Accounting Conservatism and Incentives:
Intertemporal Considerations*

Jonathan Glover
Columbia School of Business
Columbia University
jg3463@columbia.edu

Haijin Lin
C.T. Bauer College of Business
University of Houston
haijinlin@uh.edu

January 2018

* We thank Sudipta Basu, Hui Chen, Joel Demski, Eti Einhorn, Florin Sabac, Shiva Sivaramakrishnan, Phillip Stocken (the Editor), Shyam Sunder, Tsahi Versano, Alfred Wagenhofer, an anonymous referee, seminar participants at Rice University and the University of Zurich, and participants at the 10th Tel Aviv Conference in Accounting and the 18th Annual Carnegie Mellon University Accounting Mini Conference for their comments and suggestions.
Accounting Conservatism and Incentives: Intertemporal Considerations

Abstract. We study intertemporal incentive properties of conditional accounting conservatism. Conservatism has detrimental and beneficial properties. In our first model, conservatism introduces downward bias in the first period; any understatement of first-period performance is reversed in the second period. A conservative bias is not costly in the first period but instead costly in the second period when a new manager may be rewarded for the performance of his predecessor. In an extension on learning, we illustrate a beneficial role of conservatism in fine-tuning incentives. In the second model, conservatism is modeled as recognizing effort-independent bad news early and good news late. Recognizing bad news early can be optimal because of intertemporal rent shifting, which improves incentives via an “incentive spillback.” We also study overlapping projects (a multi-task setting) in which an interior accounting system can be optimal to avoid making one of the overlapping projects an incentive bottleneck.

Keywords: accounting conservatism, multi-period incentives, incentive spillback, multi-task incentives

JEL Classifications: D21, D74, D82, D86
I. INTRODUCTION

Conservatism is a pervasive feature of accounting.\(^1\) The concept of conservatism predates Pacioli and is used throughout the world (Basu 1995).\(^2\) There are many definitions of accounting conservatism. Kohler’s Dictionary (Cooper and Ijiri 1983) defines conservatism as “a guideline which chooses between acceptable accounting alternatives … so that the least favorable immediate effect on assets, income, and owner’s equity is reported.” According to Watts (2003) and Basu (1997), conservatism requires a “higher degree of verification to recognize good news as gains than to recognize bad news as losses.” Sunder (1997) notes “[t]he presence of uncertainty and the downward bias of measured current-period income, assets, and owner’s equity in the presence of uncertainty seem to be the essential aspects of conservatism.”\(^3\)

Accounting conservatism has long been criticized by accounting scholars, accounting practitioners, and others.\(^4\) As Paton and Paton (1952) write, “[i]s there anything essentially conservative … in a valuation scheme that merely shifts income from one period to the next.”

Conservatism has also been criticized by accounting standard setters.\(^5\)

In this paper, we study the intertemporal properties of conservatism with a focus on managerial incentives. Intertemporal reversals turn out to have both detrimental and beneficial

---

1. Conservatism is not a convention in the sense that it is not a matter of indifference to economic agents whether we adopt conservative, unbiased, or aggressive accounting (Sunder 1997).
2. Basu (1995) cites evidence in Penndorf (1933) that Francesco di Marco of Prato used lower of cost or market to value his inventory in 1406.
3. Bliss (1924) defines conservatism as: “anticipate no profit, but anticipate all losses.” Sanders, Hatfield, and Moore (1938) discuss conservatism, capital and income, and the form and terminology of financial statements as three general considerations in accounting. Sterling (1970) describes conservatism as accounting’s most influential valuation principle. Conservatism can also be viewed as a scaling issue, as in Demski and Sappington (1990).
4. According to Paton and Stevenson (1918, 475), lower-of-cost-or-market rule “would require some very important practical considerations to justify such an illogical procedure.” As Sanders, Hatfield, and Moore (1938, p. 12) write, “[t]he common belief that less mischief is done by understatement than by overstatement is, in the hands of honest men probably true; but with dishonest men understatement may serve their turn as well as overstatement.”
5. As early as 1980, in their Statement of Financial Accounting Concepts 2 (CON 2), the FASB seemed to view conservatism as an outdated idea that arose when the balance sheet was the only readily available financial statement (FASB 1980, paragraph 93).
incentive properties. We study simple two-period models of moral hazard with a risk-neutral principal who designs short-term linear (in periodic performance) contracts for a risk-neutral agent subject to limited liability.

