Properties of Optimal Accounting Rules in a Signalling 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

Preliminary and incomplete. Comments welcome.

Abstract

We characterize the properties of optimal accounting rules in a signaling game where an impatient entrepreneur sells shares to competitive investors. The entrepreneur can signal her private information about the fundamental of the firm by retaining a fraction of the shares. In addition, she can commit to disclosing information according to a set of accounting rules chosen ex ante. Information disclosure reduces signaling cost so that perfect diclosure is optimal. However, perfect disclosure requires disclosing infinite amount of information (measured by reduction of Shannon's entropy), which is usually unrealistic. When disclosure can only reveal finite amount of information, the optimal accounting rule features an infimum and a summary statistic of the fundamental. The infimum can be interpreted as being consistent with various conservative accounting rules while the statistic summarizes the most relevant information determined by the signalling game.

*We thank Ivan Marinovic and Steve Morris for their helpful comments and suggestions.

1 Introduction

In this paper we study the properties of optimal accounting rules that a firm would like to commit to ex-ante in a signalling game. [1]In our model, a firm in need of capital is selling its shares to outside investors to raise cash. Information asymmetry between the firm and the investor leads to firm signaling its private information through the percentage of shares it retains. However, such signalling via retaining shares is costly because the firm is impatient and discounts the future cash more than the investor. We introduce the possibility of a firm, before learning of the private information, to commit to disclosing information that is informative about the firm value following a pre-specified disclosure rule. Such disclosure helps to reduce the information asymmetry between the firm and outside investors, thus decreasing the inefficiency induced by signalling-via-share-retaining. We first document that disclosure rules that are dominated in a Blackwell sense can never be optimal, implying that a study of the properties of optimal disclosure rules requires controlling for the quantity of information allowed by the rules. Using the mutual information concept from information theory to control for the quantities of disclosure, we investigate the qualitative properties of optimal disclosure rules and find that the qualitative properties can be interpreted as being consistent with the financial reporting principles adopted in reality.

Ever since Akerlof (1970), information asymmetry has been documented as one of the major frictions in the economy that could possibly lead to a market shut-down. Numerous subsequent research (e.g., Leland and Pyle (1977, DeMarzo and Duffie (1999)) showed that although information asymmetry may not always result in a market shut-down, it always introduces costs by resulting in firms' socially inefficient decisions. For example, in Leland and Pyle (1977) where firms sell shares to external investors to raise cash, private information can be conveyed through the percentage of shares retained by the firm. However, such signaling is inefficient, as the selling price is lower than that without information asymmetry because of the "lemons" problem. Similarly in DeMarzo and Duffie (1999) where firms in need of cash sell securities to external investors, the inefficiency stems from firms' higher discount rate of the future than investors but inability to sell all shares to investors. It is thus an important question whether or not there exist other mechanisms to alleviate or even eliminate such inefficiency.

Disclosure, either voluntarily adopted by the firm or mandatorily imposed by the regulator, is one of the mechanisms that have been proposed. The classical work of Grossman (1981) and Milgrom (1981) have illustrated that voluntary disclosure can completely eliminate the inefficiency caused by information asymmetry if (1) the uninformed party knows that the informed party has the relevant information, (2) disclosure is costless and (3) disclosure can be ex-post verified. Subsequent research resorts to real-life settings to justify relaxing the first and second assumption, resulting in partial disclosure. [2] However, most of the papers in this literature do not relax the third assumption, i.e., disclosure is credible as it can be ex-post verified, which also does not seem to be realistic. In fact, in the setting of Leland and Pyle (1977), voluntary disclosure is not credible because it can not be ex-post verified. That brings the question of whether ex-ante commitment to credible ex-post disclosure can help alleviate the inefficiency caused by information asymmetry. In reality, ex-ante commitment to ex-post disclosure as reflected in the form of corporate financial reporting is an important element of the capital markets around the world. In the United States as well as in a lot of other countries, all publicly traded firms are required to periodically disclose financial information according to a pre-specified accounting standard set by the regulators and ex-post verified by the auditors. [3] The objective of such mandatory disclosure is "to provide financial information about the reporting entity that is useful to existing and potential investors ... in making decisions about providing resources to the entity." (FASB (2010)). Because in general such ex-post disclosure cannot perfectly reveal firm's private information (i.e., the third assumption is not satisfied), the question of what disclosure rule or what qualitative properties of verifiable information should be included in financial reports is the most efficient in reducing inefficiency is a fundamentally important question in accounting, finance and economics that surprisingly does not receive much attention in the literature. [4]

We use a setting similar to that of DeMarzo and Duffie (1999) to character-

[2] Subsequent work that relaxes some of the assumptions include Dye (1985), Jovanovic (1982), Jung and Kwon (1988) and Verrecchia (1983). Most of their results are partial voluntary disclosure, i.e., bad news are suppressed and good news are disclosed. However, in their settings voluntary disclosure has at least some credibility whereas in our settings it has none.

[3] In the U.S., the accounting rule is U.S. Generally Accepted Accounting Principles (GAAP). A lot of other countries, including countries in the European Union, now adopt International Financial Reporting Standard (IFRS). Corporate financial reports are accompanied by an auditor's letter, usually stating that the firms' accounting policies are consistent with whatever accounting rules that the firms need to bind to.

[4] Please see below for a more extensive review of literature. An exception is Bertomeu et. al. (2011) that studies security design and voluntary disclosure jointly but disclosure is assumed to be truthful in the sense of Dye (1985).

ize the properties of the optimal disclosure rule that a firm commits to ex-ante, i.e., before observing private information and selling shares. More specifically, we consider an initial public offering (IPO) setting where a privately informed and impatient firm sells shares to outside investors for cash. [5] The firm made a disclosure according to the pre-specified disclosure rule and then sells a percentage of its ownership to the capital market. Investors price the firm based on the firm's disclosure, the pre-specified disclosure rule and the percentage the firm offered to sell. As in DeMarzo and Duffie (1999), retaining shares is costly. However, the cost can be reduced by committing ex-ante to disclosing information according to specified disclosure rules. If such disclosure can perfectly reveal firm's private information, the firm will make such disclosure and first-best efficiency is achieved. [6] However, disclosure cannot be perfect because in reality, financial reports cannot perfectly reveal a firm's private information. Proprietary reasons notwithstanding, firms are usually operated under highly complex and uncertain environment. Disclosures are based on verifiable transactions as well as estimates and predictions that is correlated with but not identical to the private information the firm may possess. As will be discussed in more detail later, the purpose of mandatory disclosure is to set up a rule for the firm to disclose verifiable information that can (partially) convey private information. We view this as a reasonable assumption since in a lot of countries including the U.S., firms' financial reports are audited and thus the information contained in their disclosures are verifiable. However, because of other factors of which the uncertainty is not resolved when firms are disclosing, the verifiable information are at best a noisy signal of the firm's private information.

Our main result is that so long as disclosure cannot perfectly reveal private information, the optimal disclosure rule always has the following features: 1) disclosure of a lower bound of firm value and 2) a summary statistic of the distribution of firm value conditional on that the firm value is higher than this lower bound. Our results can be interpreted as consistent with the fact that financial reports, prepared by accounting rules that are in general conservative, provide a summary of firm value.

