Accounting Conservatism, Aggregation, and Information Quality*

Qintao Fan†

Xiao-jun Zhang‡

Abstract

This paper demonstrates that conservative aggregation in accounting often improves the overall quality of information produced, and therefore enhances the welfare of accounting information users. We study the optimal accounting policy when a firm can control the quality of accounting information through costly and noncontractible action. In our model, the accounting system not only affects the quality of reported information ex post, but also the quality of information generation ex ante. It is shown that the desirable accounting has two key features: (i) the accounting report aggregates, rather than reports directly, the underlying information; and (ii) the accounting has a conservative bias. By invoking conservative aggregation, which serves as a commitment to an apparently inefficient accounting scheme given the ex post information quality, firms are induced to spend more effort controlling information quality. In equilibrium, the level of conservatism might even increase with the opportunity loss associated with being overly prudent.

JEL Classification: D81; M41; M48

---

*We are grateful to two anonymous referees and the editor for valuable comments. We also wish to thank Gary Biddle, Qi Chen, Sunil Dutta, Jennifer Francis, Edward Riedl, Jim Wahlen, Yun Zhang, and participants at the 2005 UNC-Duke Fall Camp and the 2006 HKUST Summer Camp for comments and suggestions.

†University of California at Berkeley, fan@haas.berkeley.edu

‡University of California at Berkeley, xzhang@haas.berkeley.edu

1 Introduction

This paper studies the effect of aggregation and conservatism on accounting information quality. Aggregation refers to the practice of summarizing raw data into key financial measures with a limited amount of disclosure. This process involves data compression, and leads inevitably to a loss of information. Conservatism, which Sterling (1970) rates as the most influential accounting principle, is defined in the Statement of Financial Accounting Concepts No. 2 as "a prudent reaction to uncertainty... If two estimates of amounts to be received or paid in the future are about equally likely, conservatism dictates using the less optimistic estimate."

The combination of aggregation and conservatism introduces bias and may impair the usefulness of accounting information to decision making. Relative to the true probability distribution of the underlying economic states, a conservative accounting system generates signals that increase the likelihood of classifying firms as being in an unfavorable state.[1] Such bias, unless it can be quantified or adjusted for by the end user, runs the risk of producing misleading information. It is noted by the Financial Accounting Standards Board that conservatism "tends to conflict with significant qualitative characteristics, such as representational faithfulness, neutrality, and comparability (including consistency)." Similar views are expressed by Hendriksen and Van Breda 1992: "Conservatism is, at best, a very poor method of treating the existence of uncertainty in valuation and income. At its worst, it results in complete distortion of accounting data."

The above arguments focus on comparing the information quality under conservative and neutral accounting regimes, taking as given the quality of the underlying information. What's

[1] As stated by the Financial Accounting Standards Board in the SFAC No.2, conservatism causes "possible errors in measurement in the direction of understatement rather than overstatement of net income and net assets."

missing is the recognition that the information quality might be endogenous. This paper takes the analysis one step further. We analyze the case where the information originator, typically the firm, covertly controls the quality of information generated at a private cost. After the information is generated, the accounting aggregation is executed. In this situation, the accounting system not only affects the quality of reported information ex post, but also the quality of information generation ex ante.

We show that conservative aggregation in accounting often increases the quality of accounting information. We model the accounting system as an information processing scheme in the presence of uncertainty. It is assumed that the firm prefers a favorable accounting report, whereas users of the accounting information are more concerned with its accuracy for decision making. Given this divergence in preferences, we show that a conservative accounting system positively affects the firm's propensity to provide more accurate information. This conclusion is derived from the following rationale: given that the information originator prefers to be classified as having a favorable state of affairs, his expected payoff decreases with a conservative accounting system. However, this decrease is less severe when the underlying information signal is more accurate. Hence, an increased level of conservatism enhances the firm's motivation to provide accurate information.

Understanding aggregation and conservatism is without a doubt important to the accounting profession. Our analysis brings forth one fundamental insight. By imposing conservative aggregation, the accounting system effectively links the financial reporting outcome to the unobservable precision of the underlying information. This mechanism provides an incentive for the firm to increase information quality ex ante, which in turn enhances the welfare of accounting information users. Through comparative statics analysis, we further show that in certain circumstances, the level of conservatism even increases with the opportunity loss of being overly prudent. This result strongly favors the adoption of a conservatively biased system.

The rest of the paper is organized as follows. Section 2 briefly discusses the related literature and highlights the intended contribution of this paper. Section 3 describes the basic model and analyzes the information user's primary decision problem with an exogenously fixed information quality. Section 4 discusses our major results with respect to conservatism and aggregation when the information quality is endogenously derived. Section 5 concludes the paper.

2 Related Literature

Our study is closely related to two areas of accounting research: aggregation and conservatism. In the literature on aggregation, it is often argued that data reduction is needed to ease information processing costs for end users (Butterworth 1972). This line of research often assumes that aggregation reduces overall information accuracy. Early research, for example, examines the loss in user payoff arising from aggregation (Ronen 1971, Butterworth 1972, Ijiri 1975, Feltham 1975). More recently, Dye and Sridhar (2004) show how aggregate accounting reports can reduce earnings management incentives. We support this line of research by demonstrating how aggregation, which prevents investors from observing detailed data, can serve as a commitment device to achieve ex ante optimum and increase information quality. Moreover, our analysis reveals that in order for aggregation to increase reporting quality, accounting needs to be conservative. This highlights an intrinsic link between the two important properties of accounting, namely, conservatism and aggregation.

This paper also contributes to the literature on accounting conservatism. There exists a variety of informal explanations on legal, tax, and debt contracting causes for conservatism (see Devin 1963, Watts 2003 and references therein).[2] Theoretical research on this subject

[2]Gigler, Kanodia, Sapra, and Venugopalan (2009) provide a formal model on the link between accounting conservatism and debt contracts.

has focused on explaining conservatism's efficient contracting role in agency models, which typically involve moral hazard and risk aversion (see, for example, Bushman and Indjejikian 1993, Kwon, Newman and Suh 2001, and Dutta and Zhang 2002).[3] Christensen and Demski (2004) argue that when the accounting signal can be unmistakably interpreted as "good" or "bad" news, managers have an incentive to selectively report "good" news because their compensation frequently depends on this outcome. To discipline the manager's reporting incentives, it is more important to acquire additional information after receiving a favorable report. On average, such conditional recognition creates a kind of conservative bias. Gigler and Hemmer (2001) argue that firms operating under less conservative financial reporting regimes are more likely to engage in timely preemptive disclosure to facilitate risk sharing between firm manager and shareholders than firms under more conservative regimes. Chen et al. (2007) study a similar agency setting in which the accounting information used for incentive contracting is also used by potential investors to value the firm. They show that imposing a conservative accounting noise dampens the firm owner's incentive to optimistically bias earnings, which in turn improves risk sharing between the firm owner and the manager.

Similar to Chen et al. 2007 and Christensen and Demski 2004, in our paper, the party who reports the accounting information obtains private gains by presenting good news, which conflicts with the social benefits of more informative accounting reports. While Christensen and Demski (2004) study additional information production after the information generation, we, like Chen et al. 2007 and Gigler and Hemmer 2007, study the interaction between conservative accounting systems and endogenous information production in the ex ante stage. Our conclusion is similar to that of Chen et al. 2007, in that accounting conservatism can positively influence the firm owner's reporting incentive. However, unlike Chen et al. 2007 which focuses on opportunistic mean-increasing bias in the accounting reports, we study noise-reducing effort

[3] While most papers in this stream of research focus on the design of conservative accounting systems(i.e., GAAP conservatism), Lin (2006) studies a two-period agency model where a manager's *discretionary* conservative choice in the form of higher first period's depreciation can signal his private information of project type.

choice. Furthermore, in Chen et al. 2007, the firm owner benefits from information accuracy because it leads to more efficient incentive contracts. In contrast, we study how conservatism affects the quality of information used by firm outsiders in decision making. In our setting, the firm owner doesn't directly care about the unobservable quality of the accounting information.

In the broader context, this paper is also related to the growing literature on collective choice and endogenous information production. For example, in a committee decision setting, Li (2001) shows that by making it harder to take the committee's consensus action, each individual committee member has more incentive to increase his fact-finding efforts. Both our paper and that of Li 2001 address the issue of how *ex post* inefficient decision-making can alleviate the *ex ante* incentive problem with information acquisition. However, unlike Li 2001 which focuses on the free-riding problem among committee members, we examine how a biased data generator can increase the quality of raw data. Li (2001) assumes that every committee member cares and influences the classification accuracy. Our model, in contrast, assumes that the information originator has sole control of information quality and cares only about maximizing the probability of a favorable classification. As a result, in Li's model the optimal decision rule is either aggressive or conservative, whereas in ours it is strictly biased downwards.