The notion of conservatism we study can be thought of as conservatism dictated by accounting policies imposed by a firm’s board of directors or as conservatism baked into accounting standards such as lower of cost or market rather than auditor or managerial conservatism. Our conservatism is conditional—it depends on an underlying economic event and results in more timely recognition of bad news than good news (see Basu 1997).

In the first model, conservative (aggressive) bias is a possible understatement (overstatement) of performance that will be reversed in a second period. Bias distorts the information content of the performance measures, so is costly (relative to unbiased accounting) if different agents are employed in each period. A conservative bias is not costly in the first period because the first-period agent is never rewarded in error; instead, the cost arises in the second period when a new manager may be rewarded for the performance of his predecessor—when good performance from the current period is comingled with good performance from the previous period. Under an aggressive bias, it is the cost of providing incentives to the first-period agent that is costly because he may be rewarded in error.

If instead the same agent is retained in both periods, the second-period reversal of any understated first-period performance allows conservative accounting to replicate the performance of unbiased accounting. Because the same bonus rates are (optimally) employed in both periods, the comingling of first- and second-period performance in the second period is not costly. In contrast, aggressive accounting cannot replicate unbiased accounting, since the principal cannot
fully reverse (implicitly “claw back”) any first-period overpayment because of the agent’s limited liability.

When the principal makes performance-contingent turnover (replacement) decisions, conservatism can no longer replicate the performance of unbiased accounting. Under unbiased accounting, the principal optimally replaces the agent following first-period poor performance. Such replacement decisions allow the principal to use the agent’s retention rent to reduce the cost of providing incentives in the first period. However, under a conservative bias, performance-contingent replacement comes with the cost of the first-period agent no longer being in the firm’s employ when a first-period understatement is corrected (reversed) in the second period.

We then introduce learning. If the same agent is retained in both periods and the principal learns about the agent’s productivity over time, a conservative bias strictly dominates not only an aggressive bias but also unbiased accounting. A conservative bias allows the principal to effectively retroactively fine-tune the first-period compensation once uncertainty about the agent’s type has been resolved by applying the second-period bonus rate to good performance that was underreported in the first period and reversed in the second period.

In our second model, conservatism is not bias but rather the choice to recognize bad shocks (news) early and good shocks late. The shocks recognized by the accounting systems are independent of the agent’s effort (e.g., price-level adjustments driven by macroeconomic factors). Conservative (aggressive) accounting shifts the agent’s rents from the first (second) to the second (first) period. When the same agent is retained in both periods, the accounting choice is irrelevant. Any extra rents created in the second period are offset by the reduced rents in the first period and vice versa. However, when performance-contingent turnover is introduced, conservative accounting dominates both aggressive and unbiased accounting. The reason is that
conservative accounting maximizes the benefit to the agent of keeping his job (his retention rent) and, hence, creates the largest possible incentive spillback, allowing the principal to reduce the agent’s first-period bonus. In some cases, it is optimal to recognize both good shocks and bad shocks in the same period, but the optimal system continues to be a conservative one in that shocks are recognized in the first (second) period when the aggregate shock is negative (positive).

In an extension of the timely news recognition model, we study overlapping projects that give rise to a multi-task setting (Holmstrom and Milgrom 1991). The main result that emerges from the multi-task setting is that an interior accounting system (neither maximally conservative nor maximally aggressive) can be optimal because extreme systems can make one of the overlapping projects a bottleneck.