The first element of the optimal rule is consistent with many accounting rules that exhibit a "lower-of-cost-or-market" feature, considered to be a mani-

[5] As will be discussed below, the IPO setting is not essential. What is essential is that the manager has a channel other than disclosure to signal her private information

[6] In fact, if there is no exogeneous cost of disclosure, Grossmand (1981) and Milgrom (1981) implies that firms will disclosure this information voluntarily and no mandatory requirement is needed.

festation of accounting conservatism (see, e.g., Beyer (2013)). More specifically, various accounting rules that exhibit such feature have the effect that tends to make balance sheet a lower bound of the firm value. For example, most accounting rules require firms to treat research expenditures as expenses and none of the expenditures can be considered as the asset of the firm. However, it is not rare to see research activities turning into profitable projects which will bring future benefits to the firm. Thus, treating the value of research activities as having a zero value provides a lower bound of the value of research activities. Another example is the impairment accounting rules or so-called "lower of cost or the market" rule. Under such accounting rule, quite a few categories of assets (e.g., inventory, tangible assets, goodwill) have to be tested annually to see if the fair value [7] is lower than recorded acquisition value. If the fair value is lower, then the asset has to be recorded on the balance sheet using the lower fair value, i.e., "write-down" the value of the asset. However, if the fair value is higher than the recorded acquisition value, the firm cannot record the higher value on its balance sheet and has to record the asset at its book value, i.e., 'write-up" is not allowed. [8] The "lower of cost or market" rule is thus another example of balance sheet values being a lower bound of firm value. Our results are thus consistent with the conservatism principle embedded in many disclosure standards and we derive our results in a fairly general setting with information asymmetry being the only friction.

Our result is also consistent with financial reports providing a summary of firm value conditional on the firm's value verifiably exceeding the lower bound. U.S. GAAP requires that auditors issue an opinion regarding whether the firm can continue as a going concern and most firms' financial reports indicate that the presented numbers are based on the assumption that those firms will continue as a going concern. The lower bound in our setting can be interpreted as providing a lower bound of the firm value if the firm is liquidated today (i.e., the liquidation value cannot be lower than this lower bound). The summary statistic is then a summary of firm value based on the firm continuing and its value exceeding this lower bound. [9] In plain words, the summary statistic provide more information about the firm value, e.g., whether the firm value is closer

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

[8]Under IFRS, in certain circumstances, write-up after a previous write-down is allowed. However, the asset's value cannot exceed its original acquisition value.

[9]We caution against making too much inference from this moment condition as the exact form of moment condition will change when the signalling settings change. However, as will be discussed later, the lower bound result is robust.

to the lower bound or the (commonly known) upper bound. In our setting, a posterior distribution is uniquely determined given the summary statistic and the lower bound as those are sufficient statistics of the distribution.

The intuition of our result is as follows. As is typical in the signalling models in the presence of information asymmetry, outsiders are most concerned about firms with low fundamentals misleading them by disclosing overly optimistic information. The inefficiency from signalling thus comes from firms with more favorable information involving in inefficient behavior (in our setting, retaining more shares) to differentiate them from those with less favorable information. Thus the most inefficiency-reducing way of disclosure is to reassure outside investors that the firm value cannot be lower than a verifiable lower bound, resulting in the first part of our result. Note that disclosing upper bound will not help since such disclosure cannot give outsiders the reassurance and they are still concerned about the firm being a lower type than that implied by the disclosure. The second part is due to the fact that conditional on firm value exceeding the lower bound, a specific summary statistic (more specifically, the $-\frac{\rho}{1-\rho}$th moment) is a sufficient statistic for ex-ante firm value. Since disclosure cannot perfectly reveal private information, the most efficient disclosure is to provide the sufficient statistic.

Our paper is related to and makes several contributions to the economics, finance and accounting literature. First, our paper is, to the best of our knowledge, the first study of how ex ante commitment to an ex-post disclosure rule can help reduce the inefficiency introduced by signalling in a fairly general framework. Bhattacharya (1979), Miller and Rock (1985) and Leland and Pyle (1977) documents how firms can costly signal its private information to outsides either through dividend or through the fraction of shares retained. The authors acknowledge but do not consider whether credible disclosure can help alleviate the signaling cost. Kanodia and Lee (1998) explicitly considers the role of mandatory disclosure in an investment setting and showed that endogenously imperfect mandatory disclosure is essential in supporting a signalling-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 the 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 is thus equivalent to variation of the precision. In a normally distributed world, the variation of the precision is equivalent to variations of the quantity of information. While CARA-normal framework is widely used in the literature to address how much information should be disclosed in strategic settings, [10] it cannot be used to address the qualitative properties of optimal disclosure, e.g., which part of the disclosure should be more informative while keeping the total quantity of information unchanged. In this paper, we allow for a more general information structure to focus on the qualitative properties of optimal disclosures in a signalling setting. [11]

Secondly, our paper is also 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, which is assumed to be flexible in the sense that any kind of rules that induce any Bayes plausible posterior distribution [12] can be specified ex-ante. [13] Thus qualitative properties of information structure play an important role in addition to quantitative properties (as characterized by the variance in a CARA-normal framework). To focus on the qualitative properties, we need a measure of quantitative properties to control for the amount of information contained in any 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 (see, e.g., Sims (2003,2005) for excellent discussions of the concept of entropy and mutual information and how it can be adapted to address various economics problems) and solve for the optimal qualitative properties of the information structure when we control for the quantity of information contained in the information structure. This notion of informativeness has been used in Yang (2012) to study how flexible information structure affect the outcome of coordination games. Yang (2012) showed that compared with the normally distributed information structure (which he called the "rigid information structure") usually assumed in global games, introducing flexibility in the information structure can qualitatively change the properties

[10] See Kanodia (2006), Beyer et. al. (2010) and Stocken (2013) for an excellent summary of literature on accounting disclosures in strategic settings.

[11] It is worth mentioning that even in strategic settings where there is an optimal intermediate quantity of information, the optimal qualitative properties of information is still unresolved.

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

[13] 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 investors and firm signal their private information through other channels, e.g., investment. So long as the signalling equilibrium is fully revealing, our results will not change qualitatively.

of equilibrium. We believe this notion is particularly suitable for our setting since different disclosure rules can generate different information systems 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 (2013) uses the conservatism notion in Gigler et. al. (2009) to study how this qualitative property of accounting bias affects the qualitative properties of other information an individual investor is acquiring. Clearly such qualitative properties cannot be addressed in a CARA-normal framework where the only variation is the quantity of information.

