Properties of Optimal Accounting Rule 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

Abstract

We characterize the properties of optimal accounting rule 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 signaling game. 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.

*We thank Jeremy Bertomeu, Anne Beyer, Pingyang Gao, Wayne Guay (the editor), Nisan Langberg, Ivan Marinovic, Brett Trueman, an anonymous referee and participants at the 2013 Minnesota-Chicago Accounting Theory Conference and 2014 Financial Accounting and Reporting Midyear Conference for their helpful comments and suggestions. We appreciate Tim Gray for his excellent editorial assistance and Fuqua School of Business for financial support.

1 Introduction

This paper studies the properties of optimal accounting rule 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.[1] 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 rule, which is specified prior to 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 initial public offering (IPO) and highlight its qualitative properties.[2] 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 disclosure 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 in, e.g., 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.[3] In other words, "quantitative" properties of signals

[1] In the model, the firm's impatience is characterized by a low discount factor. In practice, the firm's impatience can be interpreted as being hit by a liquidity shock or having access to better investment opportunities. See DeMarzo and Duffie (1999) for further justifications of the assumption that the firm is more impatient than outside investors.

[2] While it may appear that the results in our paper apply only to an initial public offering (IPO) setting, what is crucial for 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, our results will remain qualitatively the same.

[3] This notion of informativeness is related to, but not exactly the same as the notion of timeliness. Gigler et. al. (2009) argue that bad news is recognized more timely but often prematurely because of less stringent recognition criteria. Good news, on the contrary, is recognized less timely 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 in Gigler et. al. (2009) but emerges endogenously as one of the properties of the optimal accounting rule in our model. Timeliness itself is an interesting topic and we leave it for future research.

refer to the *overall* informativeness of signals whereas "qualitative" properties of signals concern how the overall informativeness is *distributed*, e.g., whether good realization of signals are more informative than bad realization of signals or vice versa. Previous studies, relying on additive information structures, can only address quantitative properties, that is, whether more or less disclosure is desirable.[4] This is because the distribution of the noise is the same regardless of the underlying fundamentals, resulting in good news as informative as bad news. Our model captures the qualitative features in our model by allowing more flexible information structure thus more general distribution of noise.

Since the firm is less patient, selling shares to outside investors creates gains for both parties in our setting. 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 retains more shares, which is costly and 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 disclosure 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, which is measured by the reduction of entropy.[5] 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.

The interpretation of the lower bound is consistent with accounting conservatism. Accounting conservatism, as argued in Watts (2003), is defined as

[4] 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 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.

[5] For readers who are interested in the microfoundations of information theory, please refer to Shannon (1948).

requiring stricter criteria for recognizing gains than losses. This results in two effects. First, good news is more informative than bad news, as argued in Gigler 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.[6] One example is the impairment accounting rule or the so-called "lower of cost or the market" rule. Under such an accounting rule, quite a few categories of assets (e.g., inventory, tangible assets, goodwill) have to be tested periodically to see whether the fair value[7] is lower than recorded acquisition cost (subject to depreciation or amortization). If the fair value is lower, then the asset has to be recorded on the balance sheet using the lower fair value, that is, the asset has to be written down. However, if the fair value is higher than the recorded acquisition cost, the firm cannot record the higher value on its balance sheet and has to record the asset at its book value, that is, the firm cannot write-up the value of the asset.[8] The "lower of cost or market" rule results in the balance sheet values being lower bounds of firm values.[9] 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.[10]

[6] From now on, all discussions of the lower bound of the firm value refer to the firm value exceeding the lower bound on average.

[7] When there is an active and orderly market for the asset, fair value is equal to market value.

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

[9] 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 on 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 provide some discussion of this issue as well as the similarities and differences between our results and the "lower of cost or market" principal in Section 4.4.