3 Model Setup

Accounting information is provided by a firm to its end users, who are assumed to be sufficiently numerous that mutual contracting (regarding disclosed accounting data) is prohibitively expensive. We assume that an independent third party, the auditor/accountant, verifies the existence and truthfulness of the underlying data and exactly follows an accounting process defined by a set of pre-specified rules. However, the users and the auditor cannot control the underlying process generating the raw data, and hence cannot control the precision of the raw data. This assumption allows us to neutralize factors associated with earnings management, and to focus on the issue of conservatism and ex ante information quality control.

As a reasonable abstraction of the real world situation, we assume that the users and the firm have divergent preferences over the accounting signal. A representative user's expected utility increases with signal accuracy, whereas the firm's payoff increases with the favorableness of the signal. Note that the latter assumption is easily justified when the utility of the current owners of the firm depends on its market valuation, as is typically assumed in prior literature.

We model accounting as a two-step process. In the first step, after nature draws the state, a signal (i.e, raw data) is obtained regarding that state. The precision of the raw data is controlled by the unobservable action of the firm's manager. In the second step, the accounting process transforms the raw data into a financial report. End users then make appropriate economic decisions based on this published report.

To capture the aggregation aspect of accounting in a simple manner, we use the following model featuring binary classification based on a continuous signal.[4] Specifically, we assume that there are two possible states of nature: $x = 0$ (the bad state) and $x = m$ (the good state), with $m > 0$.[5] Let β denote the *a priori* probability of $x = m$, which is known by both the firm and the users. The raw information y is a noisy but unbiased signal of the state x: $y = x + \epsilon$ where ϵ is normally distributed with zero mean and precision h.[6]

A key notion in our paper is that the information quality h can vary. For simplicity,

[4]The issues of accounting conservatism and information aggregation are closely related. Aggregation ensures that users cannot fully infer the underlying information from the accounting reports. Without such a feature, it is difficult to establish any economic significance for accounting when it is merely an invertible data transformation.

[5]See, for example, Kwon, Newman and Suh (2001), and Gigler and Hemmer (2001) for analysis of conservatism using similar binary structures.

[6]The normality assumption facilitates many mathematical derivations. However, major results hold qualitatively for any distribution function that's symmetric, unimodal and covers $(-\infty, +\infty)$.

we assume that h can take two values, $\underline{h}$ and $\overline{h}$. $\underline{h}$ can be interpreted as a low level of information accuracy achieved through ordinary mechanisms such as internal control. $\overline{h}$ represents the higher level of accuracy arising from the firm's noncontractible and costly effort. More specifically, with probability r, the firm's reporting system generates y with precision $\overline{h}$. With probability $1-r$, the signal y is of quality $\underline{h}$. r is assumed to be at the firm's discretion, and is unobservable.

After observing the noisy signal $y = x + \epsilon$, the accountant produces a report z based on the data y. The accounting policy specifies how this reporting process should be done, In other words,

$$z = g(y),$$

where $g(\cdot)$ denotes the accounting rule. Assume that a representative end user takes one of two possible actions: $a = a_1$ or $a = a_2$. a_1 is assumed to be the correct action to take when the good state ($x = m$) occurs, and a_2 is assumed to be the correct action to take when the bad state ($x = 0$) occurs. Concretely, these actions could represent banks lending (or not lending) to the firm, or analysts issuing a buy/sell recommendation. As we will see later, this binary framework is designed to motivate information aggregation by reducing the welfare loss brought on by transforming the raw information signal y into binary summary measure z. In Appendix B, we demonstrate how this binary setup can be generalized.

3.1 User's primary decision problem

The decision problem of the representative end user, given the noisy accounting signal, can be formulated and solved as follows. Let $U(a, x)$ denote the representative user's utility as a function of the state-action combination (a, x), with

$$U(a,x) \equiv \begin{cases} U_1, \text{ when } a = a_1 \text{ and } x = m \\ U_1 - L_1, \text{ when } a = a_2 \text{ and } x = m \\ U_2, \text{ when } a = a_2 \text{ and } x = 0 \\ U_2 - L_2, \text{ when } a = a_1 \text{ and } x = 0 \end{cases}$$

where $L_1 > 0$ and $L_2 > 0$ denote the losses resulting from incorrect actions. Occurrences of L_1 and L_2 are referred to as type I and type II errors respectively. Notice that the objective function of the representative user is fairly general. For much of the analysis, no further specification of the user's decision set or utility function is required. The users form a decision rule $a(z)$ based on the report z. It is straightforward to show that, as typical in decision-making models under uncertainty, expected utility maximization is equivalent to expected loss minimization from type I and type II errors.

To illustrate the benchmark decision rule of a representative end user, consider the case where the precision of y (i.e., r) is exogenously fixed and the accounting rule is to simply report the underlying information y. That is,

$$z = y \tag{1}$$

We denote this benchmark accounting as R^a. If the state is actually m, then $\epsilon = y - m$. If the state is 0, then $\epsilon = y$. Let $\phi(\cdot)$ denote the density function of a standard normal variable (i.e., with mean zero and variance one). Then, by applying Bayes' rule,

$$\frac{\Pr(x=0|z)}{\Pr(x=m|z)} = \frac{\left[r\overline{h}^{\frac{1}{2}}\phi\left(y\overline{h}^{\frac{1}{2}}\right) + (1-r)\underline{h}^{\frac{1}{2}}\phi\left(y\underline{h}^{\frac{1}{2}}\right)\right](1-\beta)}{\left[r\overline{h}^{\frac{1}{2}}\phi\left((y-m)\overline{h}^{\frac{1}{2}}\right) + (1-r)\underline{h}^{\frac{1}{2}}\phi\left((y-m)\underline{h}^{\frac{1}{2}}\right)\right]\beta}. \tag{2}$$

Thus, the user should take action a_1 if

$$\begin{aligned} &L_2\left[r\overline{h}^{\frac{1}{2}}\phi\left(y\overline{h}^{\frac{1}{2}}\right) + (1-r)\underline{h}^{\frac{1}{2}}\phi\left(y\underline{h}^{\frac{1}{2}}\right)\right](1-\beta) \\ &< L_1\left[r\overline{h}^{\frac{1}{2}}\phi\left((y-m)\overline{h}^{\frac{1}{2}}\right) + (1-r)\underline{h}^{\frac{1}{2}}\phi\left((y-m)\underline{h}^{\frac{1}{2}}\right)\right]\beta. \end{aligned} \tag{3}$$

Note that $d \log \phi\left(\epsilon h^{\frac{1}{2}}\right) / d\epsilon = -\epsilon h$, a decreasing function of ϵ. This implies

$$\frac{d \log\left(\phi((y-m)h^{\frac{1}{2}}) / \phi(yh^{\frac{1}{2}})\right)}{dy} = \frac{\phi'((y-m)h^{\frac{1}{2}})}{\phi((y-m)h^{\frac{1}{2}})} - \frac{\phi'(yh^{\frac{1}{2}})}{\phi(yh^{\frac{1}{2}})} > 0, \tag{4}$$

so $\frac{\phi((y-m)h^{\frac{1}{2}})}{\phi(yh^{\frac{1}{2}})}$ is an increasing function of y. Let w^* be the value of y such that (2) holds as an equality. The value of w^* can be shown to exist.[7] For any $y > w^*$,

$$\frac{\Pr(x=0|y)}{\Pr(x=m|y)} < \frac{\Pr(x=0|w^*)}{\Pr(x=m|w^*)}. \tag{5}$$

Therefore $L_2 \Pr(x=0|y) < L_1 \Pr(x=m|y)$, and the user will take action a_1 when $z = y > w^*$. This threshold value of y characterizes the end user's optimal actions.

The previous analysis indicates that it is possible to reformulate the user's problem as choosing a threshold w^* to minimize the expected loss. For any w, let p denote the probability of correctly classifying the state as bad when the state is indeed bad ($x = 0$), and let q denote the probability of correctly classifying the state as good when the state is indeed good ($x = m$). For a given w and r, the probabilities of correct classification are:

$$\begin{aligned} p\,(w,r) &= \Pr[y \le w | x = 0] \\ &= r\Phi\left(w\overline{h}^{\frac{1}{2}}\right) + (1-r)\Phi\left(w\underline{h}^{\frac{1}{2}}\right), \text{ and} \qquad (6) \\ q\,(w,r) &= \Pr\left[y > w | x = m\right] \\ &= r\left[1 - \Phi\left((w-m)\overline{h}^{\frac{1}{2}}\right)\right] + (1-r)\left[1 - \Phi\left((w-m)\underline{h}^{\frac{1}{2}}\right)\right]. \qquad (7) \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal variable. Notice that w and r simultaneously determine p and q. Hence, from the end user's perspective, the optimal