Thirdly, by applying the notion of informativeness to an accounting standard setting, we provide results that are consistent with two of the most important properties of financial reporting: conservatism and summarizing. The conservatism property as reflected in the lower of cost or market principle have 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 and debt contract efficiency). There are also quite a few studies that justifies conservatism as defined in various notion under various settings. Chen et. al. (2007) and Gao (2013) documents the desirability of conservative accounting in the presence of the possibility of managers manipulating accounting reports in a debt-contracting setting, while Caskey and Laux (2013) justifies conservatism as strengthening the governance role of corporate board. On the other hand, Gigler et. al. (2009) casts doubt on the beneficial effects of conservatism on debt contract efficiency, while Bertomeu et. al. (2013) and Li (2013) incorporates such factors as managerial compensation or renegotiation cost and suggest there may be an interior degree of conservatism. We focus on the lower of cost or market perspective and showed that this principle is part of the properties of optima disclosure where the only friction is information asymmetry between insiders and outside investors and this private information cannot be credibly verified.

Finally, our paper is also related to the economics literature on cheap talk game with commitment. Kamenica and Gentzkow (2011) consider a cheap talk game with commitment. In their setting, a sender commits to send a signal 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, in their cheap talk game the receiver merely takes an action after receiving the signal while in ours the sender and the receiver are involved in a more complex signaling game. Second, while Kamenica and Gentzkow (2011) focus on very general cheap talk setting, we are focusing on a more specific setting because we are more interested in the specific question of the properties of the optimal disclosure rule for a firm that signals its private information to outsiders.

The rest of the paper is organized as follows. Section 2 introduces our model while section 3 establishes our main results and Section 4 concludes. All of the proofs are in the appendix.

2 The Model

There is a risk-neutral firm and many competitive risk-neutral investors. [14] As is standard in the finance literature, we assume that the firm is impatient and has an incentive to raise cash by selling some percentage of its shares to outside investors. [15] More specifically, the firm discounts future cash flows at rate $\rho \in (0,1)$ and the outside investors do not discount future cash flows. This can be viewed as a reduced form of modelling the situation when the firm needs to sell shares for cash, e.g., the firm may strictly prefer rasing capital for new investment opportunities or when the firm is subject to some regulatory capital constraints and has to raise cash.

Let $\theta \in \left[\underline{\theta}, \overline{\theta}\right] \in \mathbb{R}_{++}$ denote the firm's value and let $\Pi \in \Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)$ denote the prior distribution of θ. The firm has informational advantage relative to outside investors in that the firm knows θ before issuing shares to the outside investors. The firm can commit to disclose information about θ before selling shares. The disclosure has to follow a pre-specified disclosure rule which will be discussed in detail later. The commitment to the disclosure rule has to be specified ex-ante, i.e., before the firm learns about θ. This assumption is meant to capture the fact that financial reporting rules are established before the firm starts operating and learns any private information. After committing

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

[15] This assumption, while uncommon in the accounting literature, is as innocuous as the assumption in quite a few papers that old investors have to sell 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.

to provide a report following the disclosure rule, the firm privately learns θ and chooses the fraction of shares to issue. The report will also provide information about θ following the pre-specified disclosure rule. The investors price the firm by making inferences about the value according to the fraction of share issued, the disclosed report and the disclosure rule that the firm commits to ex ante.

Similar in spirit to Kanodia and Lee (1998), we focus on the *informational* features of the disclosure and abstract away from the underlying measurement issue. However, in contrast to Kanodia and Lee (1998), we focus on the qualitative properties of the disclosure while they focus on the quantitative properties. In other words, we view disclosure rule as an information structure and investigate the qualitative properties of the optimal information structure. To be more specific, first note that any report generates a posterior distribution of θ conditional on the report. Since the value of the report itself is a random variable that is distributed according to the disclosure rule, each disclosure rule generates a distribution of posterior distributions which we call "information structure". While it is true that different disclosure rules can cause this posterior distribution to be more or less informative about θ in the Blackwell sense, they can also cause variations in accounting reports that may not result in a change in the reports' overall informativeness about θ. For example, accounting conservatism, as modelled in, e.g., Gigler and Hemmer (2001), Chen et. al. (2007) and Gigler et. al. (2009), makes favorable reports more informative while unfavorable reports less informative, may not result in a change in the overall informativeness of accounting reports with respect to the underlying fundamental. In fact, Gigler et. al. (2009) explicitly stated, on page 783 of their paper, that [16]"(w)hile the information content of *some* signal realizations can be ordered across liberal and conservative accounting regimes, it is not the case that the *overall* information content of a conservative accounting regime is greater or less than the overall information content of a liberal accounting regime". We thus model the disclosure rule as inducing an information structure in the most general sense that incorporates both the qualitative and quantitative properties of an information structure. Mathematically, an information structure is expressed as a conditional distribution of signals (which means accounting reports in our setting) given the fundamental state θ. Let G denote a generic signal realization. Then each G is associated with a posterior of θ. With a slight abuse of notation but without confusion, we also use G to denote the associated

[16]The words that are in in Italics are emphasized by Gigler et. al. (2009).

posterior, i.e., $G \in \Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)$. Hence, an information structure amounts to a probability distribution of posteriors, denoted by $\Lambda \in \Delta\left(\Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)\right)$. Denote $\operatorname{supp}(\Lambda) \subset \Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)$ as the set of possible posteriors under information structure Λ. Given the prior $\Pi \in \Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)$, although any information structure can be expressed as a distribution of posteriors, not any arbitrary distribution of posteriors can be derived from an information structure. A distribution of posteriors Λ is an information structure if and only if it is Bayes plausible with respect to prior Π, i.e.,

$$\Pi = \int G \cdot \Lambda\left(dG\right) \ . \tag{1}$$

Let $S\left(\Pi\right)$ denote the set of all Bayes plausible information structures, i.e.,

$$S\left(\Pi\right) = \left\{\Lambda \in \Delta\left(\Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)\right) : \Pi = \int G \cdot \Lambda\left(dG\right)\right\}.$$

To facilitate our derivation, we will stick to this posterior approach in the paper, i.e., each disclosure rule is associated with a distribution of the posterior distribution of θ, i.e., G. In plain words, this means that each disclosure rule would generate a distribution of accounting reports conditional on θ. With the knowledge of the report and the rule that generates the report, the outsider will have a posterior distribution of θ conditional on the report, which, of course, also depends on the rule. As an example, note that the characterization used in Gigler et. al. (2009) to model accounting conservatism can be viewed as a special case of our accounting rule specification. Using the notation in Gigler et. al. (2009), the posterior distribution of future cash flow x given an accounting report y is $\varphi(x|y, \delta)$. The posterior is a function of δ and is used in Gigler et. al. (2009) to capture the variation of the degree of conservatism in the accounting rule. The characterization can be equivalently viewed as each δ inducing a posterior distribution of x conditional on y, which is consistent with our specification. However, our specification is more general as each of our posterior is not a function of a one-dimensional deterministic number δ but a function of all distribution functions that is Bayes plausible with respect to the prior. In technical terms, Gigler et. al. (2009) focuses on a group of distribution functions that is a one-dimensional contour of an infinite dimensional functional space which we use in our setting. Another special case that is contained in our general formulation of the information structure is the partitioning equilibrium commonly seen in cheap talk games (e.g., Crawford and Sobel (1982)). Note that in a partitioning equilibrium, any disclosed signal is associated with a unique posterior distribution of θ determined by the partitioning equilibrium. In our case, any disclosure can be associated with not a unique but a distribution of posterior distributions of θ so long as the Bayes compatibility condition as stated by equation (1) is satisfied. The following examples provide an illustration of those different cases.

