Accounting Conservatism and Debt Contract Efficiency with Soft Information

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Xu Jiang

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Chandra Kanodia

August 2011

Acknowledgements

This work would have been impossible without the guidance and support of my advisor, Chandra Kanodia. I am deeply indebted to Chandra for his invaluable suggestion and encouragement throughout the project. Chandra’s commitment to the scholarly development of his students is admirable and inspiring.

I have also benefited greatly from Frank Gigler for his many extremely helpful comments and insights.

I would also like to acknowledge the support and helpful comments that I have received from the members of my dissertation committee: Raj Singh and Aldo Rustichini.

Finally, I gratefully acknowledge the patience and emotional support received from my parents, Guoqing Xu and Xiaochun Jiang as well as my wife, Mingzhu Li.

Abstract

This paper shows how accounting conservatism affects the efficiency of debt contracting when the optimal debt contract allows parties to renegotiate. In my model allowing for renegotiation is optimal because both borrowers and lenders have access to non-contractible "soft information", in addition to the more objective accounting information. I show that conservative accounting can increase the efficiency of debt contracting when the other soft information is also sufficiently conservative (to be defined in the paper). However, when the soft information is aggressive, conservative accounting is detrimental to the efficiency of debt contracting. Thus, whether more conservative financial reporting is good for debt contracting depends on the interaction between the informational characteristics of the accounting system and other available information. Interestingly, the result that the informational characteristics of the accounting system should be the same as those of the soft information stands in contrast to the conjecture proposed by many studies -- that accounting needs to be conservative because other information, such as that offered by management, is opportunistically aggressive.

Table of Contents

List of Tables

List of Figures

1 Introduction

Existing research offers mixed opinions on whether accounting conservatism improves debt contracting. On one hand, Watts (2003) and a large body of empirical work (e.g., Basu (1997), Ball and Shivakumar (2005), Ball, Robin and Sadka (2008), and Zhang (2008)) argue that because creditors gain nothing on the upside but lose from the downside, they benefit from a timely shift in the decision rights during technical default. And since an accounting system which anticipates future losses but not future gains facilitates the timely shift in decision rights, conservative accounting improves the efficiency of debt contracts.

On the other hand, theoretical models such as Gigler, Kanodia, Sapra and Venugopalan (2009) (hereafter GKSV), find that more conservative accounting deteriorates the efficiency of debt contracting. The first key to their finding is that accounting conservatism is characterized as affecting the information content as well as the asymmetrical timeliness of accounting information. Specifically, because more conservative accounting increases the frequency of reporting low signals relative to high signals, the information content of low signals is diminished while that of high signals is increased. The second key is their observation that for any positive expected net present value (hereafter referred to as NPV) project the expected cost of wrongly continuing a bad project is less than the cost of discontinuing a good project. Since it is therefore efficient to continue with positive NPV projects in the absence of information, an information system is most valuable when the signals that result in discontinuing the projects are more informative than signals which suggest continuing the project, i.e., when bad news is more informative than good news.

In this paper I construct a theoretical model in which accounting conservatism may or may not improve the efficiency of debt contracting. The main departure from the GKSV model which allows me to derive this result is broadening the information environment to include non-contractible "soft information" in addition to the accounting information that they model. When creditors and borrowers jointly observe this soft information in conjunction with the accounting signal, it may be in their mutual best interests to renegotiate the existing debt contract, which assigns decision rights based only on the contractible accounting information. In fact it has been empirically documented that it is common to renegotiate debt contracts following technical default rather than foreclose and the "accrual of new information" is often the trigger for renegotiation (e.g., Roberts and Sufi (2009)). For example, consider a firm that has low earnings that triggers a violation of the debt covenant. However, at the same time suppose there is an independent third party report (e.g., financial analyst report) saying that the firm has gained considerable market share. Depending on the informativeness of the report, the debtholders might let the firm continue its projects despite violation of the covenant in the renewed hope that the firm will generate enough cash flow to repay the loan.

While the soft information may be relevant to the decision of whether or not to renegotiate a debt contract, its characteristics are unaffected by the degree of conservatism in the accounting information. Nevertheless, since any renegotiation would base the project termination/continuation decision on *both* accounting and soft information, the correlation of the accounting and soft information will determine the efficiency of the optimal debt contract. The correlation will in turn be affected by the degree of conservatism in the accounting information.

The main finding in this paper is that the effect of accounting conservatism on debt contract efficiency depends on the properties of the soft information. When the soft information is sufficiently conservative, meaning that favorable soft information is relatively more informative than unfavorable, conservative accounting improves debt contract efficiency. The opposite is true when soft information is aggressive, i.e., when favorable soft information is less informative than unfavorable. Consequently the optimal degree of accounting conservatism acts as a complement, rather than a substitute, to the bias in other available information. This result stands in stark contrast to the commonly held belief that conservative accounting can be optimal in "undoing" the tendency of managers to present a favorable picture of their firms (see for example LaFond and Watts (2008), page 452).[1]

To understand the intuition underlying this counter-intuitive result, think about matching two information systems to minimize the cost of making decision errors. For instance, suppose a decision on whether to liquidate or to continue a project is being made based only on the soft information. If the information is sufficiently reliable it will be optimal to liquidate if the information is unfavorable (a low signal) and continue if it is favorable (a high signal). Therefore, if the soft information is coming from an aggressive information system, meaning that high signals are less informative than low signals, decision errors are more likely when the signal is high. Now consider augmenting the soft information with accounting information. The accounting information is most useful when it leads to the largest reduction in decision errors. Since most decision errors without the accounting information will occur when the soft information is favorable, the accounting information is most valuable when it provides strong countervailing evidence to favorable soft information, i.e., when low accounting signals are most informative. As a result, the accounting information is most useful when it too is aggressive. Notice that if we were to try to "undo" the aggressive bias in the soft information by supplementing it with conservative accounting information, low accounting signals are not very informative and therefore will not be as effective in changing the decision when it is least informed.[2]

This paper makes contributions to the accounting literature on several dimensions. First, it shows that it is not the characteristics of the bias in accounting information per se which determines the efficiency of decision making, rather it is in the interaction of the bias in accounting and the bias in other available

[1] It is worth mentioning that on page 450, Lafond and Watts (2008) specifically stated that their conjecture is "not tested in the paper", which implies that whether empirical data support their conjecture or the findings of my paper is still an open question.

[2] The same reasoning applies if we start with conservative soft information; then the most valuable accounting information is also conservative.

information. Second, I show that complementary biases are actually beneficial, overturning the folk wisdom that conservative accounting is useful in undoing optimism in management representations. Finally, this paper develops a new methodology, motivated by Milgrom (1981), which emphasizes the likelihood ratio as the key statistic in analyzing information systems.

The remainder of the paper is organized as follows. Section 2 provides an overview of the related literature. The model is presented and developed in Section 3. Benchmark cases to the more general analysis are presented in Section 4 and the main analysis in Section 5. Section 6 discusses some empirical implications of my model and section 7 concludes. Some of the more complicated proofs are relegated to the appendix.

2 Related Literature

There are several theoretical studies in addition to Watts (2003) which economically rationalize the pervasiveness of accounting conservatism. Kwon et al. (2001) shows that conservative accounting can serve as a motivating device in a principal agent framework, while Chen et al. (2007) shows that conservatism can be good because it increases the cost of inefficient earnings management. Guay and Verrecchia (2006) argue that conservative accounting leads to more information being disclosed because it acts as a legal regime which prevents managers from upwardly biasing their information. Li (2009) adds asymmetric information and explicit debt contract renegotiation to the setting in GKSV and shows that conservative accounting can be valuable when renegotiation is costly by decreasing the incidence of renegotiation. In Caskey and Hughes (2010) conservative accounting helps to mitigate an asset substitution problem.

Other theoretical papers have not been as supportive of conservative accounting. In Gigler and Hemmer (2001) conservative accounting is undesirable because it reduces the value of voluntary disclosure. They show how the effect of accounting conservatism on the earnings-returns relationship documented by Basu (1997) is consistent with their model and therefore should not be interpreted as evidence that accounting conservatism is valuable. The GKSV paper as discussed in the introduction also provides empirical predictions consistent with the finding of Basu (1997) even though conservative accounting is detrimental to debt contracting in their model.

This paper, like GKSV and Li (2009), also focuses specifically on the effect of accounting conservatism on debt contracting. Like Li (2009) I include the possibility of renegotiation but based on public non-contractible, rather than private, information. In addition I use a more general notion of conservatism, adapted from GKSV, which allows for a continuous variation in the signal space as well as in the optimal debt contract.[3]

3 The Model

Consider a firm at date 0 that needs an amount K to finance a project that will return an uncertain cash flow $\tilde{x}$ two periods later (i.e., at date 2). The firm finances the investment by issuing a bond with a face value of D and an interest rate of R to be paid after the cash flow is realized at date 2.[4] A "debt covenant", which specifies the control rights of the project conditional on any contractible information accompanies the bond.[5] As in GKSV, I assume risk-neutrality for both parties, as risk-aversion would not qualitatively affect the results. However, unlike in GKSV, I don't make any assumptions about the NPV of the project, i.e., it is not necessary that the expected return $\mathrm{E}(\tilde{x})$, exceed the liquidation value M. At an intermediate period (i.e., date 1), two pieces of information are

[3] The results in Li (2009) are in part due to the fact that the information structure prohibits the debt contract from optimally adjusting to the degree of accounting conservatism.

[4] In this paper, as in GKSV we assume that the whole investment is financed by debt, i.e., capital structure is not the focus here. There is one justification for debt to be optimal, however. Since in my model there is this additional soft information, if we assume that the additional information is observable only to current shareholders and bondholders but unknown to future shareholders (e.g., relationship banks might have some information advantage over prospective investors), then a pecking-order argument similar to that in Myers and Majluf (1984) would suggest that the optimal security is debt as by assumption there is no internally generated funds. That being said, how conservatism affects capital structure is an interesting issue and is a good future research topic.

[5] I refer to the bond in conjunction with the covenant as "debt contract".

observed by both the bondholders and shareholders of the firm. One is the same as in GKSV, i.e., the contractible accounting information denoted as $\tilde{y}$ which is correlated with future cash flow but is also affected by whether the accounting system is conservative or aggressive. The other piece, the so-called "soft" information denoted as $\tilde{s}$, is also correlated with future cash flow but is not affected by the degree of accounting conservatism. I assume this "soft" information is non-contractible, which means that it cannot be contracted on initially at date zero. However, at date 1, the two parties can renegotiate any contract based on what they have observed of the "soft" information. At date zero, when writing the initial contract, they anticipate that the renegotiation will happen at date 1. After seeing the information and renegotiating at date 1, the project will either be terminated or liquidated depending upon the original contract (in particular, the control rights specified by the original contract) and the renegotiation process. The time line of the model is shown in figure 1.