[10] The model in this paper illustrates the desirability of conservative accounting for shareholders. While it has been argued that conservative accounting is desirable for debt contracting and debtholders (e.g., Watts 2003, Zhang 2008), LaFond and Watts (2008) provide evidence that conservatism is also desirable for equityholders in the presence of information asymmetry. Since the crucial assumption of our model is the existence of a signaling game, we believe that our results will continue to hold with creditors so long as there is a signaling game between the firm and creditors. Nevertheless, we believe that it is beneficial to theoretically document the desirability of conservative accounting rules for equityholders as well, given the empirical evidence in LaFond and Watts (2008) and that Gao (2013a) already shows analytically the desirability of accounting conservatism for debtholders. Finally, given the pervasive evidence, both theoretical and empirical, that accounting conservatism is good for debtholders, showing that accounting conservatism is also good for shareholders will strengthen the argument in favor of accounting conservatism.

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 expected firm value based on the firm continuing and its value exceeding this lower bound.[11]

The intuition to disclose the lower bound is as follows. In the presence of information asymmetry, outsiders are most 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 of disclosure to reduce such inefficiency is thus to reassure outside investors that the firm's expected value cannot be lower than a verifiable lower bound. Note that disclosing an upper bound does not help since it cannot give outsiders such reassurance.

The summary statistic is sufficient to determine the firm's expected value, given any posterior distribution. Thus, to save the information resource, it is optimal to just disclose the summary statistic. More specifically, the lower bound only specifies the support for the posterior distribution of firm value. Since information is costly, to economize on the use of information, it would be optimal to disclose only the summary statistic.

Our paper makes two contributions to the accounting literature. First, to the best of our knowledge, our paper is the first study on the optimal qualitative properties of accounting rules in a standard signaling game. Our paper is thus closely related to DeMarzo and Duffie (1999) and Gigler et al. (2009). DeMarzo and Duffie (1999) studies the problem of security design in a signaling game. However, their focus is on the optimal security with no role for accounting disclosure whereas our focus is on the optimal accounting rule when the firm is issuing equity (in other words, they focus on "security design" whereas we focus on "information design"). Gigler et al. (2009) directly models accounting conservatism as changing the relative information content of good earnings versus

[11] We caution against inferring too much from the form of the summary statistic result, as the exact form of the summary statistic will change when the settings of the signaling game change. However, the lower bound result is robust.

bad earnings and show that accounting conservatism could be detrimental to the efficiency of debt contracting. In addition to studying a different setting, our modelling of information structure is more general than theirs and we actually showed that accounting conservatism can be welfare-enhancing.

Secondly, by interpreting the optimal properties of disclosure rules in an accounting standard setting, we provide results that are consistent with arguably one of the most important properties of financial reporting rules: accounting conservatism. The conservatism property as reflected in the "lower of cost or market" rule or good news being more informative than bad news has often been taken as a starting point in studying other accounting issues (e.g., Beyer (2013) on cost of capital and debt contract efficiency, Burkhardt and Strausz (2009) and Caskey and Hughes (2012) on asset substitution). We adopt the conceptual definition of accounting conservatism as requiring more stringent verification for gains than for losses. We show that the optimal accounting rule in our setting generates outcomes that are consistent with this definition of accounting conservatism. We only assume the existence of information asymmetry between the firm and the outside investors and the infeasibility of perfect disclosure.

The rest of the paper is organized as follows. The next section talks about related literature not discussed in the introduction. Section 3 introduces our model and briefly discusses the signaling game. Section 4 discusses how we model disclosure as information structure and how we measure the total amount of information. We also discuss why we do not use variance, arguably one of the most commonly used measures of uncertainty in the finance and accounting literature. The main results are then presented and Section 5 concludes. Appendix A contains more background information about the concepts of entropy and mutual information which will be used in our setting. Appendix B contains all of the proofs.

