose that there is no bad news about the value of the asset, i.e., a write-up is not allowed. Using similar logic, we can argue that the optimal accounting rule is consistent with the "lower of cost or market" principle when $I > 1.5$. However, when $I < 1.5$, our lower bound result differs from the "lower of cost or market" principle. When news is sufficiently good so that θ is sufficiently high, the firm can, with probability λ_H, disclose a lower bound of 1.5, i.e., write-ups occur with certain probability. While this may be literally different from the "lower of cost or market" principle, it is still consistent with the spirit of the principle in that the book value of any asset provides a lower value, on average, than the true value.

The financial reports, in particular the income statement, also provide a summary of firm value conditional on the firm continuing as a going concern. To the extent that the firm's value on average being higher than the balance sheet value implies that the firm is a going concern, the summary statistic then provides more information about the value of the firm based on the firm continuing and its value exceeding this lower bound on average. In particular, it provides information for investors to figure out the posterior distribution of the true value of the firm and thus all the related statistics using this posterior distribution, similar to the real life situation when earnings numbers are used in the valuation models.

As a summary, one can interpret the lower bound of the uniquely optimal accounting rule in our setting as being consistent with accounting conservatism, arguably one of the most important attributes of financial accounting. Our results are derived in a setting where there is information asymmetry between the firm and outside investors that cannot be eliminated by disclosure, without any particular structure imposed on the noise of disclosed signals or accounting rules. The findings thus provide a theoretical justification for the conjecture by LaFond and Watts (2008) that information asymmetry between the insiders and the outside investors generates a demand for conservatism.

As a final remark, to the extent that the balance sheet serves as a lower bound of the firm's net assets on average and net income serves as a summary measure of the value of the firm, we also show that both the balance sheet and the income statement matter for the firm's valuation since both the lower bound and the

summary statistic appear in the firm's expected value under equation (4) and both are essential to support the equilibrium. Ohlson (1995) also shows the value relevance of both the balance sheet and the income statement in an exogenously specified linear information framework with no strategic considerations, whereas we allow any kind of information structure in a strategic setting, and the optimal accounting rule is derived endogenously. Furthermore, instead of showing that the value of the firm is linearly related to book value, our results show that the relation can be quite complex and the relation itself is a function of the accounting rule. Future research may continue along this path and derive some empirically testable valuation models when taking into account the strategic consequences of disclosure.

5   Discussion and concluding remarks

We study the properties of optimal accounting rules in a setting where an impatient firm needs to sell shares to raise immediate cash. The firm possesses information that cannot be credibly conveyed to the outside investors, resulting in costly signaling via the percentage of shares retained by the firm. However, before observing any private information, the firm can choose to commit to an accounting rule that will provide the outside investors a noisy signal of the private information. We show that, so long as the disclosure cannot perfectly reveal the firm's private information, the uniquely optimal accounting rule always consists of 1) disclosure of a lower bound of the expected firm value and 2) a moment of the posterior belief, which, together with the lower bound, completely determines the expected firm value conditional on the disclosure. This optimal accounting rule can be interpreted as being consistent with certain features of the accounting rules that guide firms in financial reporting. In particular, the disclosure of the lower bound is consistent with the conservatism principle embedded in the accounting rules. Our results provide support for conservatism, arguably one of the most important attributes of accounting, in a setting that is particularly relevant for accounting.

Our study is, to the best of our knowledge, the first study of the optimal qualitative properties of accounting information in a systematic way. Previous studies on accounting conservatism (e.g., Chen et al. (2007), Gigler et al. (2009) and Gao (2013a)) also model conservatism as a qualitative property of accounting information that changes the relative informativeness of favorable

versus unfavorable signals. However, because of their focus on the particular attributes of conservatism, their papers do not address the qualitative properties of accounting information in a fairly general way, which is our focus.

In deriving the optimal qualitative properties of accounting information, we adopt the concept of information well established in the information theory. This choice exhibits the virtue of flexibility because it is free to choose any information structure so long as it is Bayes-plausible. As discussed in section 2, we believe that this feature makes it particularly appealing to study accounting rules. Our choice of mutual information also exhibits the virtue of comparability because different information structures can be measured using one number: the reduction of entropy. This measure ensures comparability between arbitrary information structures, which we subsequently use to study optimal accounting rules.