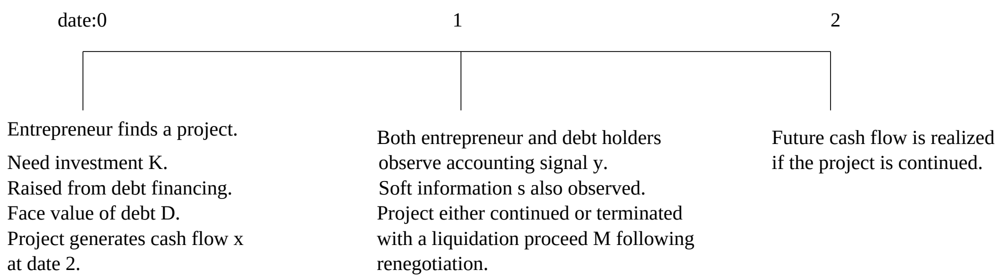

Figure 1 Timeline of the Model

I assume that x has support $[0,\infty]$, y has support $[0,\bar{y}]$ and the soft information s has support $[0,\bar{s}]$. The conditional probability density function of y given x and s is denoted by $g(y|s,x)$ and it is affected by accounting conservatism which will be discussed in more detail in the next section.The conditional probability density function of s given x is denoted by $l(s|x)$ and by assumption is not affected by accounting conservatism. Similar to GKSV (2009), I assume that high values of y and s are "good news" so that higher values of either y or s would shift the distribution of x to the right in the spirit of Milgrom (1981).[6] Since we now have two pieces of information y and s to infer values of x, I assume that for any s, there exists a unique $y \in [0, \bar{y}]$ which I denote as $y^*(s)$ such that $E(\tilde{x}|y^*(s), s) = M$ and also that for any y, there exists a unique $s \in [0, \bar{s}]$ which I denote as $s^*(y)$, which is the inverse function of $y^*(s)$ such that $E(\tilde{x}|s^*(y), y) = M$.[7] Of course, the socially efficient solution is to terminate any project for which the conditional expectation of x on y and s is smaller than M and to continue the project otherwise.

Before I go to the formal derivation of the optimal debt contract, let me briefly summarize its properties. The results show that adding non-contractible information does not change the economic implications of the optimal debt contract shown in GKSV. The optimal debt contract still leads to the socially efficient solution outlined above. The key economic insight here is that, because both pieces of information are observed by both parties, they can always renegotiate back to the socially efficient decision. And because renegotiation is costless in my model,[8] the Coase Theorem (Coase (1960)) applies and the ex post efficient decision will be made through renegotiation.[9] Its proof is quite long and tedious and is included in the appendix.

Proposition 1. *The optimal debt contract at date zero contains an earnings based covenant and the threshold value of the covenant can be any number y∈[0,$\bar{y}$]. But date 1 renegotiation results in the socially efficient solution; for any (y,s) that generates $E(\tilde{x}|y, s) > M$ the project will be continued while for any (y,s) that generates $E(\tilde{x}|y, s) < M$ the project will be terminated. Moreover, the principal amount of the bond, D, depends in part on the distribution of the bargaining*

[6] More assumptions regarding conditional distributions will be discussed later.

[7] Because of the assumption that higher values of either y or s would shift the distribution of x to the right, the uniqueness of those values is assured. The fact that $y^*(s)$ is invertible will be proved later.

[8] Thus, the setting of my model is more like a private loan with one single lender with very detailed knowledge of the firm so that the lender both has access to other "soft" information and renegotation comes at little cost. How adding renegotiation cost will affect the results of my model is an interesting future research topic.

[9] Li (2009) also derives a similar result in her setting.

powers in renegotiation and it is not always true that $D > M$.[10]

Proof. See Appendix.

□

Three remarks are worth making regarding the proposition. First, unlike in GKSV in which they explicitly derived that $D > M$, in my model because of renegotiation and the corresponding possibility of bondholders recouping all cash flows from the project in certain circumstances,[11] bondholders may accept a principal amount smaller than the liquidation value M.

Second, it is the mutual observability of s as well as the costless renegotiation that is driving the result. This result is independent of the distribution of the bargaining power between the two parties.

Thirdly, one may ask "if the threshold value of the covenant can be any number, why do we need a covenant in the first place?". The answer is that the covenant has allocational consequences, set by parties in negotiating the original debt agreement. If we assume, as in the proof of proposition 1, that whoever has the decision rights also has full bargaining power, then the tightness of covenant is directly related to the face value of debt, D. However, since this cut-off point only has allocational consequences with no effect on the efficiency of the debt contract, the exact value of the covenant has no effect on any of the main results of this paper.

[10]The above derivation is based on the assumption that $D > M$. But the same socially efficient debt contract can be derived for the case where $D < M$ and the proof is available from the author upon request. Also because D is in part determined by the distribution of bargaining power between bondholders and shareholders one cannot explicitly solve for D without assuming the distribution of bargaining power. Fortunately, the exact value of D and the associated interest rate R is not the interest of this paper. We will focus on how conservatism affects the social efficiency of the debt contract.

[11]A recent real life example is the bondholders from General Growth Properties, who actually get more than 100 cents per dollar of their orginal debt when General Growth Properties emerge from bankruptcy.

4 Accounting Conservatism And Debt Contract Efficiency: A Benchmark Case

Having derived the optimal debt contract, I next explore how accounting conservatism affects the efficiency of debt contracts. The first issue is to establish a measure of efficiency. GKSV develop an economically intuitive way of measuring the efficiency of debt contracts, namely, the sum of two types of errors which they call "undue optimism" and "false alarm". The "undue optimism" error is the expected loss from continuing the project when eventually the project generates a cash flow lower than liquidation value (i,e, it should have been terminated), while the "false alarm" error is the expected loss from terminating the project when the project would have generated a cash flow higher than liquidation value if it were continued (i.e., it should have been continued). Because in my model, the optimal debt contract also results in the efficient outcome, it can be shown in a similar manner to GKSV that the optimal contract here also results in the minimization of the sum of the same two types of errors.[12] To save space, the proof is omitted and available from the author upon request. The main point is that, the efficiency measure proposed in GKSV can also be used here. Mathematically this efficiency measure can be written as

$$\Omega \equiv \int_M^\infty \int_0^{\bar{y}} \int_0^{s^*(y)} (x-M) * f(x|y,s) * h(y,s) dyds dx$$
$$+ \int_0^M \int_0^{\bar{y}} \int_{s^*(y)}^{\bar{s}} (M-x) * f(x|y,s) * h(y,s) dyds dx \qquad (1)$$

Recall from Proposition 1 that renegotiation results in the project continuation choice being socially efficient, independent of the original covenant cut-off value y^c at date zero. Thus the sum of the two types of errors is also independent of the covenant and is only affected by the degree of accounting conservatism.

As a benchmark I first explore the case when the soft information, s, is a sufficient statistic for the pair (y, s). The sufficient statistic notion essentially says

[12] For any accounting system, the optimal contract minimizes the sum of two errors. But this summation itself is a function of how conservative the accounting system would be, which is the main focus of this paper.

that the distribution of future cash flows x conditional on s and y is the same as that conditional on s, or mathematically speaking, $f(x|y,s) = f(x|s)\ \forall y, s$. This can also be thought of as s being a signal with such high information content that s itself is sufficient to determine the distributions of future cash flows and y doesn't add anything beyond s.[13]

Intuitively one would think that because accounting information y does not have any incremental information content relative to s, accounting conservatism would not affect the efficiency of debt contract. Proposition 2 shows that this intuition is correct.

Proposition 2. *When s is a sufficient statistic for the pair (y,s), the efficiency of the debt contract is independent of the conservatism properties of the accounting system.*

Proof. When s is a sufficient statistic of the pair (y,s), we would have $f(x|y,s) = f(x|s)g(y)$. And now we have $s^*(y) = s^*$ as by definition and the property of sufficient statistics we have $E(\tilde{x}|s^*(y), y) = \int_0^\infty xf(x|y, s^*(y))dx = \int_0^\infty xf(x|s^*(y))dx \equiv M\ \forall y$ and taking derivatives of both sides with respect to y we have $\frac{\partial s^*(y)}{\partial y} = 0$ and thus we denote $s^*(y) = s^*$. Taking this notation into Equation (4.1) above and we would have

$$\Omega = \int_M^\infty \int_0^{\bar{y}} \int_0^{s^*} (x - M)f(x|s)g(y)h(y,s)dydsdx$$

$$+ \int_0^M \int_0^{\bar{y}} \int_{s^*(y)}^{\bar{s}} (M - x)f(x|s)g(y)h(y,s)dydsdx$$

$$= \int_M^\infty \int_0^{s^*} (x - M)f(x|s)ds \int_0^{\bar{y}} g(y)h(y,s)dydx$$

[13] This benchmark case is also related to the point made by some researchers who argue that accounting earnings is too noisy and lags behind other information such as analyst forecast or pro-forma earnings which might be less noisy and more timely in valuing a firm's future performance. For example, Bhattacharya et al. (2003) documents that market participants believe that firms' pro forma earnings are more representative of "core earnings" than GAAP operating income.

$+ \int_0^M \int_{s^*(y)}^{\bar{s}} (M - x) f(x|s) ds \int_0^{\bar{y}} g(y) h(y, s) dy dx$

$= \int_M^\infty \int_0^{s^*} (x - M) f(x|s) l(s) ds dx$

$+ \int_0^M \int_{s^*(y)}^{\bar{s}} (M - x) f(x|s) l(s) ds dx$

where $l(s) = \int_0^{\bar{y}} h(y, s) * g(y) dy$.

Thus, since Ω is independent of y and accounting conservatism only affects y we have the summation of errors independent of accounting conservatism.

□

I have established a benchmark case which shows that conservatism doesn't matter when s is a sufficient statistic of the pair (y, s). The GKSV case can be viewed as another benchmark which shows that conservatism diminishes debt contract efficiency when y is a sufficient statistic of the pair (y, s).[14] The more interesting case is when neither s nor y is a sufficient statistic for the other and as we will see in the next section, the relative information content of s at different values plays an important role in the effect of conservatism on debt contract efficiency.