[7] Existence can be demonstrated as follows using features of the normal distribution. At low values of w, $\phi((y-m)h^{\frac{1}{2}})/\phi(yh^{\frac{1}{2}})$ will approach infinity, hence $\frac{r\overline{h}^{\frac{1}{2}}\phi\left(y\overline{h}^{\frac{1}{2}}\right)+(1-r)\underline{h}^{\frac{1}{2}}\phi\left(y\underline{h}^{\frac{1}{2}}\right)}{r\overline{h}^{\frac{1}{2}}\phi\left((y-m)\overline{h}^{\frac{1}{2}}\right)+(1-r)\underline{h}^{\frac{1}{2}}\phi\left((y-m)\underline{h}^{\frac{1}{2}}\right)}$ will go to zero. On the other hand, at high values of w, $\frac{\phi((y-m)h^{\frac{1}{2}})}{\phi(yh^{\frac{1}{2}})}$ will approach 0 and hence $\frac{r\overline{h}^{\frac{1}{2}}\phi\left(y\overline{h}^{\frac{1}{2}}\right)+(1-r)\underline{h}^{\frac{1}{2}}\phi\left(y\underline{h}^{\frac{1}{2}}\right)}{r\overline{h}^{\frac{1}{2}}\phi\left((y-m)\overline{h}^{\frac{1}{2}}\right)+(1-r)\underline{h}^{\frac{1}{2}}\phi\left((y-m)\underline{h}^{\frac{1}{2}}\right)}$ will go to infinity.

choice of $w^*(r)$ for a given r should minimize the expected loss, that is,

$$\min_{w}\ (1-p(w,r))(1-\beta)L_2+(1-q(w,r))\beta L_1 \tag{8}$$

For a given r, the optimal classification threshold $w^*(r)$ satisfies the following first-order condition[8]:

$$(1-\beta)L_2 p_w+\beta L_1 q_w=0. \tag{9}$$

The following proposition summarizes the properties of the optimal $w^*(r)$ and also provides a useful benchmark for later derivations:

PROPOSITION 1 *For a given r, the optimal $w^*(r)$ is uniquely determined and satisfies:*

$$\frac{r\bar{h}^{\frac{1}{2}}\phi\left(w^*\bar{h}^{\frac{1}{2}}\right)+(1-r)\underline{h}^{\frac{1}{2}}\phi\left(w^*\underline{h}^{\frac{1}{2}}\right)}{r\bar{h}^{\frac{1}{2}}\phi\left((w^*-m)\bar{h}^{\frac{1}{2}}\right)+(1-r)\underline{h}^{\frac{1}{2}}\phi\left((w^*-m)\underline{h}^{\frac{1}{2}}\right)}=\frac{\beta L_1}{(1-\beta)L_2}. \tag{10}$$

(i) *$w^*(r)$ is decreasing in $\frac{\beta L_1}{(1-\beta)L_2}$. $w^*(r)\lesseqgtr\frac{m}{2}$ if $\beta L_1\gtreqless(1-\beta)L_2$.*

(ii) When $\beta L_1\neq(1-\beta)L_2$, $\left|w^*(r)-\frac{m}{2}\right|$ *decreases with r.*

(iii) *For any* $\bar{h}$ and r, *$w^*(r)\in(0,m)$ when $\underline{h}>\frac{2}{m^2}\max\left\{\ln\frac{\beta L_1}{(1-\beta)L_2},\ln\frac{(1-\beta)L_2}{\beta L_1}\right\}$.*

PROOF. See Appendix. ■

The representative user's problem in choosing the optimal classification threshold can be understood as follows. The user is willing to risk taking action a_1 if there is sufficient chance of getting the high state $x=m$. At $y=w^*$, the expected loss from getting the high state $x=m$ just balances the expected loss from getting the low state $x=0$. At any lower value of $y<w^*(r)$, the expected loss from getting a low state when taking action a_1 outweighs the loss of getting the high state when action a_2 is taken. At at any higher value of $y>w^*(r)$,

[8]Throughout the paper we use letter subscripts to denote derivatives.

the expected loss from getting a high state when taking action a_2 outweighs the loss of getting the low state if action a_1 is taken. The optimal threshold $w^*(r)$ generally doesn't equal $m/2$. The deviation of the optimal threshold $w^*(r)$ from $\frac{m}{2}$ depends on $\frac{\beta L_1}{(1-\beta)L_2}$, and can be either positive or negative. In the special case where $\beta L_1 = (1-\beta)L_2$, we have $w^*(r) = \frac{m}{2}$, which is independent of $\underline{h}$ and $\overline{h}$. Further, given the commonly known prior and the user's loss function, higher information quality r of the signal makes the optimal $w^*(r)$ closer to $\frac{m}{2}$. In other words, better information quality reduces the effect of the prior or loss function on the *ex post* optimal threshold. When $\underline{h} > \frac{2}{m^2}\max\left\{\ln\frac{\beta L_1}{(1-\beta)L_2}, \ln\frac{(1-\beta)L_2}{\beta L_1}\right\}$, for any $r > 0$, $w^*(r) \in (0, m)$. To ensure that the results are easily interpretable, throughout the paper we assume that $\underline{h} > \frac{2}{m^2}\max\left\{\ln\frac{\beta L_1}{(1-\beta)L_2}, \ln\frac{(1-\beta)L_2}{\beta L_1}\right\}$ holds.

The end user's threshold-determined action strategy gives rise to the following accounting policy, which aggregates the raw data in the following way:

$$z = g(y) = \begin{cases} m \text{ if } y > w \\ 0 \text{ if } y \le w \end{cases} \tag{11}$$

In other words, as an information aggregation mechanism, the accounting system takes the form of a threshold classification scheme. It compares the noisy accounting signal y with a threshold w : when y is greater than w, the system reports $z = m$, and when $y \le w$, the system reports $z = 0$. In this accounting strategy, denoted R^b, accountants aggregate the underlying information y into summary data, z. Such aggregation makes the end user's task easier.

Let us define the expected loss of the user as a function of w for a given r as

$$L(w, r) = (1 - p(w, r))(1 - \beta)L_2 + (1 - q(w, r))\beta L_1,$$

so that the following Corollary arises:

COROLLARY 1 *For any $w \in (0, m)$, the user's expected loss $L(w, r)$ due to wrongful classification decreases with r, that is, $L_r < 0$.*

PROOF. See Appendix. ■

Corollary 1 shows that the user's expected loss decreases with the firm's effort (r) in controlling the accounting information quality. As a result, it is in the end user's interest to motivate the firm to improve accounting information quality. Next we show how a choice of accounting policy can be used to achieve this objective. We define conservative, neutral, and aggressive accounting as follows.

Definition: Accounting is conservative (neutral, aggressive) when $w > (=, <) w^* (r)$.

The accounting $w > w^* (r)$ is conservative because, given the precision level of the underlying signal, it is more likely to issue an unfavorable report (i.e., $z = 0$).[9]

3.2 Endogenous information quality

We assume that the firm prefers to be identified with the favorable state m. In particular, the firm gets payoff S_2 when $a = a_1$, and payoff S_1 when $a = a_2$, with $S_2 > S_1 > 0$. This assumption is motivated by the observation that payoffs to a firm (or to its manager) are usually positively correlated with the perceived prospects of the firm.[10] In our setting, for a given information quality, a higher w reduces the probability of the firm being classified into the good state and reduces the firm's expected payoff.

The firm cannot bias or withhold y, but it nevertheless has control over the quality of

[9]Note that conservative accounting is defined relative to $w^*(r)$, not $m/2$. The deviation of $w^*(r)$ from $m/2$ captures other causes of conservatism analyzed in prior studies, such as the asymmetric loss function of the information users.

[10]Notice that conservatism, as defined in our model, increases the precision of good news but decreases the precision of bad news. In a more general model, it is possible that S_1 and S_2 change with the information content of the accounting report. However, the major conclusion of our paper holds as long as S_1 is less than S_2.

y through its effort r that increases the possibility of the accounting signal having a higher precision $\overline{h}$. Effort is costly, and the strictly convex and increasing cost is labelled $kc(r)$ with $c_r(r) > 0$, $c_r(0) = c(0) = 0$, $c_{rr} > 0$ $\forall r > 0$. k is a positive constant. The firm's effort and costs are not observable, and are at the firm's discretion. Thus, given $\underline{h}$ and $\overline{h}$, r essentially indexes the accounting information quality.[11]

The timing is as follows. The accounting rule is first set with a fixed threshold w. The firm then chooses effort r, nature subsequently draws state x, and the noisy signal $y = x + \epsilon$ is observed by the accountant. A report z is generated based on y, the information user chooses his action based on z, and payoffs for both parties are realized.

4 Information aggregation and accounting conservatism

For a given w, the firm chooses an optimal accounting information accuracy r^* to solve

$$\max_{r\in[0,1]}[(1-\beta)(1-p(w,r))+\beta q(w,r)]S_2+[(1-\beta)p(w,r)+\beta(1-q(w,r))]S_1-kc(r). \quad (12)$$

The sum of the first two terms, $[(1-\beta)(1-p(w,r))+\beta q(w,r)]S_2+[(1-\beta)p(w,r)+\beta(1-q(w,r))]S_1$, represents the expected private benefits to the firm: the probability of being classified in a particular state multiplied by the corresponding payoff. The last term is the cost of the internal control effort to the firm. As $q_{rr} = p_{rr} = 0$ and $kc_{rr} > 0$, this expression is globally concave in r. Denote $S_2 - S_1$ by ΔS, and let B_r denote the marginal benefits to the firm from increasing r,

$$\begin{aligned} B_r &= \frac{1}{2}(q_r - p_r)\Delta S \\ &= \Delta S\left\{\beta\left[\Phi\left((w-m)\underline{h}^{\frac{1}{2}}\right)-\Phi\left((w-m)\overline{h}^{\frac{1}{2}}\right)\right]-(1-\beta)\left[\Phi\left(w\overline{h}^{\frac{1}{2}}\right)-\Phi\left(w\underline{h}^{\frac{1}{2}}\right)\right]\right\}. \quad (13)\end{aligned}$$

[11] Since r and $c(r)$ are non-verifiable, binding contracts conditioned on the accuracy of the signal can not be written.