2 Related literature

First, our paper is related to the disclosure literature. We study how ex ante commitment to ex post disclosure can help to reduce the inefficiency caused by signaling given a fairly general information structure. Bhattacharya (1979), Miller and Rock (1985), and Leland and Pyle (1977) document how firms can costly signal their private information to outside investors either through dividend or through the fraction of shares retained. The authors acknowledge but

do not consider whether credible disclosure can help to 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. However, the information structure in Kanodia and Lee (1998) is cast in a CARA-normal framework. In such a framework, all players have a constant absolute risk aversion (CARA) utility function and 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 variation of the total quantity of information in a normally distributed world. While CARA-normal framework is widely used in the literature to address how much information should be disclosed in strategic settings,[12] 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).[13] In this paper, we allow for more general information structure to focus on the qualitative properties of optimal disclosures in a signaling setting and to generate some accounting implications.

Secondly, our paper is related to the vast literature on accounting conservatism. In addition to the papers mentioned in the introduction, there are quite a few recent studies that justify accounting conservatism defined in different notions under 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 corporate board, and Bertomeu and Magee (2015) rationalize conservatism in a model of political influence. More recently, asymmetric reporting is optimal in Armstrong et al. (2014) 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 beneficial effects of conservatism on debt contracting efficiency, while Bertomeu et al. (2013), Gao (2013a), and Li (2013) incorporate such factors as managerial compensation, earnings management or

[12] See Dye (2001), Verrecchia (2001), Kanodia (2006), Beyer et. al. (2010), and Stocken (2013) for excellent summaries of this literature.

[13] We appreciate an anonymous referee for making this point.

renegotiation cost, and suggest there may be an interior degree of conservatism.

Thirdly, our paper is related to the literature on flexible information structure and, more generally, the qualitative properties of information. In our setting, since in general private information cannot be credibly conveyed to outsiders, the only way to convey such information is through pre-specified accounting rules. The rules are assumed to be flexible in the sense that any kind of rules that induce any Bayes-plausible posterior distribution[14] can be specified ex-ante. Thus qualitative properties of information structure play an important role in addition to quantitative properties (as characterized by variance in a CARA-normal framework). To focus on the qualitative properties, however, we need a measure of quantitative properties to control for the total amount of information contained in an information structure. Blackwell's ordering is not an appropriate metric in our setting because it is not complete. We use the mutual information concept adapted from information theory[15] and look for the optimal qualitative properties of the information structure. This notion has been used in Yang (forthcoming) to study how flexible information acquisition affects the outcome of coordination games. Yang shows that, compared with the additive information structure (which he calls "rigid information structure") usually assumed in global games, introducing flexibility in the information acquisition changes the properties of equilibria qualitatively. We believe this notion of "qualitative properties" is particularly suitable for our setting since different accounting rules can generate different information structures that may have the same quantity of information but different qualitative properties. For example, in the accounting literature, Gigler et al. (2009) define accounting conservatism as changing the informativeness of high versus low signals but keeping the overall informativeness of *all* signals unchanged, while Jiang (2015) uses the conservatism notion in Gigler et al. (2009) to study how this qualitative property of accounting information affects the qualitative properties of other information. Clearly such qualitative properties cannot be addressed in a CARA-normal framework, where the only possible variation is the variance or the quantity of information.

Finally, our paper is also related to the economics literature on cheap talk games with commitment. Kamenica and Gentzkow (2011) study a cheap talk game with commitment. In their setting, a sender commits to send a signal

[14] The notion of "Bayes plausible" will be defined precisely later.

[15] See, e.g., Sims (2003, 2005) for excellent discussions of the concept of entropy and mutual information and how they can be adapted to address various economics problems.

according to a pre-specified rule to a receiver who will take an action that affects the payoff of both players. They characterize the optimal signal from the seller's point of view. There are two main differences between our paper and theirs. First, they focus on very general cheap talk setting while we are focusing on a more specific setting because we are more interested in the specific question of the properties of the optimal accounting rule for a firm that signals its private information to outsiders. Second, they do not impose any constraint on the quantity of the information while we do.[16]