5 Accounting Conservatism And Debt Contract Efficiency: The General Case

Now we come to the case where the correlation between accounting information y and the soft information s is more general. From now on, I consider the case where x is a binary signal, i.e., $x \in \{x_H, x_L\}$ $(x_H > M > x_L)$ because it simplifies the algebra without changing any results qualitatively. Also in this

[14] In GKSV, y is the only available information, which is equivalent to y being a sufficient statistic of (y, s) and the optimal contract only referring to y.

section, I will use $y^*(s)$ to denote the dividing line between terminating and continuing the project where $E(\tilde{x}|y^*(s), s) = M$. For any s, the project is terminated whenever $y < y^*(s)$ and continued otherwise.

Conservatism is assumed to change the information content of y in the manner of GKSV. Lower signals have higher information content in aggressive accounting systems while higher signals have higher information content in conservative accounting systems. Denote the level of conservatism in the accounting system δ, with higher value of δ representing more aggressive accounting. Since y^* is not only a function of the signal s, but is also a function of how conservative the system is, we write it as $y^*(s, \delta)$, with $E(\tilde{x}|y^*(s, \delta), s) = M$.

Introducing an additional piece of information requires my making the following extension to the conservatism notion in GKSV.

Assumption 1 : $\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}$ *strictly increases in* y $\forall\delta, s$.

Assumption 2 : $\frac{l(s|x_H)}{l(s|x_L)}$ *strictly increases in* s.

Assumption 3 : $\frac{\Psi(y,s|x_H,\delta)}{\Psi(y,s|x_L,\delta)} \equiv \frac{g(y|s,x_H,\delta)*l(s|x_H)}{g(y|s,x_L,\delta)*l(s|x_L)}$ *strictly increases in* s $\forall y, \delta$.

Assumption 4 : $\int_0^y g(t|s, x_i, \delta)dt$ *strictly decreases in* δ $\forall y, s, x_i, i = H, L$.

Assumption 5 : $\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}$ *strictly decreases in* δ $\forall y, s$.

Assumptions 1 and 2 are the usual monotone likelihood ratio conditions for y and s respectively. They state that, regardless of the accounting system, high values of y and high values of s are always “good news” in the sense of Milgrom (1981), i.e., higher signals increase the likelihood of future cash flow being x_H rather than x_L. Assumption 3 says that regardless of the accounting system, and for any y, higher s increases the likelihood of future cash flow being x_H rather than x_L. Assumption 4 merely states that a decrease in accounting conservatism always shifts the conditional distribution of y to the right in the sense of first order stochastic dominance. Assumption 5 is the key assumption regarding conservatism both in GKSV and in my case. It implies that a decrease in accounting conservatism not only shifts the distribution of y to the right, but also changes the information content of accounting conservatism. The same signal (y, s) that is observed in a conservative accounting system would indicate a higher likelihood of future cash flow being x_H over x_L compared with such signal generated in an aggressive accounting system. Put alternatively, $E(\tilde{x}|y, s, \delta)$ is decreasing in δ.

The following important lemma will be used repeatedly in the analysis:

Lemma 3. *$\forall(y, s)$ such that $E(\tilde{x}|y, s) \geq (\leq)M$, we have $\int_M^\infty (x - M) * \psi(y, s|x) * \varphi(x)dx \geq (\leq) \int_0^M (M - x) * \psi(y, s|x) * \varphi(x)dx$, where $\varphi(x)$ is the unconditional probability density function of x and $\psi(y, s|x)$ is the joint conditional distribution of y and s given x. When x is binary, we have $P_H(x_H - M) * \psi(y, s|x_H) \geq (\leq)P_L(M - x_L) * \psi(y, s|x_L)$.*

Proof. When $E(\tilde{x}|y, s) \geq (\leq)M$, we have $(\int_0^M + \int_M^\infty)x * f(x|y, s)dx \geq (\leq)M$, and thus we have $\int_M^\infty (x - M) * f(x|y, s)dx \geq (\leq) \int_0^M (M - x) * f(x|y, s)dx$. Using Bayes rule we have $f(x|y, s) = \frac{\psi(y,s|x)*\varphi(x)}{h(y,s)}$, and since $h(y, s)$ gets canceled out in both side, we get the expression when x is continuous. When x is binary and because $x_H > M > x_L$,we have

$\int_M^\infty (x - M) * \psi(y, s|x) * \varphi(x)dx = P_H(x_H - M)\psi(y, s|x_H)$

and

$\int_0^M (M - x) * \psi(y, s|x) * \varphi(x)dx = P_L(M - x_L) * \psi(y, s|x_L)$. And the expression thus follows. □

Basically, the lemma states that the optimal debt contract must be chosen to minimize the sum of the two types of errors. If a certain (y, s) pair causes the project to be continued, it must be that the "undue optimism" error caused by continuation is not larger than the "false alarm" error caused by termination.

Otherwise renegotiation would cause the termination of the project. Similarly, when the project is continued upon observing a certain (y, s) pair, it must be the case that the "false alarm" error is smaller than the "undue optimism".

To gain some intuition, in section 5.1 I start from the case where both y and s are binary signals and there are only two accounting systems; a conservative system and an aggressive system. In section 5.2 I introduce a new methodology which is crucial in proving the main results in the continuous setting. I present the main results in section 5.3.

5.1 When y and s are binary

In this case $x \in \{x_H, x_L\}$, $y \in \{y_h, y_l\}$ and $s \in \{s_h, s_l\}$. Subscript h represents more favorable news relative to subscript l in the sense of Milgrom (1981). Also, there are two possible accounting systems; conservative and aggressive. I denote the conservative accounting system by the superscript and/or subscript C and the aggressive accounting system by the superscript and/or subscript A. Note the purpose of this section is solely to gain some intuition in developing the assumptions used in deriving the main result of the paper. Therefore it tends to be more explanatory, at the expense of sacrificing rigor.

The efficiency difference between conservative and aggressive accounting systems must stem from differences in the continuation decision across the two settings. I assume that there is enough total information under both accounting systems such that the project is continued when observing (y_h, s_h) and terminated when (y_l, s_l) is observed. Therefore the difference in decisions only occurs when either (y_h, s_l) or (y_l, s_h) is observed. I further assume that any realization of s is more informative than the least informative y signals but less informative than the most informative y signals. As a result, the project would be continued with (y_h, s_l) in a conservative accounting system but terminated in an aggressive accounting system. Recall that this is because y_h is highly informative when generated from a conservative accounting system but not very informative when generated from an aggressive accounting system. Similar logic would imply that the project would be continued with (y_l, s_h) in a conservative accounting system but terminated in an aggressive accounting system.[15] Therefore, the difference in the sum of errors between the conservative and aggressive accounting system is:

$\Omega_C - \Omega_A = -[P_H(x_H - M) * P^C(y_l, s_h|x_H) - P_L(M - x_L) * P^C(y_l, s_h|x_L)] - [P_H(x_H - M) * P^A(y_h, s_l|x_H) - P_L(M - x_L) * P^A(y_l, s_h|x_L)]$

From Lemma 3 we know that $P_H(x_H - M) * P^C(y_l, s_h|x_H) > P_L(M - x_L) * P^C(y_l, s_h|x_L)$ and $P_H(x_H - M) * P^A(y_h, s_l|x_H) < P_L(M - x_L) * P^A(y_l, s_h|x_L)$. Thus, additional information s reduces the decision error of a conservative accounting system by continuing at (y_l, s_h) and reduces the decision error of an aggressive accounting system by terminating at (y_h, s_l). Comparing the magnitude of the reductions determines which accounting system generates the least amount of error when there is additional information s.

The expression can be rewritten as

$\Omega_C - \Omega_A = -p_H(x_H - M) * [P^C(y_l, s_h|x_H) + P^A(y_h, s_l|x_H)] + P_L(M - x_L) * [P^C(y_l, s_h|x_L) + P^A(y_h, s_l|x_L)]$.

Thus, $\Omega_C > (<)\Omega_A$ if and only if $\frac{P^C(y_l,s_h|x_H)+P^A(y_h,s_l|x_H)}{P^C(y_l,s_h|x_L)+P^A(y_h,s_l|x_L)} < (>) \frac{P_L(M-x_L)}{P_H(x_H-M)} \equiv K$.

Note from Lemma 3 $\frac{P^C(y_l,s_h|x_H)}{P^C(y_l,s_h|x_L)} > K$ and $\frac{P^A(y_h,s_l|x_H)}{P^A(y_h,s_l|x_L)} < K$. Since $\frac{P^A(y_h,s_l|x_H)}{P^A(y_h,s_l|x_L)} < \frac{P^C(y_l,s_h|x_H)+P^A(y_h,s_l|x_H)}{P^C(y_l,s_h|x_L)+P^A(y_h,s_l|x_L)} < \frac{P^C(y_l,s_h|x_H)}{P^C(y_l,s_h|x_L)}$, $\frac{P^C(y_l,s_h|x_H)+P^A(y_h,s_l|x_H)}{P^C(y_l,s_h|x_L)+P^A(y_h,s_l|x_L)}$ can be viewed as a weighted average of $\frac{P^A(y_h,s_l|x_H)}{P^A(y_h,s_l|x_L)}$ and $\frac{P^C(y_l,s_h|x_H)}{P^C(y_l,s_h|x_L)}$. Thus whether or not $\frac{P^C(y_l,s_h|x_H)+P^A(y_h,s_l|x_H)}{P^C(y_l,s_h|x_L)+P^A(y_h,s_l|x_L)}$ is greater or less than K depends on the distance of $\frac{P^A(y_h,s_l|x_H)}{P^A(y_h,s_l|x_L)}$ and $\frac{P^C(y_l,s_h|x_H)}{P^C(y_l,s_h|x_L)}$ from K, respectively. From here I divide the discussion into two cases.

[15]Because of the assumptions made above, there are certain restrictions on how the decision rule would behave under conservative relative to aggressive accounting systems. It can be shown that the rule discussed in the paper satisfies those constraints and would offer the most intuitive way to illustrate the trade-off in the binary setting. Details can be obtained from the author upon request.

Case 1: Favorable s is more informative than unfavorable s.