Our paper contributes to the now vast literature on the economic demand for accounting conservatism. For example, accounting conservatism’s properties for debt contracting have been studied by Gigler et al. (2009), Caskey and Hughes (2012), Li (2013), and others. Conservatism can facilitate trade in debt markets (Gox and Wagenhofer 2009) or in asset resale markets (Demski et al. 2009). Gao (2013) formalizes conservatism as an ex ante reaction to ex post managerial opportunism.

Closest to our paper are those in which accounting conservatism arises in equilibrium to improve managerial contracting efficiency even without reporting opportunism (for example, Christensen and Demski 2002, 2004; Kwon et al. 2001). Somewhat surprisingly, the existing models of accounting conservatism mostly focus on single-period (or essentially single-period) settings. Our news model shares features of Drymiotes and Hemmer’s (2013) model. However,
their focus is on discretionary accruals (chosen by the manager) and the stock market’s reaction to those discretionary accruals.⁶

The extension of our bias model to include a firing/retention decision is related to the “at-will contracts” in Arya, Glover and Sunder (1998) (AGS). The two models differ in two aspects. First, the principal is assumed to be able to commit to a firing/retention strategy as part of the contract in our model; in AGS, the principal cannot commit to a fire/retention strategy. The role of earnings management in AGS is to act as a substitute for the principal’s commitment. Second, we show that downward bias makes firing costlier so that an inefficient firing strategy is induced with bias; in AGS, earnings management ensures the principal adopts (ex ante) efficient firing/retention decisions.

The remainder of the paper is organized into four sections. Section II studies conservatism as bias subject to a later reversal (an intertemporal “true up” of the performance measures to cash flows). Section III studies conservatism as an information system that recognizes bad news early and good news late. We end Sections II and III with discussions about the implications of introducing agent risk aversion. Section IV studies the overlapping projects (multi-task) model, and Section V concludes. All the proofs are in the appendix.

II. CONSERVATISM AS BIAS

A risk-neutral principal contracts with a risk-neutral agent to implement a two-period project. The agent sequentially selects two productive inputs, low or high, denoted by $a_t \in \{a_{tL}, a_{tH}\}$, $t = 1, 2, a_{H} > a_{L}$. With a slight abuse of notation, $a_t$ also denotes the

⁶ Levine (1996), Bagnoli and Watts (2005), and Lin (2006) also study managerial accrual choices and show the choice of conservative accounting methods can serve as a signaling device.
agent’s disutility of the input in period $t$. Periodic unbiased performance, low or high, is denoted by $x_t \in \{0, 1\}$. The technology is described by the conditional probability of performance $x_t$ given the agent’s input. The first-period input does not affect the second-period performance measure, and the technologies are identical in both periods: $p \equiv Pr(x_t = 1|a_{tH})$ and $q \equiv Pr(x_t = 1|a_{tL}), t = 1, 2$, and $p > q$.  

In this section, most of our analysis will focus on biased performance measures. Unless the accounting system generates it, the unbiased measure, $x_t$, should be viewed as an unobservable (even to the agent) modeling construct. Only the performance measure generated by the accounting system, denoted by $y_t \in \{0, 1\}$, is publicly observable and can be verified by the courts, so is contractible. Also, the agent’s second-period input choice can be conditioned on his observation of the first-period performance, $a_2(y_1) \in \{a_{2L}, a_{2H}\}$.  

The accounting system is subject to bias. We consider two biased systems: a conservative system and an aggressive system. For simplicity, we assume a conservative (aggressive) system always produces a low (high) report when the unbiased performance measure is low (high). With probability $\varepsilon_c$, a conservative system produces a low report when the unbiased measure is high; with probability $\varepsilon_a$, an aggressive system produces a high report when the unbiased measure is low. A “true up” at the end of the second period ensures the sum of the reports equals the sum of the unbiased performance measures over two periods: $y_1 + y_2 = x_1 + x_2$. Any misreport in the first period is corrected in the second period.  