Example 1 *Suppose that the prior distribution of* θ *is uniform on* $[1, 2]$. *An example of an information structure* Λ *is as follows: Let* f_1 *be a posterior distribution of* θ *on* $[1, 2]$, f_2 *be a posterior distribution on* $[1.1, 1.5]$, f_3 *be a posterior distribution on* $[1.3, 1.7]$ *and* f_4 *be a posterior distribution on* $[1.6, 2]$. $\Lambda = \{(f_1, \pi_1(\theta)), (f_2, \pi_2(\theta)), (f_3, \pi_3(\theta)), (f_4, \pi_4(\theta))\}$, *where* $\pi_i(\theta)$ *is the probability of choosing posterior* f_i *when the true value is* θ *that satisfies* $\sum_i \pi_i(\theta) = 1$ $\forall\theta$ *and the Bayesian plausibility condition. Clearly the Bayesian plausibility condition is the most stringent constraint. For example, it can be verified that Bayesian plausibility implies that* $\pi_2(\theta) = 0\ \forall\theta > 1.5$ *and* $\pi_1(\theta) = 1\ \forall\theta < 1.1$. *Intuitively, it means that when* $\theta > 1.5$ f_2 *cannot be the posterior distribution while when* $\theta < 1.1$ f_2, f_3 *and* f_4 *cannot be its posterior distribution. As will be shown below, however, such information structure can never be optimal.*

Example 2 *The information structure in Gigler et. al. (2009) is also a special case of* Λ. *Again suppose that the prior distribution of* θ *is uniform on* $[1, 2]$. *Let* $f_{\delta,y}$ *be a posterior distribution of* θ *on* $[1, 2]$ *with* $\delta \in R$, $y \in [\underline{y}, \overline{y}]$ *with a density function* $l(\delta, y)$ *for some* $\underline{y}$ *and* $\overline{y}$. $\Lambda = \{(f_{\delta,y}, g(\delta, \delta', y))\}_{\delta,\delta' \in R, y \in [\underline{y}, \overline{y}]}$ *where* $g(\delta, \delta', y) = I_{\{\delta'=\delta\}} l(\delta, y)$ *and* $I_{\{\delta'=\delta\}}$ *is the identity function.*

Example 3 *The partitioning distribution can be seen as a special case of* Λ. *Still suppose that the prior distribution of* θ *is uniform on* $[1, 2]$. *Consider the following partitioning posterior distribution: Let* g_1 *be a posterior distribution of* θ *on* $[1, 1.5]$, g_2 *be a posterior distribution on* $(1.5, 2]$. $\Lambda = \{(g_1, \pi_1(\theta)), (g_2, 1 - \pi_1(\theta))\}$. *Bayesian plausibility would imply that* $\pi_1(\theta) = 1\ \forall\theta \in [1, 1.5]$ *and* $\pi_1(\theta) = 0\ \forall\theta \in (1.5, 2]$. *Thus, any partitioning distribution results in a unique posterior distribution* $\forall\theta$.

The timeline of the model is as follows:

$t = 0$: Nature draws θ from the prior distribution Π. Neither the firm nor the potential investors knows θ at this date. They share the common prior Π. The firm chooses an information structure Λ, i.e., a Bayes plausible distribution of posteriors with respect to Π.

$t = 1$: The firm privately learns θ. A report is published according to Λ, i.e., a posterior $G \in$supp(Λ) is drawn according to distribution Λ. The firm then chooses a fraction $q \in [0,1]$ to sell to outside investors. The competitive investors price the firm at

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

i.e., the price under perfect competition.

$t = 2$: The fundamental θ is realized and consumed by the firm and investors accordingly.

Given θ, the firm's utility is

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

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 utility. Obviously, the investors' expected utility is zero due to perfect competition. Therefore, the accounting rule that maximizes the firm's utility is also maximizing the social welfare of both the firm and the investor. Correspondingly, 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). [17]

The equilibrium of the game is defined as follows.

Definition 1 *A Bayes-Nash equilibrium of the game is a triplet* $\{\Lambda, q, p\}_\lambda$ *such that:*

(i) $q(\theta) \in \arg\max_q u_f\left(\theta, q, p(q)\right)$ *almost surely.*

(ii) $p(q) = E[\theta | q(\theta), G]$ *almost surely, where* $G \in supp(\Lambda)$ *and that* $\Pi = \int G \cdot \Lambda\left(dG\right)$.

(iii) $U_f\left(G\right) = \mathbf{E}\left[u_f\left(\theta, q\left(\theta\right), p\left(q\left(\theta\right)\right)\right) | G\right]$ $\forall G \in supp(\Lambda)$.

(iv) Λ *solves* $\sup_{\Lambda' \in S(\Pi)} \mathbf{E}_{\Lambda'}\left[U_f\left(G\right)\right]$ *subject to the constraint that the mutual information of* Λ'*, to be specified below, is no larger than* λ.

We also follow the usual convention to define the equilibrium in a signalling

[17] The fact that voluntary and mandatory disclosure coincide in our setting does not make our results invalid. In fact, the mandatory accounting rules have their roots from accounting practice established long before mandatory disclosure rules are proposed.

game to be fully revealing.

Definition 2 *A signalling game equilibrium* $\{q, p\}$ *is fully revealing if* $p(q(\theta)) = \theta$ *almost surely.*

3 Solution

To solve the problem we work backwards. First for any given posterior G we solve for the equilibrium of the signalling game at $t = 1$ and show that the equilibrium is unique and fully revealing. We then solve for the optimal properties of Λ given that any posterior $G \in$supp(Λ) induces the signalling equilibrium at $t = 1$.

3.1 Signaling Game at $t = 1$

Given posterior G, the firm and outside investors are playing a signaling game at $t = 1$. The firm's strategy is q :supp$(G) \rightarrow [0, 1]$. The investors' strategy is $p : q \rightarrow$supp(G). Our next result shows that, as is standard in the signalling literature, there is a unique fully revealing equilibrium because of the satisfaction of the single-crossing property: $u_{q\theta} < 0$.

Lemma 1 *Given posterior* G*, the signaling game has a unique equilibrium, which is fully separating and characterized by*

$$q(\theta) = \left[\frac{\widehat{\theta}(G)}{\theta}\right]^{\frac{1}{1-\rho}} \tag{2}$$

and

$$p(q) = \frac{\widehat{\theta}(G)}{q^{1-\rho}} \ , \tag{3}$$

where

$$\widehat{\theta}(G) = \inf(supp(G)) \ . \tag{4}$$

This result is similar to that documented in DeMarzo and Duffie (1999), although their focus is on optimal security design while keeping disclosure fixed whereas we focus on the optimal properties of disclosure but keep the security issued to be equity. [18]

[18] In DeMarzo and Duffie (1999), equity can be the optimal security under certain conditions.