When s_h is more informative than s_l, there isn't much difference between the likelihood ratio $\frac{P^C(y_l,s_h|x_H)}{P^C(y_l,s_h|x_L)}$ and $\frac{P(s_h|x_H)}{P(s_h|x_L)}$ because s_h is highly informative while y_l is not very informative for a conservative accounting system. Since s_h strongly suggests that the future project will be a success, $\frac{P(s_h|x_H)}{P(s_h|x_L)}$ must be much larger than K. Therefore, $\frac{P^C(y_l,s_h|x_H)}{P^C(y_l,s_h|x_L)}$ must likewise be much larger than K. However, s_l is not as informative as s_h, so $\frac{P(s_l|x_H)}{P(s_l|x_L)}$ is not much smaller than K. And since y_h is not very informative with aggressive accounting, $\frac{P^A(y_h|x_H)}{P^A(y_h|x_L)}$ is not much larger than K and therefore $\frac{P^A(y_h,s_l|x_H)}{P^A(y_h,s_l|x_L)}$ is not much smaller than K. The weighted average, $\frac{P^C(y_l,s_h|x_H)+P^A(y_h,s_l|x_H)}{P^C(y_l,s_h|x_L)+P^A(y_h,s_l|x_L)}$ is larger than K. Therefore, the conjecture is that when favorable s is more informative than unfavorable s, $\Omega_C < \Omega_A$ and a conservative accounting system is more efficient.

Case 2: Favorable s is less informative than unfavorable s.

When s_l is more informative than s_h, there isn't much difference between the likelihood ratio $\frac{P^A(y_h,s_l|x_H)}{P^A(y_h,s_l|x_L)}$ and $\frac{P(s_l|x_H)}{P(s_l|x_L)}$ because s_l is highly informative while y_h is not very informative for an aggressive accounting system. Since s_l strongly suggests that the future project will be a failure, $\frac{P(s_l|x_H)}{P(s_l|x_L)}$ must be much smaller than K. Therefore, $\frac{P^A(y_h,s_l|x_H)}{P^A(y_h,s_l|x_L)}$ must likewise be much smaller than K. However, s_h is not as informative as s_l, so $\frac{P(s_h|x_H)}{P(s_h|x_L)}$ is not much larger than K. And since y_l is not very informative with conservative accounting, $\frac{P^C(y_l|x_H)}{P^C(y_l|x_L)}$ is not much smaller than K as well, making $\frac{P^C(y_l,s_h|x_H)}{P^C(y_l,s_h|x_L)}$ not much larger than K. The weighted average, $\frac{P^C(y_l,s_h|x_H)+P^A(y_h,s_l|x_H)}{P^C(y_l,s_h|x_L)+P^A(y_h,s_l|x_L)}$ is therefore smaller than K, meaning that when favorable s is less informative than unfavorable s, $\Omega_C > \Omega_A$ and a conservative accounting system is less efficient.

Table 1 provides a summary of the conjectures for the two binary cases discussed above.[16] The general intuition gained from the binary case is that the

[16]Notice that, unlike in GKSV, no assumptions regarding the NPV of the project is imposed

Table 1: Summary of the conjectured efficiency results for binary cases

Cases	Efficiency Results
Favorable s is more informative than unfavorable s	$\Omega_C \leq \Omega_A$
Favorable s is less informative than unfavorable s	$\Omega_C \geq \Omega_A$

relative informativeness of favorable s versus unfavorable s plays an important role in whether conservatism increases or decreases the efficiency of an optimal debt contract. Since the statistical measure of informativeness is the likelihood ratio, it will be more intuitive to express the sum of errors in terms of likelihood ratios, which is the motivation for the new methodology introduced in the next section.

5.2 A new methodology

From the analysis of the optimal debt contract and Lemma 3, we know that the decision of whether or not to continue is based solely on the likelihood ratio $\frac{\psi(y,s|x_H)}{\psi(y,s|x_L)}$. The project is continued whenever $\frac{\psi(y,s|x_H)}{\psi(y,s|x_L)} > K$ and terminated whenever $\frac{\psi(y,s|x_H)}{\psi(y,s|x_L)} < K$.[17] In addition, any two (y, s) signals are "comparable" as in Milgrom (1981) because (y_1, s_1) is more favorable than (y_2, s_2) if the likelihood ratio of the former is greater than the latter. Therefore, it will be more intuitive to express the sum of errors as function of the likelihood ratios and explore the efficiency properties in terms of the properties of likelihood ratio.

First I introduce some notation and definitions.

Define $s^*(y, t, \delta)$ as the value that satisfies $\frac{\psi(y,s^*(y,t,\delta)|x_H,\delta)}{\psi(y,s^*(y,t,\delta)|x_L,\delta)} = t$ if there exists $s^*(y, t, \delta) \in [0, \bar{s}]$ that satisfies the relation.[18]

If $\frac{\psi(y,0|x_H,\delta)}{\psi(y,0|x_L,\delta)} > t$ then define $s^*(y, t, \delta) = 0$. If $\frac{\psi(y,\bar{s}|x_H,\delta)}{\psi(y,\bar{s}|x_L,\delta)} < t$ then define $s^*(y, t, \delta) = \bar{s}$.[19]

in the discussion.

[17]Recall that we have defined $K \equiv \frac{P_L(M-x_L)}{P_H(x_H-M)}$.

[18]Monotone likelihood ratio properties imply there is at most one solution.

[19]This construction preserves the continuity of $s^*(y, t, \delta)$ when there is no $s \in [0, \bar{s}]$ satisfying $\frac{\psi(y,s^*(y,t,\delta)|x_H,\delta)}{\psi(y,s^*(y,t,\delta)|x_L,\delta)} = t$.

Figure 2 A sketch for a typical $s^*(y,t,\delta)$

Figure 2 provides a sketch for a typical $s^*(y,t,\delta)$ as a function of t. The positive slope in the interior region is derived from the assumptions and is proved below as a lemma.

Lemma 4. *When* $\frac{\psi(y,s^*(y,t,\delta)|x_H,\delta)}{\psi(y,s^*(y,t,\delta)|x_L,\delta)} = t$ *for some* $s^*(y,t,\delta) \in [0,\bar{s}]$*, then we have* $\frac{\partial s^*(y,t,\delta)}{\partial t} = 1/\{\frac{\partial}{\partial s}[\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)}]|_{s=s^*(y,t,\delta)}\} > 0.$

Proof. Take the derivative of the expression $\frac{\psi(y,s^*(y,t,\delta)|x_H,\delta)}{\psi(y,s^*(y,t,\delta)|x_L,\delta)} = t$ with respect to t. Using the chain rule and rearranging terms yields the result. The sign follows from assumption 2.

□

Now if $\frac{\psi(y,0|x_H,\delta)}{\psi(y,0|x_L,\delta)} > t$ or $\frac{\psi(y,\bar{s}|x_H,\delta)}{\psi(y,\bar{s}|x_L,\delta)} < t$ then define $\frac{\partial s^*(y,t,\delta)}{\partial t} = 0$. Then for any y, $\frac{\partial s^*(y,t,\delta)}{\partial t}$ is almost everywhere continuous, which makes it Riemann integrable. This property will be exploited below.

Now I derive the expression for the distribution of t conditional on x_i for $i = H, L$, which we denote as $\Gamma(t|x_i,\delta)$. This is done by taking the derivative of the conditional cumulative distribution function of t on x_i, $\Phi(a \leq t|x_i,\delta)$. We have

$\Phi(a \leq t|x_i,\delta) = \int_0^{\bar{y}} \int_0^{s^*(y,t,\delta)} \psi(y,s|x_i,\delta)dsdy$

using the monotonicity of the likelihood ratio property with respect to s as in assumption 2.

The derivative with respect to a is

$\Gamma(t|x_i,\delta) = \frac{\partial}{\partial t}\Phi(a \leq t|x_i,\delta) = \frac{\partial}{\partial t}(\int_0^{\bar{y}} \int_0^{s^*(y,t,\delta)} \psi(y,s|x_i,\delta)dsdy)$

$= \int_0^{\bar{y}} \psi(y,s^*(y,t,\delta)|x_i,\delta) * \frac{\partial}{\partial t}s^*(y,t,\delta)dy$

There will be discontinuity points for $\frac{\partial}{\partial t}s^*(y,t,\delta)$. But since $\frac{\partial}{\partial t}s^*(y,t,\delta)$ is continuous almost everywhere, it is Riemann integrable and the expression is valid.

Denote $\underline{t}(\delta) = \frac{\psi(0,0|x_H,\delta)}{\psi(0,0|x_L,\delta)}$ and $\bar{t}(\delta) = \frac{\psi(\bar{y},\bar{s}|x_H,\delta)}{\psi(\bar{y},\bar{s}|x_L,\delta)}$ we write the sum of errors as[20]

$\Omega(\delta) = P_H(x_H - M) * P(t(\delta) \leq K|x_H) + P_L(M - x_L) * (1 - P(t(\delta) \leq K|x_L))$

$= P_L(M - x_L) + P_H(x_H - M) * \int_{\underline{t}(\delta)}^{K} \Gamma(t|x_H,\delta)dt$

$-P_L(M - x_L) * \int_{\underline{t}(\delta)}^{K} \Gamma(t|x_L,\delta)dt.$

Inserting the expression for $\Gamma(t|x_i,\delta)$ into the expression for $\Omega(\delta)$ gives

$\Omega(\delta) = P_L(M - x_L) + P_H(x_H - M) * \int_{\underline{t}(\delta)}^{K} \int_0^{\bar{y}} \psi(y,s^*(y,t,\delta)|x_i,\delta) * \frac{\partial}{\partial t}s^*(y,t,\delta)dydt$

$-P_L(M - x_L) * \int_{\underline{t}(\delta)}^{K} \int_0^{\bar{y}} \psi(y,s^*(y,t,\delta)|x_L,\delta) * \frac{\partial}{\partial t}s^*(y,t,\delta)dydt \qquad (2)$

The next section will provide my main result by analyzing how $\Omega(\delta)$ changes with δ.

5.3 When y and s are continuous

Based on the intuition from the binary examples, I construct a mathematical notion of favorable s being more informative than unfavorable s (which I denote as s being conservative) and the opposite case (s being aggressive).

$Assumption\ 6a: s\ is\ conservative\ if\ \frac{\partial^2}{\partial s \partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}] > 0\ \forall y,s,\delta.$

[20]When $t = K$,we would have both parties indifferent between liquidating and continuing the project. In the analysis that follows, we shall assume that in this case the project is liquidated, but assuming otherwise would not change the result.

$$Assumption\, 6b: s\, is\, aggressive\, if\ \frac{\partial^2}{\partial s \partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}] < 0\, \forall y, s, \delta.$$

These two important assumptions deserve explanation. First, recall from assumption 5 that a decrease in accounting conservatism (i.e., an increase in δ) decreases the likelihood ratio for any (y, s) combination. Assumptions 6a & 6b specify how this decrease in the likelihood ratio varies with s. Assumption 6a states that, when the decrease in the likelihood ratio caused by an increase in conservatism is less for high s values than for low s values, s is conservative. This conforms to the GKSV notion of conservatism in that the bias in accounting has less effect on the informativeness of the total available information when the other (non-accounting) information is by itself most informative, i.e., when s is favorable. Assumption 6b is just the opposite. When the decrease in the likelihood ratio caused by the increase in conservatism is less for low s values than for high s values, s is aggressive because low s values are relatively more informative than high s.