A conservative accounting system introduces downward bias and shifts the recognition of some good performance from the first to the second period. In the second period, good performance is less informative about current effort since good performance might mean good current-period performance or good prior-period performance whose recognition was delayed.
For now, the information system can be thought of as exogenous, although we will eventually treat it as the first move of the game—one taken by the principal. The next move is the principal’s offer of a short-term (single-period) contract to the agent. We restrict attention to short-term “at-will” contracts to focus on the role of accounting rather than contracts in fostering relationships. For tractability and to prevent the principal from undoing the bias (e.g., by exploiting moving support or specifying payments as a function of the intertemporal sum of the performance measures), we restrict compensation $s_t$ to be a linear function of performance measure in that period. Specifically, the compensation in period $t$ is comprised of a salary and a bonus: $s_1 = \alpha_1 + \beta_1 \times \max\{0, y_1\}$ and $s_2 = \alpha_2(y_1) + \beta_2(y_1) \times \max\{0, y_2\}$. The second-period intercept (salary) and slope (bonus coefficient) can vary with $y_1$.

The agent’s payoff is given by $s_1 + s_2 - a_1 - a_2$. The agent’s expected compensation across the two periods is denoted by $E[s_1 + s_2|a_1, a_2(0), a_2(1); \varepsilon]$, where $\varepsilon$ is the bias introduced by the accounting system, $a_1$ is the agent’s first-period action, and $a_2(0)$ is the agent’s second-period input following a first-period low report ($y_1 = 0$), and $a_2(1)$ is his input following a high report ($y_1 = 1$). The agent’s expected compensation in the second period is denoted by $E[s_2|a_2(y_1); \varepsilon, y_1, a_1]$. The principal consumes the residual, resulting in a payoff of $y_1 + y_2 - s_1 - s_2 = x_1 + x_2 - s_1 - s_2$. Everything is common knowledge, except for the agent’s hidden actions (and the periodic unbiased performance measures unless the accounting system is unbiased). The principal’s objective is to motivate the agent to supply a high input in both periods at the least expected cost. Motivating high inputs is optimal for the principal if the agent’s marginal productivity in the performance measures is sufficiently large compared to his marginal disutility for effort. The sequence of the events is presented in Figure 1.
Figure 1: Time-line of the Base Model

For any system $\varepsilon \in \{\varepsilon_c, \varepsilon_a\}$, the principal’s problem is to minimize her expected incentive cost subject to the following constraints. The individual rationality constraint (IR1) ensures the contract is sufficiently attractive to the agent as of the beginning of the first period—the first-period contract and expected second-period contract must together provide the agent with an expected utility of at least his two-period reservation utility $U_1 + U_2$. The incentive compatibility constraint (IC1) ensures it is *ex ante* incentive compatible for the agent to supply high effort in both periods. The constraint (IR2) ensures the contract is sufficiently attractive to the agent as of the beginning of the second period, while the constraint (IC2) ensures it is incentive compatible for the agent to supply high input in the second period given the first-period performance. Finally, the payment to the agent must be non-negative in each period—the principal pays the agent, not the other way around (constraints NN1 and NN2). We first solve the principal’s second-period (short-term) contracting problem and then, taking that solution as given, solve the principal’s first-period problem. Formally, we write the principal’s Program(s) $P$ as follows and denote the solution(s) by $\{\alpha_1, \alpha_2(\cdot), \beta_1, \beta_2(\cdot)\}$.