We believe our focus on the qualitative properties of information structure is especially relevant for financial reporting because accounting rules often must trade off between different qualitative properties of information with the impact on the quantitative properties being less straightforward. Conservatism versus aggressiveness is one example, while a principle-based accounting standard versus a rule-based accounting standard is another. From this point of view, our paper can be seen as a first step in a line of future research incorporating more institutional details to generate additional insights related to optimal accounting rules. Specifically, in our model, accounting disclosure is still modelled as a black box. Although we can show that the optimal accounting rule can be interpreted to be consistent with conservatism, we cannot show in more detail how the lower bounds are directly related to accounting conservatism applied to specific accounting measurements. To answer those questions, we need to "open the blackbox of accounting measurement" (Gao (2013b)). This seems to be the natural next step in examining the relationship between the accounting rules and the optimal qualitative properties of information systems.

6   Appendix A

Following Sims (2003, 2006), we measure the quantity of information according to information theory, building on Shannon (1948). Information, by its nature, is the reduction of uncertainty. We first introduce the concept of entropy to measure the uncertainty associated with a random variable. Then the informa-

tion conveyed by an information structure can be easily derived as the difference between the prior entropy and the posterior entropy.

Let X be a discrete random variable distributed with probability weights $p(x)$, $x \in \mathbf{X}$. The Shannon's entropy of X is determined by its distribution, given by

$$H(X) = -\mathbf{E}\left[\log p(x)\right] = -\sum_{x \in \mathbf{X}} p(x) \cdot \log p(x) \ .$$

This functional form is not arbitrary. It is derived axiomatically from information theory and relates closely to real applications like coding and information transmission. The base of the logarithm is not essential as it just changes the unit of entropy. When the base is 2, the unit is called a bit. A single toss of a fair coin has an entropy of one bit. If the base is $e = 2.71828$, the unit is called a nat. One bit equals $\log 2 \approx 0.6931$ nat. Since the base of the logarithm is not essential, we stick to the natural logarithm in this paper.

Let Y be a signal of X. A Bayes-plausible information structure is given by $\{q(y), p(x|y)\}_{y \in \mathbf{Y}}$ such that

$$p(x) = \sum_{y \in \mathbf{Y}} q(y) p(x|y)$$

for all $x \in \mathbf{X}$, where $p(x|y)$ denotes the posterior belief about X after observing $Y = y \in \mathbf{Y}$, and $q(y)$ denote the (marginal) probability of observing $Y = y$. Similarly, the posterior uncertainty, after observing $Y = y$, is measured by posterior entropy

$$H(X|y) = -\sum_{x \in \mathbf{X}} p(x|y) \cdot \log p(x|y) \ .$$

Thus the expected/average posterior entropy is

$$H(X|Y) = \sum_{y \in \mathbf{Y}} q(y) \cdot H(X|y) \ .$$

The information about X conveyed by signal Y is just the difference between the prior and posterior entropies, called mutual information between X and Y, given by

$$I(X|Y) = H(X) - H(X|Y) \ .$$

For a continuous random variable θ with density function $p(\theta)$, the entropy

is defined as

$$h\left(\theta\right) = -\int\limits_{\theta\in\mathbb{R}} p\left(\theta\right)\cdot\log p\left(\theta\right)\cdot d\theta\ .$$

The mutual information of this case can be defined accordingly.

7 Appendix B

Proof of Lemma 1:

Proof. The proof follows the proof of proposition 2 in DeMarzo and Duffie

(1999). Since from equation (1), we know that $p(q) = E[\theta + \varepsilon|q(\theta), G] = E[\theta|\left[\frac{\widehat{\theta}(G)}{\theta}\right]^{\frac{1}{1-\rho}}, G] = \theta$, as $\left[\frac{\widehat{\theta}(G)}{\theta}\right]^{\frac{1}{1-\rho}}$ is monotone with respect to θ for any fixed G. Thus the equilibrium is fully revealing, and what is left to be shown is that $q(\theta) = \left[\frac{\widehat{\theta}(G)}{\theta}\right]^{\frac{1}{1-\rho}}$ maximizes $u_f\left(\theta, q, p\right) = q\cdot(p - \rho\theta)$. The first-order condition with respect to q gives $p - \rho\theta + q\frac{dp}{dq} = 0$. Since p is fully revealing, we have $p = \theta$. Thus we have an ordinary differential equation of $\frac{dp}{p} = \frac{1-\rho}{q}dq$ with the boundary condition $p(1) = \widehat{\theta}\left(G\right)$, since the lowest type has nothing to gain from retaining any of the shares. Solving would give us $p\left(q\right) = \frac{\widehat{\theta}(G)}{q^{1-\rho}}$. Since $p(\theta) = \theta$, we also have $\theta = \frac{\widehat{\theta}(G)}{q^{1-\rho}}$, resulting in $q(\theta) = \left[\frac{\widehat{\theta}(G)}{\theta}\right]^{\frac{1}{1-\rho}}$. Finally, the second-order condition with respect to q is satisfied because of single-crossing properties from Mailath (1987). This concludes the proof. ∎