Figures 3a and 3b give a graphical illustration of the likelihood ratio as a function of s for fixed y and two accounting systems; when s is conservative and when s is aggressive. Note that the difference between the two lines represents a shift in the likelihood ratio caused by changing from aggressive to conservative accounting. When s is conservative, high s are very informative while low s is not very informative, thus the difference shrinks when s becomes larger. The opposite result holds when s is aggressive.[21]

[21] The graphs are just for illustration purpose and the fact that the slopes are positive are not always true. However, the main point in those graphs is that more informative s is going to shrink the change in informativeness caused by a change in δ.

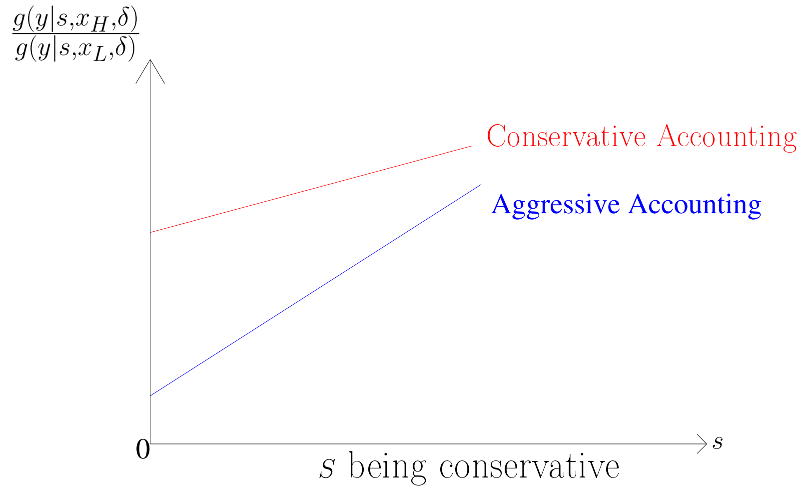

Figure 3a An illustration of s being conservative

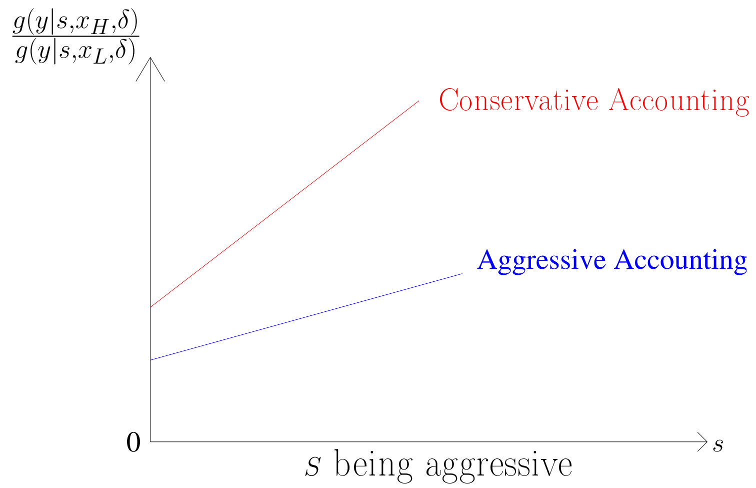

Figure 3b An illustration of s being aggressive

Equipped with these new assumptions, I present the main result of the paper.

Proposition 5. *Suppose assumptions 1,2,3 and 5 hold. Then, if s is aggressive (i.e., assumption 6b holds), then increasing accounting conservatism decreases the efficiency of debt contract. When s is sufficiently conservative (i.e., assumption 6a holds), increasing accounting conservatism increases the efficiency of debt contract .*

Proof. See Appendix. □

The proposition formalizes the intuition obtained in the binary examples and extends it to the more general case.[22] As has been discussed in the binary case, whether more conservative accounting is good or bad depends on whether the other information, s, is conservative or aggressive. In fact it can be seen directly from the proof of proposition 5 that the derivative of the efficiency measure with respect to the conservatism measure depends solely on the sign of $\frac{\partial^2}{\partial s \partial \delta}[\frac{\Psi(y,s|x_H,\delta)}{\Psi(y,s|x_L,\delta)}]$, the main ingredient in determining whether s is conservative or aggressive.

When s is conservative, favorable s is more informative than unfavorable s, implying that a not-so-informative unfavorable y doesn't matter much when s is favorable because the likelihood ratio, and therefore the decision rule, would not vary much with y. However, when s is unfavorable it is not very informative; a lot of decision error comes from liquidating when seeing unfavorable s. This means that a highly informative signal which tells the decision maker when to continue will greatly improve efficiency. This corresponds to a conservative accounting system, because (high) earnings which indicate continuation are most informative under conservative accounting.

When s is aggressive, similar logic would imply that improvements to efficiency would be greater for cases where the decision maker would have continued with less informative favorable s but a highly informative unfavorable y changes the decision. This corresponds to the aggressive accounting system being most efficient.

The only thing missing from this intuitive explanation of the result is the asymmetry that s has to be sufficiently conservative for conservative accounting to enhance the efficiency but s only need to be aggressive for aggressive accounting system to be more efficient. My conjecture is that there is an intrinsic deficiency of conservative accounting over aggressive accounting when there is only accounting information. Only when s is sufficiently conservative can this deficiency be over-

[22]From the proof we can also see the advantages of using the new methodology (i.e., use the likelihood ratio as the integration variable); only assumptions regarding the changes in likelihood ratio with respect to change in key variables need to be imposed. We don't need either the assumption on the NPV of the project or the assumption regarding the cumulative distribution function of conditional distributions of y as imposed in GKSV.

come for conservative accounting to be efficiency-enhancing. GKSV documented such deficiency but they assume positive NPV. Since my result is independent of the NPV, I am further conjecturing that the result of GKSV can be generalized without their NPV assumption, which is beyond the topic of this paper and is left for future research.

6 Empirical Implications

My main result suggests several empirical implications. First, it offers an explanation for why we observe variations in accounting conservatism across firms, industries and countries: specifically, cross-sectional differences in the relative informativeness of the other information that is observed by both parties lead to different levels of accounting conservatism being optimal.

Second, although this paper models s as non-contractible, soft information, the entire analysis goes through even when s is contractible. Thus, I would predict that the more conservative the firm's accounting system, the more conservative the other non-accounting information that is included in the debt contract. Take dividend constraints for example. Since maintaining or increasing dividends is less informative than decreasing dividends, the information conveyed in a firm's dividend choice is aggressive.[23] Therefore, a potential testable implication of my theory is that firms that have dividend constraints in their debt covenants will in general have less conservative accounting. Another example of other information not affected by accounting conservatism could be cash flow (or earnings) forecasts provided by analysts and/or managers. My model would imply that more negatively skewed forecasts are associated with more conservative accounting, not because analysts or managers cannot see through the more conservative accounting, but rather because conservative accounting is chosen as an optimal response

[23]it is well known that dividend is sticky (e.g., Myers (1984), page 581). Numerous papers (for example, Healy and Palepu (1988)) document that dividend omissions (one specific case of dividend decrease) would lead to greater market reaction than dividend initiations (one specific case of dividend increase), even after controlling for the magnitude of changes.

to the conservative bias in forecasts.

Third, my model is consistent with empirical studies that show accounting becomes more conservative after the Sarbanes-Oxley Act (SOX) (Lobo and Zhou (2006)). The reason is that, if the non-accounting information environment becomes more "conservative" in the post SOX period because managers and analysts tend to be more cautious in the face of increased legal liabilities and more hostile legal environment, my theory predicts that a more conservative accounting system is efficiency-enhancing.

7 Conclusion

This paper adds to our theoretical understanding of the relationship between accounting conservatism and debt contract efficiency. GKSV considered a debt contracting model with accounting information and showed that conservative accounting decreases the efficiency of debt contracts. I consider the case when there is not only accounting information but also other information that is helpful in predicting future cash flows. I show that conservatism can increase the efficiency of debt contracts when the other information is also sufficiently "conservative" and that aggressive accounting is optimal when the other information is also aggressive. The result stands in contrast with the common notion that accounting has to be conservative because other "soft" information sources are too optimistic to be trusted fully (LaFond and Watts (2008)). In establishing the main result, my paper also introduces a new methodology which directly exploits the properties of likelihood ratios that is potentially applicable to more general settings. Specifically, the methodology may prove beneficial in extending my analysis to settings with asymmetric information and/or endogenous capital structure. The model can also be extended to the case where the bias in the "soft" information is endogeneously determined. These extensions might provide further insights into the effect of accounting conservatism on the efficiency of debt contract.

References

[1] Ball, R., A. Robin and G. Sadka, 2008, "Is Financial Reporting Shaped by Equity Markets or by Debt Markets? An International Study of Timeliness and Conservatism" *Review of Accounting Studies* 13, 168-205.

[1] Ball, R. and L. Shivakumar, 2005, "Earnings Quality in U.K. Private Firms" *Journal of Accounting and Economics* 39, 83-128.

[2] Basu, S., 1997, "The Conservatism Principle and Asymmetric Timeliness of Earnings" *Journal of Accounting and Economics* 24,3-37.

[3] Bhattacharya, N., E.L. Black, T.E. Christensen and C. R. Larson 2003, "Assessing the relative informativeness and permanence of pro forma earnings and GAAP operating earnings" *Journal of Accounting and Economics* 36, 285-319.

[4] Caskey, J. and J. Hughes, 2010, "Assessing the Impact of Alternative Fair Value Measures on the Efficiency of Project Selection and Continuation" *Working Paper, University of California at Los Angeles.*

[5] 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.

[6] Coase, R. H., 1960, "The Problem of Social Cost" *Journal of Law and Economics* 3, 1-44.

[7] Gigler, F. and T. Hemmer, 2001, "Conservatism, Optimal Disclosure Policy and the Timeliness of Financial Reports" *Journal of Accounting Research* 76, 471-493.