The results are also quite intuitive. Since $u_f(\theta, q, p)$ satisfies the single-crossing property, given any posterior G, a firm with higher θ will signal to the outside investors by retaining a higher portion of shares as a lower θ firm will find it more expensive to mimic firms with higher θ. Thus the equilibrium is fully revealing and $p(q(\theta)) = \theta$. However, despite the equilibrium being fully revealing, the equilibrium is socially inefficient. The inefficiency is induced by firms being more impatient than outside investors. In the first-best case when θ is publicly known, we have $q(\theta) = 1\ \forall\theta$ and $p(\theta) = \theta$. The intuition is that the impatient firm will sell all shares to the investor to get as much cash for today as possible. From equation (2), $q(\theta) \leq 1$ thus there will be inefficiencies except for the lowest type because firms with signals higher than $\widehat{\theta}(G)$ will retain some of their shares. For fixed ρ, the larger the difference between $q(\theta)$ and 1 the more inefficiency induced by signalling. What is worth noting is that it may seem that disclosure plays no other role than providing the lower bound since θ is perfectly revealed through q. However, this is because we choose a setting when firm uses the shares retained as a signalling device. We can use a setting that is more commonly used in accounting (e.g., Kanodia and Lee (1998), Gao (2010)) where the firm sells all shares to new investors and therefore cannot use the shares retained as a signalling device. In this case, disclosure will play an additional role of providing information for valuation of the firm. However, so long as the signalling equilibrium is fully revealing (which is the case in Kanodia and Lee (1998)), our main results will not change qualitatively.

From equation (3), given fixed θ, the lower this lower bound, the higher the percentage of shares retained by the firm and the more inefficiency induced by signalling. [19] The intuition comes from the driving force of the inefficiency: high types engage in inefficient behavior to prevent low types from mimicking since mimicking is more expensive for low types. In our setting, the inefficiency comes from retaining shares since the firm discounts future cash relative to current cash while outsiders do not. To see this, note that when $\rho \to 1$, $q(\theta) \to 0\ \forall\theta > \widehat{\theta}(G)$, i.e., all types of firms retain all their shares and there is no inefficiency loss because there is no discounting. Also note that for the lowest possible type that is consistent with G, i.e., $\theta = \widehat{\theta}(G)$, we will have $q\left(\widehat{\theta}(G)\right) = 1\ \forall\rho$, i.e, the lowest type always efficiently sells all the shares. Thus, the lower the minimum value that θ would attain according to G, the more concerned of the investors about the potential downside loss. Therefore, higher types has to retain more

[19]Note that q is the percentage of shares sold to outside investors.

shares to prevent lower types from mimicking. This driving force will also play a central role in our main result of properties of optimal disclosure to be discussed later.

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

$$\begin{aligned} U_f(G) &= \mathbf{E}\left[u_f(\theta, q(\theta), p(q(\theta)))\,|G\right] \\ &= (1-\rho)\cdot\left[\widehat{\theta}(G)\right]^{\frac{1}{1-\rho}}\int \theta^{\frac{-\rho}{1-\rho}}\,dG(\theta)\ . \end{aligned} \tag{5}$$

We will later explore the optimal properties of accounting rule that generates the distribution of G endogenously.

3.2 Benchmark Result

Before proceeding to our main result, we first establish a benchmark result that compares information structure that can be Blackwell ranked. Note that this is different from Blackwell's Theorem since Blackwell's Theorem concerns an individual's decision-making problem under uncertainty whereas in our setting we have two players playing a signalling game and it is ex ante not obvious that more information in the Blackwell sense is desirable. Nevertheless, the result shows that information structures that are dominated in the Blackwell sense can never be optimal. Before proceeding we first characterize a property of the firm's expected utility that will become important later on.

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

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

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

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

The convexity of $U_f(G)$, which stems from its functional form, guarantees that for any information structure, the optimal posterior distribution G that maximizes the firm's expected utility have a unique lower bound as otherwise we can always carry out a convex combination, which is still Bayes compatible and do better.

We next establish our benchmark when there is no constraint on the amount of information an information structure can convey.

Proposition 1 *Let Λ_1 and Λ_2 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

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

This result is intuitive. In our setting the only friction is the information asymmetry between the firm and outside investors and the inefficiency comes from the firm being unable to sell all the shares to outsiders. Disclosing more information in the Blackwell sense can cause the information structure to be finer, resulting in more accurate disclosure in the sense that the lower bound of each posterior distribution in the information structure will be (weakly) closer to true values of θ. The signalling cost will then be reduced. Thus, an information structure that Blackwell dominates the other cannot do worse. An implication from this proposition is that the inefficiency will be completely eliminated if the firm's private information can be perfectly disclosed to outside investors and full disclosure will be the equilibrium solution. In this case for each $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$ Λ assigns probability 1 to a posterior distribution on $[\theta, \overline{\theta}]$ and 0 to any other posterior, resulting in fully revealing θ from the posterior distribution. Note that this assumption is equivalent to assuming that the firm's private information is verifiable to the public. Then insights from Grossman (1981) and Milgrom (1981) would result in firms fully disclosing all their private information, which is precisely what the following corollary states. The proof is omitted as it directly follows from proposition 1.

Corollary 1 *The firm's optimal information structure Λ^* is characterized by $supp(\Lambda^*) = \left\{\delta_\theta : \theta \in \left[\underline{\theta}, \overline{\theta}\right]\right\}$ and $\Lambda^*\left(\theta\right) = \Pi\left(\theta\right)$, where δ_θ is the Dirac distribution at point θ.*

Corollary 1 is consistent with the fair value method allowed in financial accounting rule for marketable securities. To the extent that the market value of the marketable securities perfectly reveals the fundamental value of those assets, then the optimal information structure should result in disclosing the fair value. However, to the extent that market value or fair value is at best an imperfect estimate of the fundamental value of those underlying assets, [20]

[20] This is the premise underlying many theoretical studies on fair value accounting, e.g., Plantin et. al. (2008).

corollary 1 will not apply and we turn to more general cases.

3.3 General Case

The benchmark result we established is intuitive but not so applicable to real settings. It states that more information (in the sense of Blackwell's ordering) is better, but many information structures are not even comparable with respect to Blackwell's ordering. To achieve a deeper understanding, we search for the optimal information structure while fixing the amount of information. We adopt the concepts of entropy and mutual information from information theory which has been used in economics (Sims (2003,2005)).

Specifically, the uncertainty of any distribution $G \in \Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)$ is measured by its entropy

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

where g is the density of G. Note that entropy is largest for uniform distribution while lowest for degenerate distribution. The information content of an information structure $\Lambda \in S(\Pi)$ is the average amount by which the entropy of the posterior distribution is less than the entropy of prior Π, which is defined as the mutual information associated with Λ, given by[21]

$$I(\Lambda) = h(\Pi) - \mathbf{E}_\Lambda [h(G)] \ . \tag{6}$$

In general, we focus on information structures that conveys information no more than $\kappa > 0$, i.e., $I(\Lambda) \leq \kappa$. Because of proposition 1 we will have $I(\Lambda) = \kappa$ as otherwise we can always choose a Blackwell dominant information structure that contains more information in the mutual information sense and do better according to the benchmark result. [22] Consequently, we focus on the qualitative properties of information structures while keeping the overall amount of information constant.