Before proceeding we first characterize a property of the firm's expected utility that will be used in subsequent proofs.

Lemma 2 $U_f\left(G\right)$ *is convex over* $\Delta\left(\left[\underline{\theta},\overline{\theta}\right]\right)$. *Specifically, it is strictly convex, i.e.,*

$$U_f\left(w\cdot G_1 + (1 - w)\cdot G_2\right) < w\cdot U_f\left(G_1\right) + (1 - w)\cdot U_f\left(G_2\right)$$

if $\widehat{\theta}\left(G_1\right) \neq \widehat{\theta}\left(G_2\right)$, *and* $w\in(0,1)$; *otherwise*

$$U_f\left(w\cdot G_1 + (1 - w)\cdot G_2\right) = w\cdot U_f\left(G_1\right) + (1 - w)\cdot U_f\left(G_2\right)\ .$$

Proof of Lemma 2:

Proof. The case $w = 1$ or 0 is obvious. So we focus on the case $w\in(0,1)$. According to (4), in the signaling equilibrium the firm's expected payoff under

belief $w \cdot G_1 + (1 - w) \cdot G_2$ is

$$U_f \left(w \cdot G_1 + (1 - w) \cdot G_2\right)$$
$$= (1 - \rho) \cdot \left[\widehat{\theta}\left(w \cdot G_1 + (1 - w) \cdot G_2\right)\right]^{\frac{1}{1-\rho}} \int \theta^{\frac{-\rho}{1-\rho}} d\left[w \cdot G_1(\theta) + (1 - w) \cdot G_2(\theta)\right] .$$

If $\widehat{\theta}(G_1) = \widehat{\theta}(G_2)$, then

$$\left[\widehat{\theta}\left(w \cdot G_1 + (1 - w) \cdot G_2\right)\right]^{\frac{1}{1-\rho}} = \left[\widehat{\theta}(G_1)\right]^{\frac{1}{1-\rho}} = \left[\widehat{\theta}(G_2)\right]^{\frac{1}{1-\rho}}$$

and

$$U_f \left(w \cdot G_1 + (1 - w) \cdot G_2\right)$$
$$= w(1 - \rho) \cdot \left[\widehat{\theta}(G_1)\right]^{\frac{1}{1-\rho}} \int \theta^{\frac{-\rho}{1-\rho}} dG_1(\theta) + (1 - w)(1 - \rho) \cdot \left[\widehat{\theta}(G_2)\right]^{\frac{1}{1-\rho}} \int \theta^{\frac{-\rho}{1-\rho}} dG_2(\theta)$$
$$= w \cdot U_f(G_1) + (1 - w) \cdot U_f(G_2) ;$$

otherwise, without loss of generality, let $\widehat{\theta}(G_1) < \widehat{\theta}(G_2)$, then

$$\left[\widehat{\theta}\left(w \cdot G_1 + (1 - w) \cdot G_2\right)\right]^{\frac{1}{1-\rho}} = \left[\widehat{\theta}(G_1)\right]$$