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

[9] Guay, W. and R. Verrecchia, 2006, "Discussion of an Economic Framework for Conservative Accounting and Bushman and Piotroski (2006)" *Journal of Accounting and Economics* 42, 149-165.

[10] Healy, P., and K. Palepu., 1988, "Earnings information conveyed by dividend initiations and omissions" *Journal of Financial Economics* 21, 149-175.

[11] Hölmström B., 1979, "Moral Hazard and Observability" *Bell Journal of Economics* 10, 74-91.

[12] Kwon, Y., P. Newman and Y. Suh, 2001, "The Demand for Accounting Conservatism for Management Control" *Review of Accounting Studies* 6, 29-51.

[13] LaFond, R. and R. L. Watts, 2008, "The Informational Role of Conservatism" *Accounting Review* 83, 447-478.

[14] Li, J., 2009, "Accounting Conservatism and Debt Contracts: Efficient Liquidation and Covenant Renegotiation" *Working Paper, Carnegie Mellon University* .

[15] Lobo, G.J. and J. Zhou, 2006, "Did Conservatism in Financial Reporting Increase after the Sarbanes-Oxley Act? Initial Evidence" *Accounting Horizons* 20, 57-73.

[16] Milgrom, P., 1981, "Good News and Bad News: Representation Theorems and Applications" *Bell Journal of Economics* 12, 380-391.

[17] Myers, S.J., 1984, " The Capital Structure Puzzle" *Journal of Finance* 39, 575-592.

[18] Myers, S.J. and N.S. Majluf, 1984, " Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have" *Journal of Financial Economics* 13, 187-221.

[19] Roberts, M. and A. Sufi, 2009, "Renegotiation of Financial Contracts: Evidence from Private Credit Agreements" *Journal of Financial Economics* 93, 159-184.

[20] Venugopalan, R., 2004, "Conservative in Accounting: Good or Bad?" *Working Paper, University of Illinois at Urbana-Champaign.*

[21] Watts, R., 2003, "Conservatism in Accounting Part I: Explanations and Implications" *Accounting Horizons* 17, 207-221.

[22] Zhang, J., 2008, "The Contracting Benefits of Accounting Conservatism to Lenders and Borrowers" *Journal of Accounting and Economics* 45, 27-54.

8 Appendix

Proof of Proposition 1:

Proof. Starting from date 1 after both parties observe y and s, I derive the date zero contract taking date 1 renegotiation into account. Also without loss of generality, let me assume that whenever decision right is transferred to one party, that party has the full bargaining power. This assumption is for simplicity but as has been discussed in the paper, it turns out that assuming any distribution of bargaining power between the parties would not affect the optimal debt contract. I also assume that initially the decision right is with the shareholders but it will later be shown as well that who holds the decision right first doesn't make any difference for the optimal debt contract.

Denote y^c as the covenant that both parties agreed on date zero. At date 1, after both parties observe y and s, the decision right will be shifted to the bondholders whenever at date 1 the observed signal y is smaller than y^c while stockholders retain the decision rights otherwise.[24] Also assume that for the moment $D > M$. This assumption will not affect the final result.[25] Further assume that if one party is indifferent between different decisions, the party would choose the decision that the other party prefers the most.

Consider the following four possible cases:

Case 1: when $y < y^c$ and $s < s^*(y)$.

In this case the bondholders hold the decision right and since $E(\tilde{x}|y, s)$is increasing in both y and s, $E(\tilde{x}|y, s) < M$. Thus bondholders get all the liquidation value M and shareholders get zero.

Case 2: when $y > y^c$ and $s < s^*(y)$.

[24] The situation when y is at the covenant value is not important because first, this is a point of zero measure and second, you can assume either party has the decision right and still get to the socially efficient solution.

[25] If we assume $D < M$ we can proceed exactly in the same manner and get exactly the same result. The proof is available from author upon request.

In this case the shareholders have the decision rights and their expected payoff without renegotiation if the project continues is $U(D,y,s) = \int_D^\infty (x - D)f(x|y,s)dx$. Thus the shareholders would always want the project to continue without renegotiation since their expected payoff in this case is always positive. However, in this case we have $E(\tilde{x}|y,s) < M$. And, since it's straightforward to show that $U(D,y,s) < E(\tilde{x}|y,s)$, $U(D,y,s) < M$. In this case, the bondholders, without renegotiation, would get $V(D,y,s) \equiv \int_0^D xf(x|y,s)dx + \int_D^\infty Df(x|y,s)dx = E(\tilde{x}|y,s) - U(D,y,s) < M - U(D,y,s) < M$. Since the shareholder has all the bargaining power, it would offer to bondholders a deterministic amount of $V(D,y,s)$ out of the liquidation amount M. Bondholders, indifferent to both continuing and terminating, would agree to the deal while shareholders would be better off since $M - V(D,y,s) > E(\tilde{x}|y,s) - V(D,y,s)$. Thus, the project will be liquidated, with bondholders getting $V(D,y,s)$ and shareholders getting $M - V(D,y,s)$.

Case 3: when $y > y^c$ and $s < s^*(y)$.

In this case the shareholders have the decision right and $E(\tilde{x}|y,s) > M$. Thus, the shareholders are going to continue the project without renegotiation since any renegotiation would not result in any Pareto improvement. The bondholders would get in expectation $V(D,y,s)$ while the shareholders would get in expectation $E(\tilde{x}|y,s) - V(D,y,s)$.

Case 4: when $y < y^c$ and $s > s^*(y)$.

In this case the bondholders have the decision rights and $E(\tilde{x}|y,s) > M$ as well as $E(\tilde{x}|y,s) > V(D,y,s)$. Thus, the bondholders have the decision rights and all the bargaining power. Depending on the value of s, we can have either $V(D,y,s) < M$ or $V(D,y,s) > M$ and the payoffs to both parties are different.[35]

Denote $s^{**}(y)$ as the value of s as a function of y such that $V(D,y,s^{**}(y)) = M$. Note that since $V(D,y,s)$ increases in both y and s, $V(D,y,s^{**}(y)) = M = V(D,y^c,s^*(y^c)) > V(D,y,s^*(y^c))$ for any $y < y^c$. Thus we have $s^{**}(y) > s^*(y^c)$.

[35]since $V(D,y,s) < D$ for any y and s and $M < D$, it is not obvious which one is larger.

Divide case 4 into two sub cases.

Case 4a. When $y < y^c$ and $s^*(y) < s < s^{**}(y)$.

Since bondholders have the decision right and their expected payoff when the project continues without negotiation is smaller than M, they would demand all the proceeds from continuing the project as $E(\tilde{x}|y,s) > V(D,y,s)$. If shareholders do not agree, they would liquidate the project and get M while shareholders still get nothing. Thus bondholders get $E(\tilde{x}|y,s)$ while shareholders get nothing. Note the bondholder's threat to liquidate is credible since, if shareholders reject, bondholders would liquidate the project because $M > V(D,y,s)$.

Case 4b. When $y < y^c$ and $s > s^{**}(y)$.

Now since $M < V(D,y,s)$, the threat from bondholders to liquidate the project is not credible since they would in expectation get more from continuing the project. Thus, whatever renegotiation deal the bondholders propose, shareholders would reject unless it can give them better than $E(\tilde{x}|y,s) - V(D,y,s)$. However, bondholders would never propose such a deal because this would give them less than $V(D,y,s)$. Thus, there would be no renegotiation and both parties are going to follow the original contract. Bondholders get $V(D,y,s)$ while shareholders get $E(\tilde{x}|y,s) - V(D,y,s)$.

The payoff or expected payoff from date 1 after observing both y and s is now summarized in the following table.

Table 2: Payoff matrix after y and s observed in date 1

(y,s)	Bondholder's Payoff	Equity Holder's Payoff
$y < y^c,\ s < s^*(y)$	M	0
$y > y^c,\ s > s^*(y)$	$V(D,y,s)$	$M - V(D,y,s)$
$y > y^c,\ s < s^*(y)$	$V(D,y,s)$	$E(\tilde{x}|y,s) - V(D,y,s)$
$y < y^c,\ s^*(y) < s < s^{**}(y)$	$E(\tilde{x}|y,s)$	0
$y < y^c,\ s > s^{**}(y)$	$V(D,y,s)$	$E(\tilde{x}|y,s) - V(D,y,s)$

Having solved the expected payoff for both parties at date 1, we can go back to solve the optimal debt contract, characterized by the principal value D^* and the covenant value y^* such that:

$$(y^*, D^*) \in \underset{(y^C,D)}{argmax}[\int_0^{y^C}\int_{s^{**}(y)}^{\bar{s}}(E(\tilde{x}|y,s) - V(D,y,s))h(y,s)dyds$$
$$+\int_{y^c}^{\bar{y}}\int_0^{s^*(y)}(M - V(D,y,s))h(y,s)dyds$$
$$+\int_{y^c}^{\bar{y}}\int_{s^*(y)}^{\bar{s}}(E(\tilde{x}|y,s) - V(D,y,s))h(y,s)dyds]$$

subject to

$$\int_0^{y^c}\int_0^{s*(y)} M * h(y,s)dyds + \int_{y^c}^{\bar{y}}\int_0^{s^*(y)} V(D,y,s)h(y,s)dyds$$
$$+\int_{y^c}^{\bar{y}}\int_{s^*(y)}^{\bar{s}} V(D,y,s)h(y,s)dyds + \int_0^{y^c}\int_{s^*(y)}^{s^{**}(y)} E(\tilde{x}|y,s)h(y,s)dyds$$
$$+\int_0^{y^c}\int_{s^{**}(y)}^{\bar{s}} V(D,y,s)h(y,s)dyds \geq K(1+R).$$

It is clear from the constraint that the higher the payoff to debtholders, the higher $V(D,y,s)$ and thus the lower the payoff to bondholders. Therefore, the constraint must bind and thus we insert the binding constraint into the objective function and get

$$(y^*, D^*) \in argmax_{(y^C,D)}[-K(1+R) + \int_0^{\bar{y}}\int_0^{s^*(y)} M * h(y,s)dyds$$
$$+\int_0^{\bar{y}}\int_{s^*(y)}^{\bar{s}} E(\tilde{x}|y,s) * h(y,s)dyds]$$

It is straightforward to observe that the objective function is independent of y^c and D.

□

In order to prove proposition 5, we need to prove the following lemma first. The lemma is stated and then proved.

Lemma 6. $\int_0^{\bar{y}} \frac{d}{d\delta}[\psi(y, s^*(y,t,\delta)|x_i,\delta)] * \frac{\partial s^*(y,t,\delta)}{\partial t}dy = 0, \quad i = H, L$ *and almost all* t.