The first order condition for (12) is:

$$B_r - kc_r = 0, \tag{14}$$

which is sufficient, as well as necessary, and admits a unique solution.[12] Again, we see the dissonance of preferences: in choosing the optimal r^*, the firm equates its marginal payoff, $B_r = (\beta q_r - (1-\beta)p_r)\Delta S$, with its marginal cost of effort, kc_r. Given that a firm always prefers the favorable classification, it may disregard the benefits of higher quality accounting information to the user.

Taking the derivative of B_r with respect to w, we have

$$\begin{aligned} B_{rw} = \Delta S \Big\{ & \left[\beta \underline{h}^{\frac{1}{2}} \phi\left((w-m)\underline{h}^{\frac{1}{2}}\right) + (1-\beta)\, \underline{h}^{\frac{1}{2}} \phi\left(w\underline{h}^{\frac{1}{2}}\right)\right] \\ & - \left[\beta \overline{h}^{\frac{1}{2}} \phi\left((w-m)\overline{h}^{\frac{1}{2}}\right) + (1-\beta)\, \overline{h}^{\frac{1}{2}} \phi\left(w\overline{h}^{\frac{1}{2}}\right)\right] \Big\}. \end{aligned} \tag{15}$$

Therefore $B_{rw} \neq 0$ when

$$\underline{h}^{\frac{1}{2}} \left[\beta\phi\left((w-m)\underline{h}^{\frac{1}{2}}\right) + (1-\beta)\,\phi\left(w\underline{h}^{\frac{1}{2}}\right)\right] \neq \overline{h}^{\frac{1}{2}} \left[\beta\phi\left((w-m)\overline{h}^{\frac{1}{2}}\right) + (1-\beta)\,\phi\left(w\overline{h}^{\frac{1}{2}}\right)\right] \tag{16}$$

Thus, except for very special circumstances, the firm's choice of information precision (r) would be affected by the threshold w.

4.1 Desirability of information aggregation

Note that the information user's first order condition is:

$$\begin{aligned} \frac{\partial L}{\partial w} &= \frac{\partial[(1-\beta)\,(1-p)\,L_2 + \beta(1-q)L_1]}{\partial w} \\ &= -(1-\beta)L_2 p_w - \beta L_1 q_w + \frac{\partial r}{\partial w}\frac{\partial[(1-\beta)\,(1-p)\,L_2 + \beta(1-q)L_1]}{\partial r} \end{aligned} \tag{17}$$

[12] In the remainder of the paper, we focus on the case where (12) has an interior solution and (14) holds. Corner solution is an easy extension but doesn't offer any interesting insight.

Consider the case of a disaggregated accounting policy, denoted R^d, in which the underlying signal y is directly disclosed. Since we assume that the number of end users is large enough so that contracting with them individually on the use of accounting information is prohibitively expensive, each user will, based on his conjecture of the unobservable r (which must be correct in equilibrium), use the *ex post* optimal decision rule $w^*(r)$. Therefore, $-(1-\beta)L_2 p_w - \beta L_1 q_w = 0$. However, this w^* is not optimal *ex ante*. From Corollary 1, $\frac{\partial[(1-\beta)(1-p)L_2+\beta(1-q)L_1]}{\partial r} > 0$ for any $w \in (0, m)$. Condition (16) further implies $\partial r / \partial w \neq 0$. Therefore the *ex ante* optimal level $w^{**} \neq w^*(r)$.

In contrast, consider an accounting regime with aggregation (R^a) where the underlying data y is summarized into a bivariate signal z. Since only z is reported, users can no longer use the *ex post* optimal decision rule w^* (given that y is not observable). Therefore, the accounting rule R^a can achieve the *ex ante* optimum by setting

$$z = \begin{cases} 0 \text{ when } y \leq w^{**} \\ m \text{ when } y > w^{**} \end{cases} \tag{18}$$

More formally, we have the following proposition.

PROPOSITION 2 *Assume condition (16) holds such that $B_{rw} \neq 0$. Let $\{w^d, r^d\}$ denote the equilibrium choice of w and r under the disaggregated accounting regime (R^d), and let $\{w^a, r^a\}$ denote the optimal choice of w and the induced r under the aggregated accounting regime (R^a). The end users' expected loss is less under the aggregated accounting scheme, R^a, than it is under the disaggregated accounting scheme, R^d:*

$$L(w^a, r^a) < L(w^d, r^d)$$

In addition, under R^a, the accounting is biased in the sense that the optimal choice of the threshold w^a is different from that of the ex post optimal threshold for the user's primary decision problem, $w^*(r^a)$, given the firm's induced choice of r^a.

PROOF. See Appendix. ■

In essence, the aggregation feature of the accounting policy R^a serves as a commitment device. With endogenous information quality r, accounting aggregation increases the accuracy of the raw data. This result highlights a key benefit of accounting aggregation, which in addition to reducing information processing costs, improves the overall quality of the information..

We would like to note that, in our model, a bivariate decision setting for the representative user is used to simplify the exposition. In a more general setup, aggregation in the form of (18) will likely incur additional cost due to information loss (e.g., Ronen 1971, Butterworth 1972, Ijiri 1975, Feltham 1975). The benefit of aggregation due to the increase in the quality of raw data needs to be weighed against those added costs.

4.2 Accounting conservatism and information quality

An interesting aspect of the above aggregated accounting scheme is its bias. In this section we further explore the issue.

PROPOSITION 3 *For any fixed $\underline{h}$, as long as $\overline{h}$ is large enough, there always exists an interval $W(\underline{h};\overline{h})$ such that $B_{rw} > 0$ for all $w \in W$. Furthermore, such an interval expands as $\overline{h}$ increases, that is, $\overline{h}^1 > \overline{h}^2$ implies that $W(\overline{h}^2;\underline{h}) \subset W(\overline{h}^1;\underline{h})$, and approaches $(0, m)$ as $\bar{h}$ approaches infinity.*

PROOF. See Appendix. ■

$B_{rw} > 0$ implies that a higher classification threshold w leads to a higher marginal benefit of increasing r for the firm. Given that c_r is unrelated to w, from the usual comparative statics argument, the following Corollary shows that the choice of r increases with w.

COROLLARY 2 *When $B_{rw} > 0$, the user's optimal r^* increases with w, that is, $\frac{dr^*}{dw} > 0$.*

PROOF. This again follows the standard monotone comparative statics argument. Notice that total differentiation of (14), the firm's first order condition in choosing r^*, yields

$$\frac{\partial r^*}{\partial w} = \frac{[\beta q_{rw} - (1-\beta)p_{rw}]\Delta S}{kc_{rr}} = \frac{B_{rw}\Delta S}{kc_{rr}} > 0 \tag{19}$$

Thus, when $B_{rw} > 0$, a higher w leads to a higher choice of r. ■

When the user chooses the optimal classification threshold, he takes into account the effect of w on the firm's information quality effort. The following proposition shows that w^{**} is greater than the optimal *ex post* $w^*(r)$.