and

$$U_f \left(w \cdot G_1 + (1 - w) \cdot G_2\right)$$
$$= (1 - \rho) \cdot \left[\widehat{\theta}(G_1)\right]^{\frac{1}{1-\rho}} \left[w \cdot \int \theta^{\frac{-\rho}{1-\rho}} dG_1(\theta) + (1 - w) \cdot \int \theta^{\frac{-\rho}{1-\rho}} dG_2(\theta)\right]$$
$$= w(1 - \rho) \cdot \left[\widehat{\theta}(G_1)\right]^{\frac{1}{1-\rho}} \int \theta^{\frac{-\rho}{1-\rho}} dG_1(\theta) + (1 - w)(1 - \rho) \cdot \left[\widehat{\theta}(G_1)\right]^{\frac{1}{1-\rho}} \int \theta^{\frac{-\rho}{1-\rho}} dG_2(\theta)$$
$$< w(1 - \rho) \cdot \left[\widehat{\theta}(G_1)\right]^{\frac{1}{1-\rho}} \int \theta^{\frac{-\rho}{1-\rho}} dG_1(\theta) + (1 - w)(1 - \rho) \cdot \left[\widehat{\theta}(G_2)\right]^{\frac{1}{1-\rho}} \int \theta^{\frac{-\rho}{1-\rho}} dG_2(\theta)$$
$$= w \cdot U_f(G_1) + (1 - w) \cdot U_f(G_2) .$$

This concludes the proof. ∎

Proof of Proposition 1:

Proof. Since Λ_2 dominates Λ_1 in the sense of Blackwell's ordering, there exists an information structure from supp(Λ_2) to supp(Λ_1), expressed as $\Gamma \in \Delta(\Delta(\text{supp}(\Lambda_2)))$, a distribution of probability measures over supp(Λ_2) such

that

$$\Lambda_2 = \int \nu \Gamma (d\nu) \qquad (9)$$

and

$$\text{supp}(\Lambda_1) = \left\{ \int_{\text{supp}(\nu)} G\nu (dG) : \nu \in \text{supp}(\Gamma) \right\} , \qquad (10)$$

where ν denotes a typical member of $\Delta(\text{supp}(\Lambda_2))$. Then

$$
\begin{aligned}
\mathbf{E}_{\Lambda_1}[U_f(G)] &= \int_{\text{supp}(\Lambda_1)} U_f(G) \cdot \Lambda_1(dG) \\
&= \int_{\text{supp}(\Gamma)} U_f \left(\int_{\text{supp}(\nu)} G\nu(dG) \right) \cdot \Gamma(d\nu) \\
&\leq \int_{\text{supp}(\Gamma)} \int_{\text{supp}(\nu)} U_f(G)\nu(dG) \cdot \Gamma(d\nu) \\
&= \int_{\text{supp}(\Lambda_2)} U_f(G) \cdot \Lambda_2(dG) \\
&= \mathbf{E}_{\Lambda_2}[U_f(G)] ,
\end{aligned}
$$

where the second equality follows (10), the inequality follows the convexity of $U_f(\cdot)$, and the third equality follows (9).

Since $U_f(\cdot)$ is strictly convex, the inequality will be strict if Λ_2 strictly dominates Λ_1 in the sense of Blackwell's ordering, i.e., $\text{supp}(\nu)$ contains at least two distinct G.

This concludes the proof. ∎

Proof of Proposition 2:

Proof. The firm's problem is

$$\sup_{\Lambda \in S(\Pi)} \mathbf{E}_{\Lambda}[U_f(G)]$$

$$\text{s.t. } I(\Lambda) \leq \kappa.$$

Note that the objective $\mathbf{E}_{\Lambda}[U_f(G)]$ is linear in Λ, and the domain $\{\Lambda \in S(\Pi) : I(\Lambda) \leq \kappa\}$ is convex. So the optimization problem is well defined. The Lagrangian is

$$L = \mathbf{E}_{\Lambda}[U_f(G)] - \mu_\kappa \cdot I(\Lambda),$$

where μ_κ is the Lagrangian multiplier for $I(\Lambda) \leq \kappa$. Proposition 1 and Corollary 1 imply that the constraint binds; thus $\mu_\kappa > 0$. The firm's optimal information

structure can be solved from the dual problem

$$\sup_{\Lambda \in S(\Pi)} L = \mathbf{E}_\Lambda \left[U_f \left(G \right) \right] - \mu_\kappa \cdot I(\Lambda)$$

for appropriately chosen $\mu_\kappa > 0$. Here μ_κ can be interpreted as the marginal cost of information, and $\mu_\kappa \cdot I(\Lambda)$ is the total information cost incurred. Since

$$I(\Lambda) = h \left(\Pi \right) - \mathbf{E}_\Lambda \left[h \left(G \right) \right]$$

and $h \left(\Pi \right)$ is independent from the choice of information structure Λ, the problem reduces to

$$\sup_{\Lambda \in S(\Pi)} \mathbf{E}_\Lambda \left[U_f \left(G \right) \right] + \mu_\kappa \cdot \mathbf{E}_\Lambda \left[h \left(G \right) \right].$$