3 The model

3.1 Model setup

The game involves a risk-neutral firm and a continuum of competitive risk-neutral investors.[17] The firm is less patient and hence has an incentive to raise cash by selling some percentage of its shares to outside investors.[18] Specifically, the firm discounts future cash flows at rate $\rho \in (0,1)$ and the outside investors 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 zero mean.[19] We assume that $\theta \in \left[\underline{\theta}, \overline{\theta}\right] \in \mathbb{R}_{++}$, and the investors share a common prior belief $\Pi \in \Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)$. We assume Π is absolutely continuous with respect to Lebesgue measure and fully supported on $\left[\underline{\theta}, \overline{\theta}\right]$. The 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

[16] For examples of other accounting applications of this literarure, see, e.g., Friedman et al. (2015) and Huang (2015).

[17] The assumptions are standard in the literature and reflect the fact that we focus on informational issues and abstract away from risk-sharing issues.

[18] 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.

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

information according to it. Specifically, the accounting rule is modelled as an information structure which specifies the conditional distribution of 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 more detail later.

$t = 1$: Nature draws θ according to the common prior Π and the firm learns θ. A report is published according to the accounting rule and published to the investors, who then form their posterior belief, denoted by G accordingly. 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 without confusion. 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.

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

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

$$\begin{aligned} & E[\rho\left(1-q\right)\left(\theta+\varepsilon\right)+qp|\theta] \\ = \quad & q\left(p-\rho\theta\right)+\rho\theta \ . \end{aligned}$$

Since the second term has no strategic effect, just let

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

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 disclosure 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).[20]

[20] The fact that optimal mandatory disclosure coincides with optimal voluntary disclosure in our setting does not make our results less plausible. In fact, mandatory accounting rules have

Given G, the firm's strategy is $q_G : \sup p(G) \to [0,1]$ 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 in 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*[21] *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

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

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

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

their roots in accounting practice voluntarily adopted by firms long before they are proposed.

[21] 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 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.

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 rule 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 absent information constraint the firm always rank accounting rules according to Blackwell's ordering. We then introduce a measure of information and characterize the optimal accounting rule subject to some information constraint.

4.1 A general approach of modelling accounting rule

We consider a general approach to modelling disclosure. In the literature, 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 $\Lambda(G)=\begin{cases} \lambda_H & \text{if } G=G_H \\ \lambda_L & \text{if } G=G_L \\ 0 & \text{otherwise} \end{cases}$. By definition Λ is Bayes plausible, i.e., equation (5) holds.*

Example 2 *(Partition)* [22] *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 as an information structure with three signal realizations, H, M and L, such that $\Pr(s=L|\theta)=\begin{cases} 1 & \text{if } \theta\in[1,1.3] \\ 0 & \text{otherwise,} \end{cases}$ $\Pr(s=M|\theta)=\begin{cases} 1 & \text{if } \theta\in(1.3,1.7] \\ 0 & \text{otherwise} \end{cases}$, 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 in Gigler et al. (2009), $\widetilde{x}$ denote 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,*

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

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*

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

where

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

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.[23] 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.*

[23] 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.

Corollary 1 provides a potential justification of fair value accounting where managers disclose their best estimates to outside investors in order 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 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 information

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

4.3.1 Measuring uncertainty

Although in the literature, variance is often used to measure uncertainty, it may fail to capture some key ingredients of uncertainty in certain circumstances. We provide an example to highlight this point. Suppose there are $N > 2$ states indexed by $i \in \{1, 2, ..., N\}$ and let's consider two probability distributions: (i) states 1 and N occur with $\frac{1}{2}$ probability and the rest $N-2$ states occur with 0 probability; (ii) all states occur with equal probability. The variance associated with probability distribution (i) is $var(i) = \frac{1}{2}(1 - \frac{N+1}{2})^2 + \frac{1}{2}(N - \frac{N+1}{2})^2 = \frac{(N-1)^2}{4}$ whereas the variance associated with probability distribution (ii) is $var(ii) = \sum_{j=1}^{N} \frac{1}{N}(j - \frac{N+1}{2})^2 = \frac{(N-1)(N+1)}{12}$. 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 that under probability distribution (ii). This is counterintuitive in the sense that under equally probable outcomes, the latter should be more uncertain than the former because there are only two equally probable outcomes under probability distribution (i) whereas under probability distribution (ii) there are $N > 2$ equally probably outcomes.