**Program P—Period One**

$$\min_{\alpha_1, \beta_1} E[s_1 | a_{1H}; \varepsilon]$$

---

7 See Innes (1990) and Sappington (1983) for standard models of limited liability in principal-agent relationships.
Subject to

\[ E[s_1 + s_2|a_{1H}, a_{2H}, a_{2H}; \epsilon] - a_{1H} - a_{2H} \geq \bar{U}_1 + \bar{U}_2; \]  \quad (IR1)

\[ E[s_1 + s_2|a_{1H}, a_{2H}, a_{2H}; \epsilon] - a_{1H} - a_{2H} \geq E[s_1 + s_2 - a_2|a_{1i}, a_{2j}, a_{2k}; \epsilon] - a_{1i} \]
for all \( i, j, k \in \{L, H\} \);  \quad (IC1)

\[ s_1 = \alpha_1 + \beta_1 \times \max\{0, y_1\} \geq 0 \text{ for all } y_1; \]  \quad (NN1)

where \( s_2 \) is the solution to Program P – Period Two.

**Program P—Period Two**

\[
\min_{a_{2}(y_1), b_{2}(y_1)} E[s_2|a_{2H}; \epsilon, y_1, a_{1H}]
\]

Subject to

\[ E[s_2|a_{2H}; \epsilon, y_1, a_{1H}] - a_{2H} \geq \bar{U}_2 \text{ for all } y_1; \]  \quad (IR2)

\[ E[s_2|a_{2H}; \epsilon, y_1, a_{1H}] - a_{2H} \geq E[s_2|a_{2H}; \epsilon, y_1, a_{1H}] - a_{2L} \text{ for all } y_1; \]  \quad (IC2)

\[ s_2 = \alpha_2(y_1) + \beta_2(y_1) \times \max\{0, y_2\} \geq 0 \text{ for all } y_1, y_2. \]  \quad (NN2)

For simplicity, we assume \( \bar{U}_1 = \bar{U}_2 = a_L = 0 \). This ensures the incentive compatibility and the non-negative constraints dominate the individual rationality constraints. Program P assumes the same agent is employed in both periods. Because it helps illustrate the role of conservatism, we will also characterize the solution assuming a new agent is hired in the second period. In this case, the incentive compatibility (and individual rationality) constraints have to hold period-by-period.

Bias contaminates performance measurement, comingling the effects of the first- and second-period inputs in the second-period performance measure. Whether or not this comingling
is costly depends on the nature of the bias (aggressive vs. conservative) and whether the same agent is employed in both periods.

Under a conservative system, if the same agent is retained in both periods, the principal optimally offers the same bonus rate in each period as she would under unbiased measurement:

\[ \beta_1 = \beta_2 (\cdot) = \frac{a_H}{p-q}. \]

The first-period bonus is not sufficient to motivate high input in the first period, since some good performance is mistakenly reported as bad performance. However, the under-reporting is reversed in the second period. The reversal increases the probability of the agent receiving a bonus in the second period; the “overpayment” (an extra rent) serves to fill in the gap in the first-period incentives, because the second-period overpayment is a function of the agent’s first-period effort. The intertemporal reversal creates an incentive spillback—contracts in later periods help resolve agency conflicts in early periods because of the bias reversal/correction. Of course, this spillback arises only if the same agent is retained in the second period.

If instead the principal contracts with a new agent in the second period, the first-period bonus rate has to be increased to

\[ \beta_1 = \frac{a_H}{(p-q)(1-\varepsilon_c)} \]

However, the expected cost of providing incentives in the first period remains

\[ p(1 - \varepsilon_c) \frac{a_H}{(p-q)(1-\varepsilon_c)} = \frac{pa_H}{p-q}, \]

as under unbiased measurement. The reason a conservative bias does not create additional first-period rents is that the effect of conservatism on the probability of reporting good performance is multiplicative with effort, preserving the likelihood ratio associated with good performance \( q(1 - \varepsilon_c)/p(1 - \varepsilon_c) = q/p. \) To motivate the agent hired for the second period, the principal must use a bonus rate of

\[ \beta_2 (y_1) = \frac{a_H}{p-q}; \]

the second-period agent earns additional rents relative to those he would earn under unbiased measurement, since, under
a conservative bias, the agent hired in the second period is sometimes rewarded for good
performance generated by the first-period agent that was not recognized until the second period.
The lack of a spillback incentive from the second period to the first makes a conservative system
costlier than an unbiased system.