PROPOSITION 4 *When $B_{rw} > 0$ so that the firm's choice of information quality r increases in w, the optimal choice of the threshold level w^{**} is higher than the optimal threshold level for the user's primary decision problem, $w^*(r)$, given the firm's induced choice of r. In other words, the accounting system embodied by the threshold level w^{**} is conservative.*

PROOF. See Appendix. ■

This proposition shows that, given the firm's preference of being classified in the favorable state, a more conservative classification threshold increases the information quality control effort by the firm. To understand this intuition, consider the benchmark case in which the end users of the financial information have a symmetric loss function(i.e., $L_1 = L_2$), and the two states occur with equal likelihood(i.e., $\beta = 1/2$). From Proposition 1, it's easy to see that the optimal classification threshold should be $w^* = \frac{m}{2}$, regardless of h. With such accounting, $z = 0$ and $z = m$ occur with equal likelihood, correctly reflecting the underlying probability of the true states. However, with such an unbiased accounting method, the firm has no incentive to exert any effort to control information. The marginal benefit of increasing r is thus

$$\begin{aligned} B_r &= [\beta q_r - (1-\beta)p_r]\Delta S \\ &= \frac{1}{2}\left[\Phi\left((w-m)\underline{h}^{\frac{1}{2}}\right) + \Phi\left(w\underline{h}^{\frac{1}{2}}\right)\right]\Delta S - \frac{1}{2}\left[\Phi\left((w-m)\overline{h}^{\frac{1}{2}}\right) + \Phi\left(w\overline{h}^{\frac{1}{2}}\right)\right]\Delta S \end{aligned}$$

Note that

$$
\begin{aligned}
&\Phi\left((w-m)h^{\frac{1}{2}}\right)+\Phi\left(wh^{\frac{1}{2}}\right)\\
&=\Phi\left((w-\frac{m}{2}-\frac{m}{2})h^{\frac{1}{2}}\right)+\Phi\left((w-\frac{m}{2}+\frac{m}{2})h^{\frac{1}{2}}\right)\\
&=\Phi\left([\frac{m}{2}+(w-\frac{m}{2})]h^{\frac{1}{2}}\right)+\left[1-\Phi\left(-(w-\frac{m}{2}-\frac{m}{2})h^{\frac{1}{2}}\right)\right]\\
&=1+\Phi\left([\frac{m}{2}+(w-\frac{m}{2})]h^{\frac{1}{2}}\right)-\Phi\left([\frac{m}{2}-(w-\frac{m}{2})]h^{\frac{1}{2}}\right)\\
&=1+\int_{\frac{m}{2}-(w-\frac{m}{2})}^{\frac{m}{2}+(w-\frac{m}{2})}\phi\left(xh^{\frac{1}{2}}\right)dx
\end{aligned}
\tag{20}
$$

where the second equality is by the symmetry of the normal distribution function. This implies

$$
B_r=\int_{\frac{m}{2}-(w-\frac{m}{2})}^{\frac{m}{2}+(w-\frac{m}{2})}\phi\left(x\underline{h}^{\frac{1}{2}}\right)dx-\int_{\frac{m}{2}-(w-\frac{m}{2})}^{\frac{m}{2}+(w-\frac{m}{2})}\phi\left(x\overline{h}^{\frac{1}{2}}\right)dx
\tag{21}
$$

Therefore, when $w=\frac{m}{2}$, $\Phi\left((w-m)h^{\frac{1}{2}}\right)+\Phi\left(wh^{\frac{1}{2}}\right)=1$ and $B_r=0$. The firm's incentive to increase signal precision is low because, from its perspective, the increase in the probability of obtaining the favorable classification is *exactly* offset by the increase in probability of obtaining the unfavorable classification. So even though the user unambiguously benefits from more precise information, the firm has little incentive to incorporate such benefits in choosing the control effort r.

The situation changes if we apply a conservative accounting policy. By setting w equal to w^* plus a small positive deviation $\varepsilon>0$, such that $w>w^*$, the marginal benefit of increasing r is positive (i.e., $B_r>0$). Note that under this accounting regime, report $\{z=0\}$ would be issued when signal $w=\frac{m}{2}$ is received which indicates an equal likelihood of the true state being 0 or m. That is, the accounting reports the less favorable of the two equally likely outcomes. With such a conservative accounting system, the overall probability of issuing an unfavorable report is high. The firm, however, will have an incentive to increase signal precision since $B_r>0$. As a result the optimal accounting policy (w^{**}) is conservative.

It is worth pointing out the generalizability of the above argument from binary states to multiple states. As long as it is the case that, between any two states of affairs, the firm always prefers to be identified with the more favorable state, a more conservative classification rule will likely improve the firm's propensity to generate more accurate information. In Appendix B, we provide a scenario with three states of nature and three decision alternatives to illustrate this rationale.

4.3 Comparative Statics Analysis

Although conservatism, as a deviation from the optimal accounting scheme given perceived information quality, can increase the incentive to improve information quality, it is costly because classification is too pessimistic given the *ex post* information quality. A trade-off needs to be made between the *ex ante* incentive for improving information quality and the *ex post* accuracy given the information quality. The optimal threshold level w^{**} must therefore be chosen to strike a balance between these two forces. Next we examine how the optimal threshold level w^{**} changes with respect to the changes in the losses associated with Type I and Type II errors in our model.

PROPOSITION 5 *When $B_{rw} > 0$, ceteris paribus, the user's decision-making threshold w^{**} increases with L_2 and increases (decreases) with L_1 if $\frac{\Delta S}{k}$ is large(small).*

PROOF. See Appendix. ■

When the firm has no control over information quality, the optimal decision-making threshold w^* increases in L_2 and decreases in L_1. This is a natural consequence of the fact that information users seek to balance the opportunity losses associated with Type I and Type II decision-making errors. This result doesn't necessarily hold when the firm can control the information quality. Holding everything else constant, an increase in L_1 has two offsetting effects: On the one hand, a larger L_1 increases the loss associated with a Type I error (i.e., being overly cautious and misclassifiying the firm as being in the bad state when the state is actually good), and hence tends to decrease the decision-making threshold. On the other hand, because $\frac{\partial L}{\partial r}$ increases in the magnitudes of L_1 and L_2, as L_1 increases, the information user cares more about information quality and will want to commit to a higher decision-making threshold to induce more information quality control effort. Which effect dominates depends on the magnitude of $\frac{dr}{dw}$, that is, the sensitivity of the firm's information quality control effort to the choice of the decision-making threshold. Notice that the firm doesn't directly care about the information user's payoff: the first-order condition of the firm's information quality control effort reveals that $\frac{dr}{dw}$ is amplified by $\frac{\Delta S}{k}$, but is independent of L_1 or L_2. Hence, when $\frac{\Delta S}{k}$ is large enough, the information quality control effect dominates the direct effect of decision making efficiency, and the optimal decision-making threshold will increase with L_1. Notice that an increase in L_2 unambiguously increases the decision-making threshold because both of the above two effects — improving decision-making efficiency and inducing more information quality control effort — tend to increase L_1.

5 Conclusion

This study has shown that when a firm can control the quality of its reported financial information through noncontractible action, an accounting policy that aggregates raw information in a biased fashion can increase the quality of accounting information. Our analysis takes into account the endogenous nature of the information generation process. The accounting system, which functions as a classification system in the presence of uncertainty, serves not only as an information aggregation scheme ex post but also incites firms to improve their information quality ex ante. Comparative statics reveal that in equilibrium, the level of conservatism might even increase with the opportunity loss associated with being overly prudent (e.g., misclassifying a good firm as being in an unfavorable state of affairs).

In this paper, we restricted the purpose of public reporting policy to mandating the disclosure of high-quality information. Our conclusions may not hold if the objectives of standard-setting include other considerations such as enhancing corporate control or facilitating litigation. Our study also relies on the assumption that the average preference of investors can be captured by the behavior of a representative investor. We do not study the more general setting where different types of investors intend different uses of the same accounting information. Prior studies (e.g., Demski 1974) have shown that since alternative accounting standards often lead to different wealth distributions among individuals, a strict criterion of Pareto improvement does not yield much insight when comparing alternative accounting standards. Consistent with this view, we make no claim that our model offers a comprehensive explanation of accounting conservatism. Our purpose has been to highlight the potentially positive effect of conservative aggregation on the information quality control effort put forth by firm insiders, an effect which to the best of our knowledge has been overlooked in the literature. Our result should be of interest to accounting academics, practitioners, and standard setters who are concerned with factors affecting the accuracy of accounting information.

Appendix 1:

Preliminaries on Normal Density

Denote the density function of a standard normal variable (i.e., with mean 0 and variance 1) by $\phi(\cdot)$ and the associated cumulative distribution function by $\Phi(\cdot)$. Then for a normal variable $w \sim N(m, \frac{1}{h})$, denote its density function by $f(w) = h^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right)$ and its cumulative distribution function by $F(w) = \Phi\left((w-m)h^{\frac{1}{2}}\right)$. The derivative of f with respect to h is:

$$f_h(w) = \left(\frac{1}{2h} - \frac{1}{2}(w-m)^2\right) f(w) = \left(\frac{1}{2h} - \frac{1}{2}(w-m)^2\right) h^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right).$$

and

$$F_h(w) = \frac{1}{2}(w-m)h^{-\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right)$$

Proof of Proposition 1.

PROOF. Let $L(w,r)$ denote $(1-p(w,r))(1-\beta)L_2 + (1-q(w,r))\beta L_1$, the expression to be minimized in (8).

(i) This obviously holds since $\frac{\phi((y-m)h^{\frac{1}{2}})}{\phi(yh^{\frac{1}{2}})}$ increases with y and equals 1 at $y = \frac{m}{2}$.

(ii) Notice that $\frac{\phi(w^* h^{\frac{1}{2}})}{\phi((w^*-m)h^{\frac{1}{2}})} = \exp(-mh(w^* - \frac{m}{2}))$ increases in h if $w^* < \frac{m}{2}$ and decreases in h if $w^* > \frac{m}{2}$.