Following the information theory[24], we use Shannon's Entropy to measure the uncertainty associated with a probability distribution p over Θ. Specifically, the entropy associated with p is given by

$$h(p) = -\sum_{i=1}^{N} p_i \log(p_i). \tag{6}$$

The entropy associated with probability distribution (i) is $h(i) = -\frac{1}{2}\log(\frac{1}{2}) - \frac{1}{2}\log(\frac{1}{2}) = \log 2$ and the entropy associated with probability distribution (ii) is $h(ii) = -\sum_{i=1}^{N} \frac{1}{N}\log(\frac{1}{N}) = \log N$. When all the states are equally probable, 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.

4.3.2 Measuring information

Information can be measured by the reduction of entropy since it reduces the uncertainty of the underlying state. Let $\Lambda \in S(\Pi)$ be an information structure. For any resulting posterior belief $G \in$supp(Λ), the posterior uncertainty is measured by its associated entropy,

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

where g is the density of G[25]. We use the notation $\mathbf{E}_G$ to economize expression (6) and $\mathbf{E}_G$ denotes the expectation under posterior belief G. The expected posterior uncertainty is then given by $\mathbf{E}_\Lambda [h(G)]$. Note that the posterior uncertainty is measured by the prior entropy $h(\Pi)$. Then, the amount of information is given by[26]

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

[24] The seminal work is Shannon (1948). See Cover and Thomas (1991) for more rigorous justification of entropy as a measure of uncertainty.

[25] More rigorously, g is the Radon-Nikodym derivatiuve of G.

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

4.4 Main results

As discussed before, we impose a constraint

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

for some $\kappa \geq 0$ in order to capture the infeasibility of perfect disclosure. The firm's problem is to choose a $\Lambda \in S(\Pi)$ to maximize $U_f(G)$ given by expression (4) 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 by $\frac{\Lambda(G_1)}{\Lambda(G_1)+\Lambda(G_2)} G_1 + \frac{\Lambda(G_1)}{\Lambda(G_1)+\Lambda(G_2)} G_2$, we create a new disclosure 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 concern.

The summary statistic of a posterior G corresponds to firm's equilibrium expected payoff $U_f(G)$ in order 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 disclosure rule in example 2, i.e.,*

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

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* $\pi(\theta) = \begin{cases} \frac{1}{2} & \text{if } \theta \in [1, 1.5) \\ \frac{3}{2} & \text{if } \theta \in [1.5, 2] \\ 0 & \text{otherwise} \end{cases}$. *An accounting rule that is consistent with Proposition 2 is* $\Lambda(G) = \begin{cases} \frac{1}{2} & \text{if } G = G_H \\ \frac{1}{2} & \text{if } G = G_L \\ 0 & \text{otherwise} \end{cases}$ *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* $\Pr(s = H|\theta) = \begin{cases} 0 & \text{if } \theta \in [1, 1.5) \\ \frac{2}{3} & \text{if } \theta \in [1.5, 2] \end{cases}$ *and the probability of disclosing $s = L$ is given by* $\Pr(s = L|\theta) = 1 - \Pr(s = H|\theta) = \begin{cases} 1 & \text{if } \theta \in [1, 1.5) \\ \frac{1}{3} & \text{if } \theta \in [1.5, 2] \end{cases}$. *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 the property of accounting 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 as good news disclosures. This is consistent with the general interpretation for 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 be equal to the recognized gain, making the good news very informative. In contrast, aggressive reporting would imply that the bad news be more informative than good news, whereas neutral reporting would imply that bad news is as informative as good news. 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 at the time of incurring such expenditure. 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 is also related to the "lower of cost or market" principle applicable to assets that are easier to measure than, e.g., internally developed intangible assets. 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 that 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., 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 underlying 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 is different 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 completely eliminated by disclosure, without any particular structure imposed in 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 accounting 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 are relevant 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 explore further 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 rule 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, those papers do not address the qualitative properties of accounting information in a fairly general way, which is the focus of our paper.