Since $\kappa > 0$ we are intuitively comparing accounting rules that cannot perfectly reveal the firm's private information. We believe this is a reasonable assumption for modelling firm's disclosure. For example, the firm may have favorable information regarding the future demand of its products. However, such information cannot be credibly verified. However, the firm can choose to build up its inventories. This build-up of inventory and, eventually, the sale of

[21]We refer readers to Cover and Thomas (1991) for further details.

[22]It can be shown that if information structure A dominates information structure B in the Blackwell sense, A have a higher amount of mutual information then B.

the goods, are verifiable and will be reflected in the disclosed financial reports. However, since the future selling price depends on a lot of uncertain factors (e.g., macroeconomic conditions) many of which may be beyond the firm's control, the eventually reported numbers cannot be used to verify the firms' private information. [23] However, the distribution of the reported numbers will be affected by firm's inventory build-up decision which is in turn affected by the firm's private information. Thus, disclosed financial reports are correlated with firm's private information but cannot perfectly reveal it.

Next proposition illustrates our main result when the firm's disclosure cannot perfectly reveal private information.

Proposition 2 *For all $\kappa > 0$, the firm will commit to a unique optimal information structure, denoted by Λ_κ, 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 over $\left[\widehat{\theta}(G), \overline{\theta}\right]$ with respect to the moment $\int_{\left[\widehat{\theta}(G), \overline{\theta}\right]} \theta^{\frac{-\rho}{1-\rho}} dG(\theta)$. In other words, ex post firm will disclose $\widehat{\theta}(G)$ and a number that is equal to $\int_{\left[\widehat{\theta}(G), \overline{\theta}\right]} \theta^{\frac{-\rho}{1-\rho}} dG(\theta)$.*

The first part of the result states that the optimal disclosure rule will induce posterior distributions with different lower bound of the support. The reason is that, as will be stated below, the lower bound plays a crucial role in alleviating inefficiency cost caused by signalling. Thus an optimal information structure should economize on this by inducing posterior distributions with different lower bounds. Posteriors with same lower bound is not the most efficient way of providing crucial information in reducing signalling cost given the constraint on the total amount of information that can be provided. This follows from Lemma 2. The convexity of $U_f(G)$ implies that for each G, the lower bound has to be unique. Note that this result implies that disclosure rules that induce different posterior distributions with fixed support is never optimal in our setting.

The second part of Proposition 2 says that the uniquely optimal disclosure rule is to report 1) the lower bound of the support of the posterior distribution of θ; and 2) a moment which turns out also to be a sufficient statistic of the fundamental conditional on that the firm value being not lower than this lower bound. The investor will then perceive the posterior to be the maximum entropy distribution with the constraint of the lower bound and the moment, whereas both the lower bound and the moment will in general depend on κ.

[23] Put it in another way, bad earnings numbers cannot be used to verify that a firm possesses bad private information as the firm can be claiming that the numbers were caused by bad luck.

The intuition of the main result is as follows. From the analysis of the signalling game we know that the inefficiency of signalling comes from high types involving in suboptimal behavior that is costly for them but more costly for low types to mimic. Thus, given that mandatory disclosure cannot completely reveal the private information, the most inefficiency-reducing way is to disclose the lower bound so that outside investors will be assured that the firm value cannot be any lower. Disclosing upper bound, on the other hand, will not help since outside investors are concerned about low types over-reporting and upper bound disclosure will not help them alleviate such concern, which is our first result. Our result depends crucially on the fact that in this signalling game, the inefficiency comes from high type while there is no inefficiency loss for the low type. So long as this fact holds and the signalling game is fully revealing, the lower bound result will be one of the properties of the optimal disclosure rule, regardless of the actual signalling devices the firm may use. Thus, even in a setting similar to Kanodia and Lee (1998) when a firm sells all the shares to outside investors but use investment as a signalling device, our results will not change qualitatively. However, in a signalling setting when the signalling equilibrium is fully revealing and there is no inefficiency loss for the highest type, then our results will reverse and the optimal disclosure rule will disclose an upper bound. We believe such setting is rare in accounting since the major issue related to disclosure in the presence of information asymmetry is of parties with bad private information trying to pretend that they have better information.

For the second part, note that under any posterior G, the firm's expected utility is $U_f(G) = (1-\rho)\cdot\left[\widehat{\theta}(G)\right]^{\frac{1}{1-\rho}}\int_{\mathrm{supp}(G)}\theta^{\frac{-\rho}{1-\rho}}dG(\theta)$. Thus, the firm's value is solely determined by $\widehat{\theta}(G)$ and the moment $\int_{\mathrm{supp}(G)}\theta^{\frac{-\rho}{1-\rho}}dG(\theta)$. Since mandatory disclosure cannot perfectly reveal θ, the optimal information structure should provide information in the form of the moment, which is our second result. The maximum entropy condition comes from the fact that without any additional information the best response for the investor is to choose the most smooth distribution in the sense of maximizing entropy.

Example 4 *Follow example 1 suppose that the prior distribution of θ has the probability density function f on $[1,2]$ and $\rho = \frac{2}{3}$. An example of an optimal information structure Λ is as follows:* $\Lambda = \{(1, 0.7, \pi_1(\theta)), (1.1, 0.65, \pi_2(\theta)),$

$(1.3, 0.5, \pi_3(\theta)), (1.6, 0.4, \pi_4(\theta))\}$*, where $\pi_i(\theta)$ is the probability of disclosing the specific lower bound and moment when the true value is θ that satisfies $\sum_i \pi_i(\theta) = 1\ \forall\theta$ and the Bayesian plausibility condition. Thus, if the disclosed* *posterior is* $(1, 0.7)$*, investors perceive that the posterior distribution of* θ *conditional on this disclosure, denoted by* G_1*, is the most smooth distribution on* $[1, 2]$ *in the sense of entropy maximizing with the constraint that* $\int_1^2 \theta^{-2} dG_1(\theta) = 0.3$*. Similarly, when the disclosed posterior is* $(1.3, 0.5)$*, investors perceive that the posterior distribution of* θ *conditional on this disclosure, denoted by* G_1*, is the most smooth distribution on* $[1.3, 2]$ *in the sense of entropy maximizing with the constraint that* $\int_{1.3}^2 \theta^{-2} dG_1(\theta) = 0.5$*. The number of the posterior distributions, the probability of choosing each posterior distribution, the lower bounds of each posterior distributions and the magnitude of the moment are jointly determined by the total quantity of information contained in the information structure.*

Figure 1 shows the shape of the entropy-maximizing distribution G_1 given the support of the prior distribution of θ is on [1.2], the lower bound of the posterior is 1 and how the shape varies with the moment. [24] From the graph is is very clear that although uniform distribution always maximizes the entropy given no moment constraint, the moment constraint places additional structure on the distribution. More interestingly, the more the moment differs from the moment of the uniform distribution, the more the information is contained in the posterior as investors will be more certain whether θ is close to the lower bound or upper bound.