Further, from (i), $w^* \lesseqgtr \frac{m}{2}$ when $\beta L_1 \gtreqless (1-\beta)L_2$. Rearrange (11) and we get

$$\begin{aligned} &r\bar{h}^{\frac{1}{2}}\left[\phi\left(w^*\bar{h}^{\frac{1}{2}}\right)(1-\beta)L_2 - \phi\left((w^*-m)\bar{h}^{\frac{1}{2}}\right)\beta L_1\right] \\ &= (1-r)\,\underline{h}^{\frac{1}{2}}\left[\phi\left((w^*-m)\underline{h}^{\frac{1}{2}}\right)\beta L_1 - \phi\left(w^*\underline{h}^{\frac{1}{2}}\right)(1-\beta)L_2\right]. \end{aligned}$$

For the above equality to hold, it must be true that

$$sign\left[\frac{\phi\left(w^*\bar{h}^{\frac{1}{2}}\right)}{\phi\left((w^*-m)\bar{h}^{\frac{1}{2}}\right)} - \frac{\beta L_1}{(1-\beta)L_2}\right] = sign\left[\frac{\beta L_1}{(1-\beta)L_2} - \frac{\phi\left(w^*\underline{h}^{\frac{1}{2}}\right)}{\phi\left((w^*-m)\underline{h}^{\frac{1}{2}}\right)}\right],$$

so the following must also be true:

– when $\beta L_1 > (1-\beta)L_2$, $\frac{\phi\left(w^*\underline{h}^{\frac{1}{2}}\right)}{\phi\left((w^*-m)\underline{h}^{\frac{1}{2}}\right)} < \frac{\beta L_1}{(1-\beta)L_2} < \frac{\phi\left(w^*\bar{h}^{\frac{1}{2}}\right)}{\phi\left((w^*-m)\bar{h}^{\frac{1}{2}}\right)}$.

– when $\beta L_1 < (1-\beta)L_2$, $\frac{\phi\left(w^*\underline{h}^{\frac{1}{2}}\right)}{\phi\left((w^*-m)\underline{h}^{\frac{1}{2}}\right)} > \frac{\beta L_1}{(1-\beta)L_2} > \frac{\phi\left(w^*\bar{h}^{\frac{1}{2}}\right)}{\phi\left((w^*-m)\bar{h}^{\frac{1}{2}}\right)}$.

So

$$\begin{aligned} L_{rw} &= \left[\underline{h}^{\frac{1}{2}}\phi\left(w\underline{h}^{\frac{1}{2}}\right)(1-\beta)L_2 - \underline{h}^{\frac{1}{2}}\phi((w-m)\underline{h}^{\frac{1}{2}})\beta L_1\right] \\ &- \left[\bar{h}^{\frac{1}{2}}\phi\left(w\bar{h}^{\frac{1}{2}}\right)(1-\beta)L_2 - \bar{h}^{\frac{1}{2}}\phi\left((w-m)\bar{h}^{\frac{1}{2}}\right)\beta L_1\right] \\ &\Rightarrow L_{rw}\begin{cases} < 0 \text{ when } w^* < \frac{m}{2} \\ > 0 \text{ when } w^* > \frac{m}{2} \end{cases}. \end{aligned}$$

By the standard monotone comparative statics argument, given that L_w decreases with r when $w^* < \frac{m}{2}$ and increases with r when $w^* > \frac{m}{2}$, the unique optimal w^* increases with r when $w^* < \frac{m}{2}$ and decreases with r when $w^* > \frac{m}{2}$, that is, $\frac{\partial w^*}{\partial r} \lesseqgtr 0$ when $w^* \gtreqless \frac{m}{2}$, or to put it equivalently, when $\beta L_1 \neq (1-\beta)L_2$, $\left|w^* - \frac{m}{2}\right|$ decreases with r.

(iii) When $r = 0$, from (11), $w^* = \frac{m}{2} + \frac{1}{m\underline{h}} \ln \frac{\beta L_1}{(1-\beta)L_2}$. From (ii), we know $\left|w^* - \frac{m}{2}\right|$ decreases in r. Hence, $0 < \frac{m}{2} + \frac{1}{m\underline{h}} \ln \frac{\beta L_1}{(1-\beta)L_2} < m$ constitutes a sufficient condition for $w^* \in (0, m)$ for any $r \geq 0$. Straightforward mathematics reveals that this condition entails $\underline{h} > \frac{2}{m^2} \max \left\{ \ln \frac{\beta L_1}{(1-\beta)L_2}, \ln \frac{(1-\beta)L_2}{\beta L_1} \right\}$. ■

Proof of Corollary 1

PROOF. Let L_r denote the derivative of the expected loss L with respect to r, then

$$\begin{aligned} -L_r &= (1-\beta)L_2 p_r + \beta L_1 q_r \\ &= (1-\beta)L_2 \left[\Phi\left(w\overline{h}^{\frac{1}{2}}\right) - \Phi\left(w\underline{h}^{\frac{1}{2}}\right) \right] + \beta L_1 \left[\Phi\left((w-m)\underline{h}^{\frac{1}{2}}\right) - \Phi\left((w-m)\overline{h}^{\frac{1}{2}}\right) \right] \\ &> 0 \text{ for any } w \in (0, m). \end{aligned}$$

since for any $w \in (0, m)$, $\Phi\left(w\overline{h}^{\frac{1}{2}}\right) > \Phi\left(w\underline{h}^{\frac{1}{2}}\right)$ and $\Phi\left((w-m)\underline{h}^{\frac{1}{2}}\right) > \Phi\left((w-m)\overline{h}^{\frac{1}{2}}\right)$.

Hence, the expected loss L decreases with r for any $w \in (0, m)$. ■

Proof of Proposition 2

PROOF. With accounting policy $R^d : z = y$, signal y is fully revealed. Hence the representative end user will take the following action:

$$a = \begin{cases} a_1 \text{ if } y \geq w^* \\ a_2 \text{ if } y > w^* \end{cases}$$

where $w^*(r) \equiv$ argmin $[(1-\beta)(1-p)L_2 + \beta(1-q)L_1]$. In equilibrium, the firm chooses r^* such that $\Delta S\ [\beta q_r - (1-\beta)p_r] = kc_r$. Consider the situation where the user chooses w to efficiently motivate management to control information quality. Totally differentiating ΔS $[\beta q_r - (1-\beta)p_r] - kc_r = 0$ with respect to r and w yields

$$\frac{dr^*}{dw} = \frac{-(\beta q_{rw} - (1-\beta)p_{rw})\Delta S}{-kc_{rr}} = \frac{-B_{rw}\Delta S}{-kc_{rr}} > 0.$$

As a result

$$\frac{\partial L}{\partial w} = -(1-\beta)L_2 p_w - \beta L_1 q_w + \frac{dr}{dw}\frac{\partial[(1-\beta)(1-p)L_2 + \beta(1-q)L_1]}{\partial r}.$$

By Corollary 1,

$$\frac{\partial[(1-\beta)(1-p)L_2 + \beta(1-q)L_1]}{\partial r} = L_r \neq 0,$$

In addition, condition (16) implies

$$\frac{dr}{dw} = \frac{[\beta q_{rw} - (1-\beta)p_{rw}]\Delta S}{kc_{rr}} = \frac{B_{rw}\Delta S}{kc_{rr}} \neq 0.$$

Thus, at the point $w = w^*(r)$, $\frac{\partial L}{\partial w} \neq 0$. Therefore, in order to satisfy the first-order condition $\frac{\partial L}{\partial w} = 0$, we must have $w^{**} \neq w^*(r)$. That is, *ex ante* commitment to biased accounting reduces the overall expected loss. The accounting rule R^a, with only signal z disclosed, can achieve this *ex ante* optimum by setting the accounting rule to

$$z = \begin{cases} 0 \text{ when } y < w^{**} \\ m \text{ when } y > w^{**} \end{cases}$$

Hence,

$$L(w^a, r^a) < L(w^d, r^d)$$

■

Proof of Proposition 3.

PROOF. The derivative of the private marginal benefit $B_r = (\beta q_r - (1-\beta)p_r)\Delta S$ with respect to w is

$$\begin{aligned} B_{rw} &= [\beta q_{rw} - (1-\beta)p_{rw}]\Delta S \\ &= \left\{\beta\left[\underline{h}^{\frac{1}{2}}\phi\left((w-m)\underline{h}^{\frac{1}{2}}\right) - \overline{h}^{\frac{1}{2}}\phi\left((w-m)\overline{h}^{\frac{1}{2}}\right)\right] - (1-\beta)\left[\overline{h}^{\frac{1}{2}}\phi\left(w\overline{h}^{\frac{1}{2}}\right) - \underline{h}^{\frac{1}{2}}\phi\left(w\underline{h}^{\frac{1}{2}}\right)\right]\right\}\Delta S. \end{aligned}$$

Now for any given $\underline{h}$ and $w \in (0, m)$, we have

$$\lim_{\overline{h}\to\infty} B_{rw} = \left\{\beta\underline{h}^{\frac{1}{2}}\phi\left((w-m)\underline{h}^{\frac{1}{2}}\right) + (1-\beta)\,\underline{h}^{\frac{1}{2}}\phi\left(w\underline{h}^{\frac{1}{2}}\right)\right\}\Delta S > 0.$$

Hence, by continuity, there exist $\overline{h}$ and an interval $W \subset (0, m)$ such that for all $w \in W$, $B_{rw} > 0$.