In deriving the optimal qualitative properties of accounting information, we use the mutual information concept adopted from information theory. This choice of mutual information 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 single 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 need to make trade-offs 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. More specifically, in our model, accounting disclosure is still modelled as a black box. Although we are able to show that the optimal accounting rule can be interpreted to be consistent with accounting 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 information 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

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

This functional form is not arbitrarily chosen. It is derived axiomatically from information theory and also closely related 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

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

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

$$h(\theta) = -\int_{\theta \in \mathbb{R}} p(\theta) \cdot \log p(\theta) \cdot d\theta \ .$$

The mutual information of this case can be defined accordingly.

For example, if $\theta \sim N(\mu_\theta, \sigma_\theta^2)$, then $h(\theta) = \frac{1}{2} \cdot [\log(2\pi) + 1] + \log \sigma_\theta$.

Suppose s is a signal of θ, and s and θ are jointly normally distributed.[27] Suppose the correlation between s and θ is ρ, then the mutual information between s and θ is $h(s|\theta) = -\frac{1}{2} \cdot \log(1-\rho^2)$.

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 $p(q) = E[\theta + \varepsilon | q(\theta), G] = E[\theta | \left[\frac{\widehat{\theta}(G)}{\theta}\right]^{\frac{1}{1-\rho}}, G] = \theta$ since $\left[\frac{\widehat{\theta}(G)}{\theta}\right]^{\frac{1}{1-\rho}}$ 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 $u_f(\theta, q, p) = q \cdot (p - \rho\theta)$. 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 $p(1) = \widehat{\theta}(G)$, since the lowest type has nothing to gain from retaining any of the shares. Solving would give us $p(q) = \frac{\widehat{\theta}(G)}{q^{1-\rho}}$. 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.

[27] In the literature, s is also assumed to be $\theta + \varepsilon$ where ε is normally distributed and independent of θ.

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

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

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

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

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

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

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

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

and

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

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

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

and

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

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 $\operatorname{supp}(\Lambda_2)$ to $\operatorname{supp}(\Lambda_1)$, expressed as $\Gamma \in \Delta(\Delta(\operatorname{supp}(\Lambda_2)))$, a distribution of probability measures over $\operatorname{supp}(\Lambda_2)$ such that

$$
\Lambda_2=\int \nu \Gamma(d \nu) \tag{9}
$$

and

$$
\operatorname{supp}\left(\Lambda_1\right)=\left\{\int_{\operatorname{supp}(\nu)} G \nu(d G): \nu \in \operatorname{supp}(\Gamma)\right\} \quad , \tag{10}
$$

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

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

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., $\operatorname{supp}(\nu)$ contains at least

two distinct G.

This concludes the proof. ■

Proof of Proposition 2:

Proof. The firm's problem is

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

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

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

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

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 $\in$supp(Λ_{κ}^{*}) 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., $w = \Lambda\left(G_1\right) / \left[\Lambda\left(G_1\right) + \Lambda\left(G_2\right)\right]$. 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 $\in$supp(Λ_{κ}^{*}). 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 $\left[\widehat{\theta}(G), \overline{\theta}\right]$ with respect to the moment $\int_{\left[\widehat{\theta}(G), \overline{\theta}\right]} \theta^{\frac{-\rho}{1-\rho}} dG(\theta)$. 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. Wiley-Interscience, New York.

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, 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.

Jiang, X., 2015. Bias in accounting Information and bias in non-accounting Information: substitutes or complements? 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,* forthcoming.