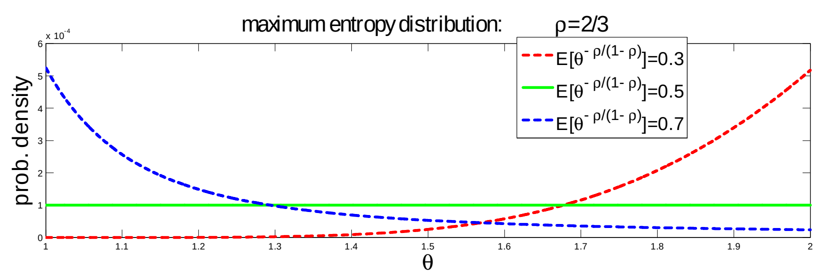

Figure 1 A graphical illustration of a maximum entropy distribution with respect to different values of the moment.

Our results of the optimal disclosure rule can be interpreted as being con-

[24] Note that this graph is not meant to capture the comparative statics of G_1 with respect to the moment since in equilibrium both the moment and the lower bound are simultaneously determined which then determine the entropy-maximizing distribution. Changing the moment alone without changing the lower bound has no meaning.

sistent with the conservatism principle embedded in the financial reports that a firm discloses periodically. Conservative accounting principles result in firms disclosing a lower bound of their net assets on the balance sheet, consistent with the first element of the optimal disclosure rule. Numerous examples exist in accounting that makes balance sheet a lower bound of firm value, e.g., treatment of research and development expenditures as expenses and the lower-of-cost-or-market rule of accounting for inventory. Our paper thus provides a rationale for the conservatism principle embedded in many disclosure standards and we derive our results in a fairly general setting with information asymmetry being the only friction.

The financial reports, in particular the income statement, also provide a summary of firm value conditional on the firm can continue as a going concern. To the extent that the firm value being higher than the lower bound implies that the firm can continue as a going concern, the summary statistic then provides more information about firm value based on firms continuing and its value exceeding this lower bound. In particular, it provides information for investors to figure out the posterior distribution and thus all the related statistics conditional on this posterior distribution, similar to the real life situation when sophisticated investors and analysts use earnings numbers in their valuation models.

As a summary, the uniquely optimal disclosure rule in our setting can be interpreted to be consistent with the conservatism, arguable one of the most important attributes of financial disclosure. Conservatism, as reflected in disclosing a lower bound, stems from the fact that inefficiency in a signalling game comes from high types taking inefficient actions to prevent low types from mimicking them. Disclosing a verifiable lower bound is then the most efficient way to prevent low types from mimicking without resorting to inefficient actions. Firm also discloses a summary statistic, which comes from the fact that the expected value of the firm can be summarized by the lower bound and a moment of the posterior distribution. Given the disclosure of lower bound, the moment is a sufficient statistic of the firm value. Because we impose no particular structure on the disclosure rules, our results are as general as the existence of information asymmetry between firm and outside investors that cannot be completely eliminated by disclosure.

To the extent that balance sheet serves as a lower bound of firms' net assets and net income serves as a summary measure of firm value, we also show that both the balance sheet and the income statement are relevant for a firm's valuation since both the lower bound and the summary statistic appears in firm's expected value under equation (5). Ohlson (1995) also shows the value relevance of both balance sheet and 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 disclosure rule is derived endogenously. Furthermore, instead of showing that firm value is linearly related to book value and net income, our results showed that the relation can be quite complex and the relation itself is a function of the optimal disclosure rule. Future research may explore further along this path and derive some empirically testable valuation models while taking into account the strategic consequences of disclosure.

4 Concluding Remarks

We study the properties of optimal disclosure 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 outside investors, resulting in costly signalling of the percentage of shares retained by the firm. However, before observing any private information, firms can choose to commit to a disclosure rule that will provide outside investors a noisy signal of the private information. We show that so long as the disclosure cannot perfectly reveal firms' private information, the uniquely optimal mandatory disclosure rule always consists of the following: 1) disclosure of a lower bound of the posterior distribution and 2) disclosure of a moment of the posterior distribution which together with the lower bound consists of a sufficient statistic of the firm value. This optimal disclosure rule can be interpreted as a justification for the accounting standard that firms use in producing financial reports. In particular, the disclosure of the lower bound is consistent with the conservatism principle embedded in the accounting rules resulting in firms' balance sheet being a reflection of the lower bound of their value. Our result theoretically justifies conservatism, one of the most important attributes of mandatory disclosure, in a setting that is particularly relevant for mandatory disclosure.

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 (2013)) 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, these papers didn't address the qualitative properties of accounting information in a fairly general way, which is the focus of our paper. Thus, rather than rationalizing conservative accounting with respect to self-defined neutral or aggressive accounting, conservative accounting rule appears as one of the optimal qualitative properties of accounting information.

We believe our introduction of the qualitative properties of information structure is especially relevant for financial reporting because financial reporting often needs to make trade-offs between qualitative properties with the impact on the quantitative properties less obvious, with conservatism versus aggressiveness being one example and principles-based disclosure standard versus rules-based disclosure standard being another. From this point of view, our paper can be seen as a benchmark with future research incorporating more institutional structure to generate more insights related to optimal accounting standards. For example, we do not consider the possibility of firm manipulating report while not violating the rules. One can, in spirit, use the two-step representation introduced in Gao (2013) to study the relation between the extent of earnings manipulation and the properties of optimal accounting rules.

5 Appendix

Proof of Lemma 1:

Proof. The proof follows the proof of proposition 2 in DeMarzo and Duffie (1999). Since from equation (2) we know that $p(q) = E[\theta | q(\theta), G] = E[\theta | \left[\frac{\widehat{\theta}(G)}{\theta}\right]^{\frac{1}{1-\rho}}, G] = \theta$ 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 $u_f(\theta, q, p) = q \cdot (p - \rho\theta)$. 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 a ordinary differential equation of $\frac{dp}{p} = \frac{1-\rho}{q} dq$ with 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. ■

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 (5), in the signaling equilibrium the firm's expected payoff under belief $w \cdot G_1 + (1-w) \cdot G_2$ is

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

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

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

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

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

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 supp(Λ_2) to supp(Λ_1), expressed as $\Gamma \in \Delta\left(\Delta\left(\text{supp}\left(\Lambda_2\right)\right)\right)$, a distribution of probability measures over supp(Λ_2) such that

$$\Lambda_2 = \int \nu\Gamma\left(d\nu\right) \tag{7}$$

and

$$\text{supp}\left(\Lambda_1\right) = \left\{\int_{\text{supp}(\nu)} G\nu\left(dG\right) : \nu \in \text{supp}\left(\Gamma\right)\right\} , \tag{8}$$

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

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

where the second equality follows (8), the inequality follows the convexity of $U_f\left(\cdot\right)$, and the third equality follows (7).