Now we show that this interval W expands as $\overline{h}$ increases. From (A1), we can rewrite B_{rw} as

$$\begin{aligned} B_{rw} &= [\beta q_{rw} - (1-\beta)p_{rw}]\Delta S \\ &= \left\{\int_{\underline{h}}^{\overline{h}}\left[-\beta\left(\frac{1}{2h} - \frac{1}{2}(w-m)^2\right)h^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right) - (1-\beta)\left(\frac{1}{2h} - \frac{1}{2}w^2\right)h^{\frac{1}{2}}\phi(wh^{\frac{1}{2}})\right]dh\right\}\Delta S. \end{aligned}$$

For a given w, let $b(h; w)$ denote the integrand in the above expression, that is,

$$b(h; w) = -\beta\left(\frac{1}{2h} - \frac{1}{2}(w-m)^2\right)h^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right) - (1-\beta)\left(\frac{1}{2h} - \frac{1}{2}w^2\right)h^{\frac{1}{2}}\phi(wh^{\frac{1}{2}}).$$

Now the key to the proof is to show that $\frac{dB_{rw}}{d\bar{h}} = b(\bar{h}; w) > 0$

We prove this by contradiction. Assume the contrary, that is, $\frac{dB_{rw}}{d\bar{h}} = b(\bar{h}; w) \leq 0$ when $B_{rw} > 0$.

First, since $B_{rw} > 0$, by the mean value theorem, there exists an $\widetilde{h} \in \left(\underline{h}, \overline{h}\right)$ such that $b(\widetilde{h}; w) > 0$. Since $b(h; w)$ is infinitely differentiable in h, given that $b(\widetilde{h}; w) > 0$ and $\widetilde{h} < \overline{h}$, there must exist an $\widetilde{\widetilde{h}} \in [\widetilde{h}; \overline{h}]$ such that $b(\widetilde{\widetilde{h}}; w) = 0$ and $\frac{db(\widetilde{\widetilde{h}};w)}{d\widetilde{\widetilde{h}}} \leq 0$

The derivative of $b\,(h;w)$ with respect to h is

$$\begin{aligned}
\frac{db(h;w)}{dh} &= \frac{h^{-2}}{2}[\beta h^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right) + (1-\beta)h^{\frac{1}{2}}\phi\left(wh^{\frac{1}{2}}\right)] \\
&- \beta\left(\frac{1}{2h} - \frac{1}{2}(w-m)^2\right)^2 h^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right) - (1-\beta)\left(\frac{1}{2h} - \frac{1}{2}w^2\right)^2 h^{\frac{1}{2}}\phi(wh^{\frac{1}{2}}) \\
&= \frac{1}{2}(w-m)^2[2-h(w-m)^2]\beta h^{-\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right) + \frac{1}{2}w^2[2-hw^2](1-\beta)h^{-\frac{1}{2}}\phi\left(wh^{\frac{1}{2}}\right) \\
&= \underbrace{\frac{1}{2}\beta(w-m)^2 h^{-\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right) + \frac{1}{2}(1-\beta)w^2h^{-\frac{1}{2}}\phi\left(wh^{\frac{1}{2}}\right)}_{a_1(h;w)} \\
&+ \underbrace{\frac{1}{2h}\left[-\beta(w-m)^2(1-h(w-m)^2)h^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right) - (1-\beta)w^2(1-hw^2)h^{\frac{1}{2}}\phi\left(wh^{\frac{1}{2}}\right)\right]}_{a_2(h;w)}.
\end{aligned}$$

Denote the first term of the above expression by $a_1\,(h;w)$ and the second term by $a_2\,(h;w)$. It's obvious that $a_1\,(h;w) > 0\ \forall\ h$ and w.

For $a_2\,(h;w)$, notice that at the point where $b(h;w) = 0$, that is,

$$\begin{aligned}
b(h;w) &= -\beta\left(\frac{1}{2h} - \frac{1}{2}(w-m)^2\right)h^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right) - (1-\beta)\,h^{\frac{1}{2}}\left(\frac{1}{2h} - \frac{1}{2}w^2\right)\phi(wh^{\frac{1}{2}}) \\
&= \frac{1}{2h}\left[-\beta(1-h\,(w-m)^2)h^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right) - (1-\beta)\left(1-hw^2\right)h^{\frac{1}{2}}\phi(wh^{\frac{1}{2}})\right] \\
&= 0 \qquad (22)
\end{aligned}$$

so we must have

- if $w > \frac{m}{2}$, $(w-m)^2 < w^2$, for $b(h;w) = 0$ to be satisfied, it must be true that $-\beta\,(1-h(w-m)^2) < 0 < -\,(1-\beta)\,(1-hw^2)$.
- if $w < \frac{m}{2}$, $(w-m)^2 > w^2$, for $b(h;w) = 0$ to be satisfied, it must be true that $-\beta\,(1-h(w-m)^2) > 0 > -\,(1-\beta)\,(1-hw^2)$.

Thus, at $b(h;w) = 0$, $a_2\,(h;w)$ must be positive because relative to (A2) it always puts less weight on the negative part and more weight on the positive part.

Therefore, $\frac{db(h;w)}{dh}$ must be positive at $b(h;w) = 0$. This contradicts with the existence of $\widetilde{\widetilde{h}}$ $\in [\widetilde{h};\overline{h}]$ where $b(\widetilde{\widetilde{h}};w) = 0$ and $\frac{db(\widetilde{\widetilde{h}};w)}{d\widetilde{\widetilde{h}}} \leq 0$.

Hence, when $B_{rw} \geq 0$, it must be true that

$$\begin{aligned}\frac{dB_{rw}}{d\overline{h}} &= b(\overline{h};w)\Delta S \\ &= \left\{-\beta\left(\frac{1}{2\overline{h}}-\frac{1}{2}(w-m)^2\right)\overline{h}^{\frac{1}{2}}\phi\left((w-m)h^{\frac{1}{2}}\right)-(1-\beta)\left(\frac{1}{2\overline{h}}-\frac{1}{2}w^2\right)\overline{h}^{\frac{1}{2}}\phi(wh^{\frac{1}{2}})\right\}\Delta S \\ &> 0.\end{aligned}$$

Therefore, an increasing $\overline{h}$ expands the interval $W \subset (0,m)$ over which $B_{rw} \geq 0$. ■

Proof of Proposition 4

PROOF. Note that

$$\frac{\partial L}{\partial w} = -(1-\beta)L_2p_w - \beta L_1q_w + \frac{dr}{dw}\frac{\partial[(1-\beta)(1-p)L_2 + \beta(1-q)L_1]}{\partial r}.$$

By Corollary 1,

$$\frac{\partial[(1-\beta)(1-p)L_2 + \beta(1-q)L_1]}{\partial r} = L_r < 0,$$

and by Corollary 2,

$$\frac{dr}{dw} = \frac{[\beta q_{rw} - (1-\beta)p_{rw}]\Delta S}{kc_{rr}} = \frac{B_{rw}\Delta S}{kc_{rr}} > 0.$$

Thus, at the point $w = w^*(r^*) \equiv \text{argmin}\ [(1-\beta)(1-p)L_2 + \beta(1-q)L_1]$,

$$sign\left(\frac{\partial L}{\partial w}\right) = sign\left(\frac{dr}{dw}\frac{\partial[(1-\beta)(1-p)L_2 + \beta(1-q)L_1]}{\partial r}\right) < 0.$$

Therefore, L is decreasing in w at the *ex post* optimal w^*. In order to satisfy the first order condition $\frac{\partial L}{\partial w} = 0$, we must have $-(1-\beta)L_2p_w - \beta L_1q_w > 0$ which means that $w^{**} > w^*(r^*)$, that is, *ex ante* commitment to conservatism is optimal. ■

Proof of Proposition 5

PROOF. Notice that the firm's information quality control effort choice is determined by (14) which is independent of L_1 and L_2. The derivative of the first-order condition of the user's threshold choice with respect to L_2 evaluated at the original w^{**}, the original effort level r^*, and the new L_2 is:

$$\begin{aligned}&\frac{\partial L}{\partial w^{**}\partial L_2}\\&= -p_w\beta + \frac{dr^*}{dw^{**}}(-\beta p_r)\\&< 0,\end{aligned}$$

since $p_r > 0$, $p_w > 0$, and $\frac{\partial r^*}{\partial w^{**}} > 0$. Therefore, the loss function decreases with L_2 at the original w^{**}, implying that the new threshold is higher than the original w^{**}.