Proof of Lemma 6:

Proof. We know that because the marginal distribution of s, $l(s|x_i)$ doesn't depend on δ, we have $\int_0^{\bar{y}} \frac{\partial}{\partial\delta}\psi(y,s|x_i,\delta)dy = 0$ for $i = H, L$ and for any s.

Note that in this case we are fixing y and s and change δ, but changing δ is equivalent to changing t (subject to some derivative term). More formally, define $\delta^*(y,s,t)$ to be the solution of the equation $\frac{\psi(y,s|x_H,\delta^*(y,s,t))}{\psi(y,s|x_L,\delta^*(y,s,t))} = t$. Now express $\frac{\partial}{\partial \delta}\psi(y,s|x_i,\delta)$ as

$$\frac{\partial}{\partial \delta}\psi(y,s|x_i,\delta) = \frac{\partial}{\partial t}\psi(y,s|x_i,\delta^*(y,s,t)) * \frac{1}{\frac{\partial}{\partial t}\delta^*(y,s,t)}.$$

Now I express $\frac{d}{d\delta}[\psi(y,s^*(y,t,\delta)|x_i,\delta)]$ in an alternative way. It is more appropriate to rewrite the expression as $\frac{\partial}{\partial \delta}[\psi(y,s^*(y,t,\delta)|x_i,\delta)]$ as it is the partial derivative holding y and t fixed. Using the $\delta^*(y,s,t)$ notation above, we can rewrite it as

$$\frac{\partial}{\partial \delta}[\psi(y,s^*(y,t,\delta)|x_i,\delta)] = \frac{\partial}{\partial s}\psi(y,s|x_i,\delta^*(y,s,t)) * \frac{1}{\frac{\partial}{\partial s}\delta^*(y,s,t)}.$$

Note that

$$\frac{1/\frac{\partial}{\partial s}\delta^*(y,s,t)}{1/\frac{\partial}{\partial t}\delta^*(y,s,t)} = \frac{\frac{\partial}{\partial t}\delta^*(y,s,t)}{\frac{\partial}{\partial s}\delta^*(y,s,t)} = -\frac{1/\frac{\partial}{\partial \delta}(\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)})|_{\delta=\delta^*(y,s,t)}}{\frac{\partial}{\partial s}(\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)})|_{\delta=\delta^*(y,s,t)}/\frac{\partial}{\partial \partial}(\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)})|_{\delta=\delta^*(y,s,t)}}$$

$$= -\frac{1}{\frac{\partial}{\partial s}(\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)})|_{\delta=\delta^*(y,s,t)}}$$

$$= -\frac{1}{\frac{\partial}{\partial s}(\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)})|_{s=s^*(y,t,\delta)}} = \frac{\partial s^*(y,t,\delta)}{\partial t}.$$

Now note that for arbitrary non-zero functions $h(s,t)$ and $l(s,t)$, we have

$$\frac{\partial^2}{\partial y \partial t}(log\frac{\psi(y,s|x_H,\delta^*(y,s,t))}{h(s,t)*t*\psi(y,s|x_L,\delta^*(y,s,t))}) = \frac{\partial^2}{\partial y \partial t}(log(t/(t*h(s,t)))) = 0$$

$$= \frac{\partial^2}{\partial y \partial s}(log(t/(t*l(s,t)))) = \frac{\partial^2}{\partial y \partial s}(log\frac{\psi(y,s|x_H,\delta^*(y,s,t))}{l(s,t)*t*\psi(y,s|x_L,\delta^*(y,s,t))}).$$

Thus integrating over y of the two second-order partial differential equations, we have

$\frac{\partial log\psi(y,s|x_H,\delta^*(y,s,t))}{\partial t} - \frac{\partial log(h(s,t)*\psi(y,s|x_L,\delta^*(y,s,t)))}{\partial t} = F(t,s)$ and

$\frac{\partial log\psi(y,s|x_H,\delta^*(y,s,t))}{\partial s} - \frac{\partial log(l(s,t)*\psi(y,s|x_L,\delta^*(y,s,t)))}{\partial s} = G(t,s)$ for some function F and G.

Because $\psi(y,s|x_H,\delta^*(y,s,t)) = t\psi(y,s|x_L,\delta^*(y,s,t))$ we have

$$(\frac{1}{t} - \frac{1}{h(s.t)}) * \frac{\partial}{\partial t}[(t - h(s,t)) * \psi(y,s|x_L,\delta^*(y,s,t))] = F(t,s) * \psi(y,s|x_L,\delta^*(y,s,t))$$

and

$$(\frac{1}{t} - \frac{1}{l(s,t)}) * \frac{\partial}{\partial s}[(t - l(s,t)) * \psi(y,s|x_L,\delta^*(y,s,t))] = G(t,s) * \psi(y,s|x_L,\delta^*(y,s,t)).$$

It follows that we have

$$(\frac{1}{t} - \frac{1}{h(s.t)}) * \frac{\partial}{\partial t}[(t - h(s,t))] * \psi(y,s|x_L,\delta^*(y,s,t))$$

$$+(\frac{1}{t} - \frac{1}{h(s.t)}) * (t - h(s,t)) * \frac{\partial}{\partial t}\psi(y,s|x_L,\delta^*(y,s,t)) = F(t,s) * \psi(y,s|x_L,\delta^*(y,s,t))$$

and

$$(\frac{1}{t} - \frac{1}{l(s.t)}) * \frac{\partial}{\partial s}(t - l(s,t)) * \psi(y,s|x_L,\delta^*(y,s,t)))$$

$$+(\frac{1}{t} - \frac{1}{l(s.t)}) * (t - l(s,t)) * \frac{\partial}{\partial s}\psi(y,s|x_L,\delta^*(y,s,t)) = F(t,s) * \psi(y,s|x_L,\delta^*(y,s,t))$$

It then follows that

$(\frac{1}{t} - \frac{1}{h(s.t)}) * (t - h(s,t)) * \frac{\partial}{\partial t}\psi(y, s|x_L, \delta^*(y, s, t)) = (F(t, s) - (\frac{1}{t} - \frac{1}{h(s.t)}) * \frac{\partial}{\partial t}(t - h(s, t))) * \psi(y, s|x_L, \delta^*(y, s, t))$

and

$(\frac{1}{t} - \frac{1}{l(s.t)}) * (t - l(s,t)) * \frac{\partial}{\partial s}\psi(y, s|x_L, \delta^*(y, s, t)) = (G(t, s) - (\frac{1}{t} - \frac{1}{l(s.t)}) * \frac{\partial}{\partial s}(t - l(s, t))) * \psi(y, s|x_L, \delta^*(y, s, t))$

Solving explicitly for $F(t, s)$ and $G(t, s)$ from the differential equations above gives

$F(t, s) = \frac{\partial}{\partial t}[-log(h(s, t))] = -\frac{1}{h(s,t)} * \frac{\partial}{\partial t}h(s, t)$ and

$G(t, s) = \frac{\partial}{\partial s}[-log(l(s, t))] = -\frac{1}{l(s,t)} * \frac{\partial}{\partial s}l(s, t)$.

Inserting into the equations above gives

$(\frac{1}{t} - \frac{1}{h(s.t)}) * (t - h(s, t)) * \frac{\partial}{\partial t}\psi(y, s|x_L, \delta^*(y, s, t)) = (-\frac{1}{t} + \frac{1}{h(s,t)} + \frac{1}{t} * \frac{\partial h(s,t)}{\partial t} - \frac{2}{h(s,t)} * \frac{\partial h(s,t)}{\partial t}) * \psi(y, s|x_L, \delta^*(y, s, t))$

and

$(\frac{1}{t} - \frac{1}{l(s.t)}) * (t - l(s, t)) * \frac{\partial}{\partial s}\psi(y, s|x_L, \delta^*(y, s, t)) = (\frac{1}{t} * \frac{\partial l(s,t)}{\partial t} - \frac{2}{l(s,t)} * \frac{\partial l(s,t)}{\partial t}) * \psi(y, s|x_L, \delta^*(y, s, t))$

Since $h(s, t)$ and $l(s, t)$ are arbitrary, choose $h(s, t)$ and $l(s, t)$ such that they are almost everywhere not equal to t and that

$$-\frac{1}{t}+\frac{1}{h(s,t)}+\frac{1}{t}*\frac{\partial h(s,t)}{\partial t}-\frac{2}{h(s,t)}*\frac{\partial h(s,t)}{\partial t}\neq 0 \; and \; \frac{1}{t}*\frac{\partial l(s,t)}{\partial t}-\frac{2}{l(s,t)}*\frac{\partial l(s,t)}{\partial t}\neq 0$$

almost everywhere. Since all the coefficients are non-zero almost everywhere, we can divide the above two equations, we have

$\frac{\frac{\partial}{\partial s}\psi(y,s|x_L,\delta^*(y,s,t))}{\frac{\partial}{\partial t}\psi(y,s|x_L,\delta^*(y,s,t))}=K(t,s)$ for some function of t and s almost everywhere.

Now we have

$$\frac{\partial}{\partial\delta}[\psi(y,s^*(y,t,\delta)|x_L,\delta)]=\frac{\partial}{\partial s}\psi(y,s|x_L,\delta^*(y,s,t))*\frac{1}{\frac{\partial}{\partial s}\delta^*(y,s,t)}$$

$$=K(t,s)*\frac{\partial}{\partial t}\psi(y,s|x_L,\delta^*(y,s,t))*\frac{1}{\frac{\partial}{\partial t}\delta^*(y,s,t)}*\frac{1/\frac{\partial}{\partial t}\delta^*(y,s,t)}{1/\frac{\partial}{\partial s}\delta^*(y,s,t)}$$

$$=K(t,s)*\frac{\partial}{\partial\delta}\psi(y,s|x_L,\delta)*\frac{1}{\frac{\partial s^*(y,t,\delta)}{\partial t}}.$$

Thus we have

$\frac{\partial}{\partial\delta}[\psi(y,s^*(y,t,\delta)|x_L,\delta)]*\frac{\partial s^*(y,t,\delta)}{\partial t}=K(t,s)*\frac{\partial}{\partial\delta}\psi(y,s|x_L,\delta)$ almost everywhere.

Now, integrating over y from 0 to $\bar{y}$ gives

$$\int_0^{\bar{y}}\frac{d}{d\delta}[\psi(y,s^*(y,t,\delta)|x_L,\delta)]*\frac{\partial s^*(y,t,\delta)}{\partial t}dy=\int_0^{\bar{y}}K(t,s)*\frac{\partial}{\partial\delta}\psi(y,s|x_L,\delta)dy$$

$=K(t,s)*\int_0^{\bar{y}}\frac{\partial}{\partial\delta}\psi(y,s|x_L,\delta)dy=0$ almost everywhere.