Suppose we have two different posteriors: $G_1, G_2 \in \mathrm{supp}(\Lambda_\kappa^*)$ and $\widehat{\theta} \left(G_1 \right) = \widehat{\theta} \left(G_2 \right)$. Let $w \in (0, 1)$ be the relative weight that Λ assigns to G_1, i.e., $w = \frac{\Lambda(G_1)}{[\Lambda(G_1) + \Lambda(G_2)]}$. Then the firm can benefit from combining G_1 and G_2, since Lemma 2 implies that $U_f \left(w \cdot G_1 + (1 - w) \cdot G_2 \right) = w \cdot U_f \left(G_1 \right) + (1 - w) \cdot U_f \left(G_2 \right)$, while less information is disclosed, as $h \left(w \cdot G_1 + (1 - w) \cdot G_2 \right) > w \cdot h \left(G_1 \right) + (1 - w) \cdot h \left(G_2 \right)$. Thus we prove the first part. Suppose G is a posterior in the optimal information structure, i.e., $G \in \mathrm{supp}(\Lambda_\kappa^*)$. To minimize information cost, $G \in \Delta \left(\left[\underline{\theta}, \overline{\theta} \right] \right)$ must maximize $h \left(G \right)$ subject to the constraint $U_f \left(G \right) = \int_{\left[\widehat{\theta}(G), \overline{\theta} \right]} \theta^{\frac{-\rho}{1-\rho}} \, dG \left(\theta \right)$. In other words, G is the maximum entropy distribution over $\left[\widehat{\theta} \left(G \right), \overline{\theta} \right]$ with respect to the moment $\int_{\left[\widehat{\theta}(G), \overline{\theta} \right]} \theta^{\frac{-\rho}{1-\rho}} \, dG \left(\theta \right)$. This concludes the proof. ∎

References

Akerlof, G.A., 1970. The market for "lemons": quality uncertainty and the market mechanism. *The Quarterly Journal of Economics* 84(3), 488–500.

Armstrong, C.S., Taylor D.A., Verrecchia R.E., 2014. Asymmetric reporting. University of Pennsylvania, working paper.

Bertomeu, J., Darrough, M., Xue J., 2013. Agency conflicts, earnings Management, and conservatism. Baruch College, working paper.

Bertomeu, J., Beyer, A., Dye, R.A., 2011. Capital structure, cost of capital, and voluntary disclosures. *The Accounting Review* 86(3), 857–886.

Beyer, A. 2013. Conservatism and Aggregation: The effect on cost of equity capital and the efficiency of debt contracts. Stanford University, working paper.

Beyer, A., Kohen, D.A., Lys, T.Z., Walther, B.R., 2010. The financial reporting environment: review of the recent literature. *Journal of Accounting and Economics* 50(2–3), 296–343.

Bhattacharya, S., 1979. Imperfect information, dividend policy, and "the bird in the hand" fallacy. *The Bell Journal of Economics* 10(1), 259–270.

Burkhardt, K., Strausz, R., 2009. Accounting transparency and the asset substitution problem. *The Accounting Review* 84(3), 689–712.

Caskey, J., Hughes, J.S., 2011. Assessing the impact of alternative fair value measures on the efficiency of project selection and continuation. *The Accounting Review* 87(2), 483–512.

Caskey, J., Laux, V., 2013. Effects of corporate governance and managerial optimism on accounting conservatism and manipulation. University of Texas at Austin, working paper.

Crawford, V., Sobel, J., 1982. Strategic information transmission. *Econometrica* 50(6), 1431–1451.

Chen, Q., Hemmer T., Zhang Y., 2007. On the relation between conservatism in accounting standards and incentives for earnings management. *Journal of Accounting Research* 45(3), 541–565.

Cover, T.M., Thomas J.A., 1991. *Elements of information theory*, second edition. New York: Wiley-Interscience.

DeMarzo, P., Duffie D., 1999. A liquidity-based model of security design. *Econometrica* 67(1), 65–99.

Demski, J.S., Sappington, D.E.M., 1990. Fully revealing income measurement. *The Accounting Review* 65 (2), 363–383.

Dye, R.A. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23(1), 123–145.

Dye, R.A. 2001. An evaluation of "essays on disclosure" and the disclosing literature in accounting. *Journal of Accounting and Economics* 32(1), 181–235.