This concludes the proof. ■

Proof of Proposition 2:

Proof. The firm's problem is

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

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

Proposition 1 and Corollary 1 imply that the constraint binds. Let $\mu_\kappa > 0$ denote the corresponding Lagrangian multiplier. The firm's optimal information structure can be solved from the dual problem

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

for appropriately chosen $\mu_\kappa > 0$. The dual problem can be rewritten as

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

For the sake of expression, we will focus on this dual problem. 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, and thus $1-w$ for G_2, respectively. Then the firm can benefit from combining G_1 and G_2, since $U_f\left(w \cdot G_1 + (1-w) \cdot G_2\right) = w \cdot U_f\left(G_1\right) + (1-w) \cdot U_f\left(G_2\right)$ according to Lemma 2 and $h\left(w \cdot G_1 + (1-w) \cdot G_2\right) > w \cdot h\left(G_1\right) + (1-w) \cdot h\left(G_2\right)$. Thus we prove i). For any G $\in$supp(Λ_κ), note that $U_f\left(G\right)$ is fixed when $\widehat{\theta}\left(G\right)$ and $\int_{\left[\widehat{\theta}(G),\overline{\theta}\right]} \theta^{\frac{-\rho}{1-\rho}} dG\left(\theta\right)$ are fixed. Therefore, given $\widehat{\theta}\left(G\right)$ and $\int_{\left[\widehat{\theta}(G),\overline{\theta}\right]} \theta^{\frac{-\rho}{1-\rho}} dG\left(\theta\right)$, we need to choose $G \in \Delta\left(\left[\underline{\theta}, \overline{\theta}\right]\right)$ such that $h\left(G\right)$ is maximized, which requires that 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, George A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism". *The Quarterly Journal of Economics*, 84(3): 488-500.

Bertomeu, Jeremy, Masako Darrough and Jason Xue. 2013. "Agency Conflicts, Earnings Management, and Conservatism" Baruch College, Working Paper.

Bertomeu, Jeremy, Anne Beyer and Ronald A. Dye. 2011. "Capital Structure, Cost of Capital, and Voluntary Disclosures." *The Accounting Review* 86(3): 857-886.

Beyer, Anne. 2013. "Conservatism and Aggregation: The Effect on Cost of Equity Capital and the Efficiency of Debt Contracts". Stanford University, Working Paper.

Beyer, Anne, Daniel A. Kohen, Thomas Z. Lys and Bevely R. Walther. 2010. "The Financial Reporting Environment: Review of the Recent Literature." *Journal of Accounting and Economics* 50(2-3): 296-343.

Bhattacharya, Sudipto. 1979. "Imperfect Information, Dividend Policy, and "the Bird in the Hand" Fallacy." *The Bell Journal of Economics* 10(1): 259-270.

Burkhardt, Katrin, and Roland Strausz. 2009. "Accounting Transparency and the Asset Substitution Problem." *The Accounting Review* 84(3): 689-712.

Caskey, Judson, and John S. Hughes. 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, Judson, and Volker Laux. 2013. "Effects of corporate governance and managerial optimism on accounting conservatism and manipulation". University of Texas at Austin, Working Paper.

Crawford, Vincent, and Joel Sobel. 1982. "Strategic Information Transmission". *Econometrica* 50(6): 1431-1451.

Chen, Qi, Thomas Hemmer, and Yun Zhang. 2007. "On the Relation Between Conservatism in Accounting Standards and Incentives for Earnings Management." *Journal of Accounting Research* 45(3): 541-565.

Cover, Thomas M. and Joy. A. Thomas. 1991. *Elements of Information Theory.* New York, NY, UAS: Wiley-Interscience.

DeMarzo, Peter, and Darrell Duffie. 1999. "A Liquidity-Based Model of Security Design." *Econometrica* 67(1): 65-99.

Dye, Ronald A. 1985. "Disclosure of Nonproprietary Information." *Journal of Accounting Research* 23(1): 123-145.

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)".

Gao, Pingyang. 2010. "Disclosure Quality, Cost of Capital and Investors' Welfare". *The Accounting Review* 85(1): 1-29.

Gao, Pingyang. 2013. "A Measurement Approach to Conservatism and Earnings Management." *Journal of Accounting and Economics, forthcoming.*

Gigler, Frank B., and Thomas Hemmer. 2001. "Conservatism, Optimal Disclosure Policy, and the Timeliness of Financial Reports." *The Accounting Review* 76(4): 471-493.

Gigler, Frank, Chandra Kanodia, Haresh Sapra, and Raghu Venugopalan. 2009. "Accounting Conservatism and the Efficiency of Debt Contracts." *Journal of Accounting Research* 47(3): 767-797.

Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure About Product Quality." *Journal of Law and Economics* 24(3): 461-483.

Jiang, Xu. 2013. "Biased Accounting Information and Private Information Acquisition." Duke University, Working Paper.

Jovanovic, Boyan. 1982. "Truthful Disclosure of Information." *The Bell Journal of Economics* 13(1): 36-44.

Jung, Woon-Oh, and Young K. Kwon. 1988. "Disclosure When the Market is Unsure of Information Endowment of Managers." *Journal of Accounting Research* 26(1): 146-153.

Kamenica, Emir and Matthew Gentzkow. 2011. "Bayesian Persuasion". *American Economic Review* 101(6): 2590-2615.

Kanodia, Chandra. 2006. "Accounting Disclosure and Real Effects". *Foundations and Trends in Accounting* 1(3): 167-258.

Kanodia, Chandra, and Deokheon Lee. 1998. "Investment and Disclosure: The Disciplinary Role of Periodic Performance Reports." *Journal of Accounting Research* 36(1): 33-55.

Kanodia, Chandra, Rajdeep Singh, and Andrew E. Spero. 2005. "Imprecision in Accounting Measurement: Can It Be Value Enhancing?." *Journal of Accounting Research* 43 (3): 487-519.

Hayne E. Leland, and David H. Pyle. 1977. "Informational Asymmetries, Financial Structure, and Financial Intermediation." *The Journal of Finance* 32(2): 371-387.

Li, Jing. 2013. "Accounting Conservatism and Debt Contracts: Efficient Liquidation and Covenant Renegotiation." *Contemporary Accounting Research*, Forthcoming.

Mailath, George. 1987. "Incentive Compatibility in Signalling games with a Continuum of Types" *Econometrica* 55(6): 1349-1365.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics* 12(2): 380-391.

Miller, Merton H., and Kevin Rock. 1985. "Dividend Policy Under Asymmetric Information." *The Journal of Finance* 40(4): 1031-1051.

Plantin Guillaume, Haresh Sapra and Hyun Song Shin. 2008. "Marking-to-Market: Panacea or Pandora's Box?" *Journal of Accounting Research* 46(2): 435-460.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50(3): 665-690.

Sims, Christopher A. 2005. "Rational Inattention: a Research Agenda." No. 2005, 34. Deutschen Bundesbank Research Center Discussion paper Series 1: Economic Studies 2005, 34.

Stocken, Phillip. 2012. "Strategic Accounting Disclosure." *Foundations and Trends in Accounting.* 7(4): 197-291.

Verrecchia, Robert E. 1983. "Discretionary Disclosure." *Journal of Accounting and Economics* 5: 179-194.

Yang, Ming. 2012. "Coordination with Rational Inattention". Duke University, Working Paper.