It then follows that

$$\int_0^{\bar{y}}\frac{d}{d\delta}[\psi(y,s^*(y,t,\delta)|x_H,\delta)]dy=\int_0^{\bar{y}}\frac{d}{d\delta}[t*\psi(y,s^*(y,t,\delta)|x_L,\delta)]dy$$

$= t * \int_0^{\bar{y}} \frac{d}{d\delta}[\psi(y, s^*(y,t,\delta)|x_L,\delta)]dy = 0$ almost everywhere.

□

Proof of Proposition 5:

Proof. The derivative of $\Omega(\delta)$ with respect to δ is

$$\frac{d\Omega(\delta)}{d\delta} = -P_H(x_H - M) * \frac{\partial \underline{t}(\delta)}{\partial\delta} * \int_0^{\bar{y}} \psi(y, s^*(y, \underline{t}(\delta), \delta)|x_H, \delta) * \frac{\partial}{\partial t} s^*(y,t,\delta)|_{t=\underline{t}(\delta)} dy$$

$$+P_H(x_H - M) * \int_{\underline{t}(\delta)}^{K} \int_0^{\bar{y}} \frac{d}{d\delta}[\psi(y, s^*(y,t,\delta)|x_H,\delta) * \frac{\partial}{\partial t} s^*(y,t,\delta)]dydt$$

$$+P_L(M - x_L) * \int_0^{\bar{y}} \psi(y, s^*(y, \underline{t}(\delta), \delta)|x_L, \delta) * \frac{\partial}{\partial t} s^*(y,t,\delta)|_{t=\underline{t}(\delta)} dy$$

$$-P_L(M - x_L) * \int_{\underline{t}(\delta)}^{K} \int_0^{\bar{y}} \frac{\partial}{\partial\delta}[\psi(y, s^*(y,t,\delta)|x_L,\delta) * \frac{\partial}{\partial t} s^*(y,t,\delta)]dydt$$

The first and third terms are zero. The reason is that because $\frac{\psi(y,0)|x_H,\delta)}{\psi(y,0)|x_L,\delta)} > \frac{\psi(0,0)|x_H,\delta)}{\psi(0,0)|x_L,\delta)} = \underline{t}(\delta)$, by our definition $s^*(y, \underline{t}(\delta), \delta)) = 0$ for all y that is non-zero and $\frac{\partial}{\partial t} s^*(y,t,\delta)|_{t=\underline{t}(\delta)} = 0$ for all y that is non-zero. Thus, the terms inside the integration sign for both the first and third terms are zero almost everywhere, resulting in those terms being zero. For the second term, we only need to consider the terms where $\frac{\partial}{\partial t} s^*(y,t,\delta) \neq 0$, i.e., where $\frac{\psi(y,s)|x_H,\delta)}{\psi(y,s)|x_L,\delta)} = t$ has solution for some $s \in [0, \bar{s}]$. For those terms taking the derivative gives

$$\frac{\partial}{\partial \delta}[\psi(y, s^*(y,t,\delta)|x_i,\delta) * \frac{\partial}{\partial t}s^*(y,t,\delta)] = \frac{\partial}{\partial \delta}[\psi(y, s^*(y,t,\delta)|x_i,\delta)] * \frac{\partial}{\partial t}s^*(y,t,\delta)$$

$$+\psi(y, s^*(y,t,\delta)|x_i,\delta) * \frac{\partial}{\partial \delta}(\frac{\partial}{\partial t}s^*(y,t,\delta)),\ i = H, L.$$

Denote $V(y,\delta)$ as the set of t such that $\frac{\partial}{\partial t}s^*(y,t,\delta) \neq 0$ for given y and δ, we now have the sum of the non-zero second and forth terms as

$$\int_{\underline{t}(\delta)}^{K} \int_0^{\bar{y}} \frac{\partial}{\partial \delta}[P_H(x_H - M) * \psi(y, s^*(y,t,\delta)|x_H,\delta)$$

$$-P_L(M - x_L) * \psi(y, s^*(y,t,\delta)|x_L,\delta)] * \frac{\partial}{\partial t}s^*(y,t,\delta)dydt$$

$$+\int_{\underline{t}(\delta)}^{K} \int_{V(y,\delta)}[P_H(x_H - M)\psi(y, s^*(y,t,\delta)|x_H,\delta)$$

$$-P_L(M - x_L)\psi(y, s^*(y,t,\delta)|x_L,\delta)]\frac{\partial}{\partial \delta}\frac{\partial}{\partial t}s^*(y,t,\delta)dydt$$

The first term of the above expression is zero because from Lemma 6, we have that

$$\int_0^{\bar{y}} \frac{\partial}{\partial \delta}[\psi(y, s^*(y,t,\delta)|x_i,\delta)] * \frac{\partial}{\partial t}s^*(y,t,\delta)dy = 0 \text{ for almost all t.}$$

Therefore

$$\int_{\underline{t}(\delta)}^{K} \int_0^{\bar{y}} \frac{\partial}{\partial \delta}[P_H(x_H - M) * \psi(y, s^*(y,t,\delta)|x_H,\delta)$$

$$-P_L(M - x_L) * \psi(y, s^*(y,t,\delta)|x_L,\delta)] * \frac{\partial}{\partial t}s^*(y,t,\delta)dydt$$

$$= P_H(x_H - M) * \int_{\underline{t}(\delta)}^{K} \int_0^{\bar{y}} \frac{\partial}{\partial \delta}[\psi(y, s^*(y,t,\delta)|x_H,\delta)] * \frac{\partial}{\partial t}s^*(y,t,\delta)dydt$$

$$-P_L(M - x_L) * \int_{\underline{t}(\delta)}^{K} \int_0^{\bar{y}} \frac{\partial}{\partial \delta}[\psi(y, s^*(y,t,\delta)|x_L,\delta)] * \frac{\partial}{\partial t}s^*(y,t,\delta)dydt$$

$= 0 - 0 = 0.$

Note that because $t \leq K$, from Lemma 3 we know that

$P_H(x_H - M)\psi(y, s^*(y,t,\delta)|x_H, \delta) < P_L(M - x_L)\psi(y, s^*(y,t,\delta)|x_L, \delta)$. Thus the sign of $\frac{d\Omega(\delta)}{d\delta}$ depends on the sign of $\frac{\partial}{\partial \delta}\frac{\partial}{\partial t}s^*(y,t,\delta)$. From Lemma 4 we know that $\frac{\partial s^*(y,t,\delta)}{\partial t} = 1/\{\frac{\partial}{\partial s}[\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)}]|_{s=s^*(y,t,\delta)}\}$.

So we have

$$sgn(\frac{\partial}{\partial \delta}\frac{\partial}{\partial t}s^*(y,t,\delta))$$

$$= -sgn(\frac{\partial}{\partial \delta}\{\frac{\partial}{\partial s}[\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)}]|_{s=s^*(y,t,\delta)}\}).$$

Now

$$\frac{\partial}{\partial s}[\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)}]|_{s=s^*(y,t,\delta)}$$

$$= \frac{\psi_s(y,s^*(y,t,\delta)|x_H,\delta)*\psi(y,s^*(y,t,\delta)|x_L,\delta) - \psi_s(y,s^*(y,t,\delta)|x_L,\delta)*\psi(y,s^*(y,t,\delta)|x_H,\delta)}{\psi^2(y,s^*(y,t,\delta),|x_L,\delta)}.$$

So we have

$$\frac{\partial}{\partial \delta}\{\frac{\partial}{\partial s}[\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)}]|_{s=s^*(y,t,\delta)}\}$$

$$= \frac{\partial}{\partial \delta}[\frac{\psi_s(y,s^*(y,t,\delta)|x_H,\delta)*\psi(y,s^*(y,t,\delta)|x_L,\delta) - \psi_s(y,s^*(y,t,\delta)|x_L,\delta)*\psi(y,s^*(y,t,\delta)|x_H,\delta)}{\psi^2(y,s^*(y,t,\delta),|x_L,\delta)}].$$

Now if assumption 6b is satisfied, $\frac{\partial^2}{\partial s \partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}] < 0$,

we have $\frac{\partial^2}{\partial s \partial \delta}[\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)}] = \frac{\partial^2}{\partial s \partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)} * \frac{l(s|x_H)}{l(s|x_L)}]$

$= \frac{\partial^2}{\partial s \partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}] * \frac{l(s|x_H)}{l(s|x_L)} + \frac{\partial}{\partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}] * \frac{\partial}{\partial s}[\frac{l(s|x_H)}{l(s|x_L)}] < 0$ as both terms are negative by assumption 2,5 and 6b. Thus

$$\frac{\partial}{\partial \delta}[\frac{\psi_s(y,s|x_H,\delta)*\psi(y,s|x_L,\delta)-\psi_s(y,s|x_L,\delta)*\psi(y,s|x_H,\delta)}{\psi^2(y,s|x_L,\delta)}] < 0 \quad \forall y, s.$$

It follows that $\frac{\partial}{\partial \delta}\{\frac{\partial}{\partial s}[\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)}]|_{s=s^*(y,t,\delta)}\} < 0$. So $sgn(\frac{\partial}{\partial \delta}\frac{\partial}{\partial t}s^*(y,t,\delta)) > 0$ and $\frac{d\Omega(\delta)}{d\delta} < 0$.

Now if assumption 6a is satisfied, $\frac{\partial^2}{\partial s \partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}] > 0$, the sign of $\frac{\partial^2}{\partial s \partial \delta}[\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)}]$ is ambiguous because the first term is positive and the second term is negative. However, if $\frac{\partial^2}{\partial s \partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}]$ is sufficiently positive, i.e., if $\frac{\partial^2}{\partial s \partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}] > \frac{-\frac{\partial}{\partial \delta}[\frac{g(y|s,x_H,\delta)}{g(y|s,x_L,\delta)}]*\frac{\partial}{\partial s}[\frac{l(s|x_H)}{l(s|x_L)}]}{\frac{l(s|x_H)}{l(s|x_L)}}$, we would have $\frac{\partial^2}{\partial s \partial \delta}[\frac{\psi(y,s|x_H,\delta)}{\psi(y,s|x_L,\delta)}] > 0$ and $\frac{d\Omega(\delta)}{d\delta} > 0$.

□