Financial Accounting Standard Board, 2010. Conceptual framework for financial reporting—Chapter 1, the objective of general purpose financial reporting, and chapter 3, qualitative characteristics of useful financial information (a replacement of FASB Concepts Statements No. 1 and No. 2).

Friedman, H.L., Hughes, J.S., and B. Michaeli, 2015. Bayesian persuasion in the presence of discretionary disclosure. University of California at Los Angeles, working paper.

Gao, P., 2010. Disclosure quality, cost of capital and investors' welfare. *The Accounting Review* 85(1), 1–29.

Gao, P., 2013a. A measurement approach to conservatism and earnings management. *Journal of Accounting and Economics* 55(2), 251–268.

Gao, P., 2013b. A two-step representation of accounting measurement. *Accounting Horizons* 27(4), 861–866.

Gigler, F.B., Hemmer, T., 2001. Conservatism, optimal disclosure policy, and the timeliness of financial reports. *The Accounting Review* 76(4), 471–493.

Gigler, F.B., Kanodia, C., Sapra, H., and Venugopalan, R., 2009. Accounting conservatism and the efficiency of debt contracts. *Journal of Accounting Research* 47(3), 767–797.

Gox, R.F., Wagenhofer, A., 2009. Optimal impairment rules. *Journal of Accounting and Economics* 48(1), 2–16.

Grossman, S.J., 1981. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* 24(3), 461–483.

Guay, W., Verrecchia, R.E., 2006. Discussion of an economic framework for conservative accounting and Bushman and Piotroski (2006). *Journal of Accounting and Economics* 42(1–2), 149–165.

Huang, Z., 2015. Optimal disclosure strategy with investor information acquisition. Duke University, working paper.

Jovanovic, B., 1982. Truthful disclosure of information. *The Bell Journal of Economics* 13(1), 36–44.

Jung, W., Kwon, Y.K., 1988. Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research* 26(1), 146–153.

Kamenica, E., Gentzkow, M., 2011. Bayesian persuasion. *American Economic Review* 101(6), 2590–2615.

Kanodia, C. 2006. Accounting disclosure and real effects. *Foundations and Trends in Accounting* 1(3), 167–258.

Kanodia, C., Lee, D., 1998. Investment and disclosure: the disciplinary role of periodic performance reports. *Journal of Accounting Research* 36(1), 33–55.

Kanodia, C., Singh, R., Spero, A.E., 2005. Imprecision in accounting measurement: can it be value enhancing? *Journal of Accounting Research* 43 (3), 487–519.

Leland, H.E., Pyle, D.H., 1977. Informational asymmetries, financial structure, and financial intermediation. *The Journal of Finance* 32(2), 371–387.

LaFond, R., Watts., R.L., 2008. The informational role of conservatism. *The Accounting Review* 83(2), 447–478.

Li, J., 2013. Accounting conservatism and debt contracts: efficient liquidation and covenant renegotiation. *Contemporary Accounting Research* 30(3), 1082–1098.

Mailath, G., 1987. Incentive compatibility in signaling games with a continuum of types. *Econometrica* 55(6), 1349–1365.

Milgrom, P.R., 1981. Good news and bad news: representation theorems and applications. *The Bell Journal of Economics* 12(2), 380–391.

Miller, M.H., Rock, K., 1985. Dividend policy under asymmetric information. *The Journal of Finance* 40(4), 1031–1051.

Plantin, G., Sapra H., Shin, H.S., 2008. Marking-to-market: Panacea or Pandora's box? *Journal of Accounting Research* 46(2), 435–460.

Shannon, C.E., 1948. A mathematical theory of communication. *Bell System Technical Journal* 27, 379–423.

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

Sims, C.A., 2005. Rational inattention: a research agenda. Deutschen Bundesbank Research Center Discussion paper Series 1:Economic Studies 2005, 34.

Stocken, P., 2012. Strategic accounting disclosure. *Foundations and Trends in Accounting* 7(4), 197–291.

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

Verrecchia, R.E., 2001. Essays on disclosure. *Journal of Accounting and Economics* 32(1), 97–180.

Watts, R.L., 2003. Conservatism in accounting part I: Explanations and implications. *Accounting Horizons* 17(3), 207–221.

Yang, M., 2015. Coordination with flexible information acquisition. *Journal of Economic Theory,* 158, 721-738.