The derivative of the first-order condition of the user's threshold choice with respect to L_1 evaluated at the original w^{**}, the original effort level r^*, and the new L_1 is

$$\begin{aligned}&\frac{\partial L}{\partial w^{**}\partial L_1}\\&= -q_w\beta + \frac{dr^*}{dw^{**}}(-\beta q_r).\end{aligned}$$

Because $-\frac{\bar{h}^{\frac{1}{2}}}{\sqrt{2\pi}} < q_w < 0$ and $0 < q_r < 1$, the above expression can be either positive or negative depending on the magnitude of $\frac{\partial r^*}{\partial w^{**}}$. Notice that $\frac{\partial r^*}{\partial w^{**}} = \frac{B_{rw}\Delta S}{kc_{rr}}$ increases in $\frac{\Delta S}{k}$. When $\frac{\Delta S}{k} = 0$, $\frac{\partial L}{\partial w^{**}\partial L_1} = -q_w\beta > 0$, implying that the loss function increases with L_1 at the original w^{**} so that the new threshold is lower than the original w^{**}. When $\frac{\Delta S}{k}$ is large enough, the above expression is negative. Hence, by continuity, we find that the optimal decision w^{**} increases (decreases) with L_1 when $\frac{\Delta S}{k}$ is large (small). ■

Appendix 2:

Consider a decision making scenario with three states of nature and three action alternatives. We assume that both the information user and the information originator have the same prior belief that all three states are equally likely. The information user's payoffs (opportunity losses) for each state-action combination are shown in the table below:

Payoff Table

	States of Nature		
Actions	State 1: $x = 0$ ($prob = 0.33$)	State 2: $x = m$ ($prob = 0.33$)	State 3: $x = 2m$ ($prob = 0.33$)
a_1	U	$U - L$	$U - 2L$
a_2	$U - L$	U	$U - L$
a_3	$U - 2L$	$U - L$	U

Notice that action a_i is the correct action to take when the actual state is i. Given state i, the opportunity loss, the difference between the actual payoff (for a chosen action) and the best potential payoff increases the further away the chosen action is from the optimal action. For example, the actions a_2 and a_3 can be thought of as capturing the degree of action beyond a simple yes/no decision (e.g., the "buy" and the "strong buy" recommendations issued by a securities analyst).

The information originator generates a signal $y = x + \epsilon$, where $\epsilon \sim N(0, h^{-1})$.

An ex post optimal decision then involves two threshold levels $\{w_1, w_2\}$ with $w_1 < w_2$ for the realized signal y: the information user would take action a_1 when $y < w_1$, action a_2 when $w_1 < y < w_2$, and action a_3 when $y > w_2$.

Given w_1 and w_2, let p_{ij}, with $i, j \in \{1, 2, 3\}$ denote the probabilities of action a_j being taken when the underlying state is i.

Then at w_1, it has to be true that the information user is indifferent between taking action

a_1 and a_2:

$$\sum_{i=1}^{3} p_{i1} \cdot |i-1| \cdot L = \sum_{i=1}^{3} p_{i2} \cdot |i-2| \cdot L$$

Given the monotone likelihood property of the normal distribution, it is straightforward to verify that at such w_1, the information user strictly prefers a_1 over a_3.

Similarly, at w_2, the information user is indifferent to taking action a_2 or a_3 (and strictly prefers a_2 over a_1):

$$\sum_{i=1}^{3} p_{i2} \cdot |i-2| \cdot L = \sum_{i=1}^{3} p_{i3} \cdot |i-3| \cdot L$$

It can be easily verified that *regardless of the signal precision*, the ex post optimal decision rules are $w_1 = \frac{m}{2}$ and $w_2 = \frac{3m}{2}$. Notice that at these two thresholds

$p_{11} = p_{33}$, $p_{21} = p_{23}$, and $p_{31} = p_{13}$, and the two equations above are satisfied.

The information originator is risk neutral and has linear utility over the actions of the information user. He obtains $S_i = \alpha + \beta \cdot i$ when action a_i is taken. Let S denote his expected utility given the common prior and the decision thresholds. Then

$$\begin{aligned}
S &= \frac{1}{3} \sum_{j=1}^{3} \sum_{i=1}^{3} p_{ij} \cdot S_j \\
&= \frac{1}{3} \left[(1 + p_{31}) \cdot (\alpha + \beta) + (1 - p_{13} - p_{31}) \cdot (\alpha + 2\beta) + (1 + p_{13}) \cdot (\alpha + 3\beta) \right] \\
&= \alpha + 2 \cdot \beta + \frac{\beta}{3} \cdot (p_{13} - p_{31})
\end{aligned}$$

Because at the ex post optimal w_1 and w_2, $p_{13} = p_{31}$. The information originator's expected utility S remains constant at $\alpha + 2\beta$ regardless of information precision, and he has no incentive to improve the information quality.

Relative to the ex post optimal thresholds $\{w_1, w_2\}$, an accounting system $\left\{w_1^{'}, w_2^{'}\right\}$ is conservative if $w_i^{'} \geq w_i$ for $i \in \{1, 2\}$. The intuition that a little conservatism can tilt the information originator's preference toward more accurate information is analogous to that of a scenario with only two states of nature. An increase in w_1 increases p_{31} and an increase in w_2 reduces p_{13}, both resulting in lower expected utility S. However, the decrease in S

would be less extreme if the signal were more accurate. For example, presume that h can take two values: with probability r, $h = \underline{h}$, where $0 < \underline{h} < +\infty$, and with probability $(1 - r)$, $h = \overline{h} = +\infty$. Then the information originator would have a strictly positive incentive to increase r, because when the information is more precise, his expected utility is less affected by the thresholds $\{w_1, w_2\}$.

It is also interesting to note that with three states of nature, aggregation could play a more significant role in inducing information quality control effort than a simple commitment mechanism. If we aggregate state 2 and state 3 by setting $w_1^{'} = \frac{m}{2}$ and $w_2^{'} = +\infty$, then $p_{13} \equiv 0$ regardless of h. Since p_{31} decreases in h, the information originator would again have a positive incentive to improve information quality. Such aggregation will likely incur additional costs due to information loss (e.g., Ronen 1971, Butterworth 1972, Ijiri 1975, Feltham 1975). The benefit of aggregation due to the increase in the accounting information quality will need to be weighed against those added costs.

References

[1] Bushman, R. and R. Indjejikian. 1993. Stewardship Value of "Distorted" Accounting Disclosures. *The Accounting Review.* October 1993: 765-782.

[2] Butterworth, J. 1972. The Accounting System as an Information Function. *Journal of Accounting Research* 10: 1-27.

[3] Chen, Q., T. Hemmer, and Y. Zhang. 2007. On the Relation between Conservatism in Accounting Standards and Incentives for Earnings Management *Journal of Accounting Research.*45: 541-565.

[4] Christensen, J. and J. Demski. 2004. Asymmetric Monitoring: Good Versus Bad News Verification. *Schmalenbach Business Review* (July).

[5] Demski, J. 1974. Choice Among Financial Reporting Alternatives. *The Accounting Review* (April): 221-232.

[6] Devine, C. 1963. The Rule of Conservatism Reexamined. *Journal of Accounting Research* 1 (Autumn): 127-38.

[7] Dutta, S. and X. Zhang. 2002. Revenue Recognition in a Multiperiod Agency Setting. *Journal of Accounting Research* 40 (March): 67-83

[8] Dye, R. and S. Sridhar. 2004. Reliability-Relevance Trade-Offs and the Efficiency of Aggregation. *Journal of Accounting Research* 42: 51-88.

[9] Feltham, G. 1975. Cost Aggregation: An Information Economic Analysis. *Journal of Accounting Research* 15 (Spring): 42-70.

[10] Financial Accounting Standards Board. 1975. Statement of Concepts No. 2: *Qualitative Characteristics of Accounting Information.* Stamford, CT: FASB.

[11] Gigler, F. and T. Hemmer. 2001. Conservatism, Optimal Disclosure Policy, and the Timeliness of Financial Reports. *The Accounting Review* (October): 471-493

[12] Gigler, F., C. Kanodia, H. Sapra and R. Venugopalan. 2009. Accounting Conservatism and the Efficiency of Debt Contracts. *Journal of Accounting Research 47 (3)*: 767-797

[13] Hendriksen, E. S. and M. F. Van Breda. 1992. *Accounting Theory.* Irwin, 5^{th} ed., Homewood. IL.

[14] Ijiri, Y. 1975. *Theory of Accounting Measurement.* American Accounting Association. Sarasota, FL.

[15] Kwon, Y.K., Newman, P. and Y. S. Suh. 2001. The Demand for Accounting Conservatism for Management Control. *Review of Accounting Studies* 6 (March): 29-51

[16] Li, H. 2001. A Theory of Conservatism. *Journal of Political Economy*, vol. 109(3): 617-636.

[17] Lin, H. 2006. Accounting Discretion and Managerial Conservatism: An Intertemporal Analysis. *Contemporary Accounting Research* 23 (4): 1017-41.

[18] Ronen, J. 1971. Some Effects of Sequential Aggregation in Accounting on Decision-Making. *Journal of Accounting Research* 9 (Autumn): 307-332.

[19] Sterling, R. R. 1970. *Theory of the Measurement of Enterprise Income.* University of Kansas Press. Lawrence, Kansas.

[20] Watts, R. 2003. Conservatism in Accounting. Working paper. University of Rochester.