Lemma 1. Assume a conservative system is in place.

(i) If the same agent is employed in both periods, the optimal bonus rates are: \( \beta_1 = \beta_2(0) = \beta_2(1) = \frac{a_H}{p-q} \). The total expected incentive cost is \( \frac{2p_a H}{p-q} \), which does not vary
with the level of downward bias.

(ii) If a new agent is employed in each period, the optimal bonus rates are: \( \beta_1 = \frac{a_H}{(p-q)(1-\epsilon_c)} \) and \( \beta_2(0) = \beta_2(1) = \frac{a_H}{p-q} \). The total expected incentive cost is \( \frac{(2+\epsilon_c)p_a H}{p-q} \),
which increases in the level of downward bias.

The situation is different under an aggressive system. With upward bias, even if the same
agent is employed in both periods, the cost of providing incentives is higher than under unbiased
measurement. The reason is that any overpayment made in the first period can only be partially
reversed/clawed back in the second period—a reduction in the second-period performance from
1 to 0 is costly to the agent, but a reduction of the second-period performance from 0 to -1
imposes no cost on the agent because of his limited liability. So, the agent earns additional rents
in the first period under an aggressive system relative to what he would earn under an unbiased
system. In the second period, the bonus rate can be adjusted so that the agent receives the same
rent from his second-period effort as he would under unbiased measurement.
Now, consider an aggressive system with a new agent hired in the second period. Because good performance in the first period is less informative under an aggressive system than under an unbiased or a conservative system, the cost of providing incentives in the first period increases. \( \beta_1 \) is increased from \( \frac{a_H}{p-q} \) to \( \frac{a_H}{(p-q)(1-\epsilon_a)} \), and the expected cost of providing incentives to the first-period agent is \( \left[ p + (1-p)\epsilon_a \right] \frac{a_H}{(p-q)(1-\epsilon_a)} > \frac{p a_H}{p-q} \). The first-period agent now earns rents for two reasons: (i) the first-period performance measure is less informative (a worse likelihood ratio) and (ii) there is no bonus reversal/clawback at all. (Again, the second-period bonus can be adjusted so that the second-period input generates the same rents in the second period as it would under an unbiased measurement.)

**Lemma 2.** Assume an aggressive system is in place.

(i) If the same agent is employed in both periods, the optimal bonus rates are: 
\[
\beta_1 = \frac{(p-q)\epsilon_a}{p(p-q)(1-\epsilon_a)}, \quad \beta_2(0) = \frac{a_H}{p-q}, \quad \text{and} \quad \beta_2(1) = \frac{[p+(1-p)\epsilon_a] a_H}{p(p-q)}. \]

The total expected incentive cost is 
\[
\left[ \frac{(p+(1-p)\epsilon_a)(p-q)\epsilon_a}{p(1-\epsilon_a)} + p \right] \frac{a_H}{p-q}, \]
which increases in the level of upward bias.

(ii) If a new agent is employed in each period, the optimal bonus rates are: 
\[
\beta_1 = \frac{a_H}{(p-q)(1-\epsilon_a)}, \quad \beta_2(0) = \frac{a_H}{p-q}, \quad \text{and} \quad \beta_2(1) = \frac{[p+(1-p)\epsilon_a] a_H}{p(p-q)}. \]

The total expected incentive cost is 
\[
\left[ \frac{p+(1-p)\epsilon_a}{1-\epsilon_a} + p \right] \frac{a_H}{p-q}, \]
which increases in the level of upward bias.

Lemmas 1 and 2 imply that for any biased information system, a long-term relationship (a repeated encounter) is preferred to a short-term one. When the principal and the agent have a long-term relationship, a conservative system is no costlier than an unbiased system. The
conservative system has a built-in correction system that works without the principal knowing whether the first-period performance was actually understated (with the exception of the moving support cases). However, an upward bias introduced by an aggressive system leads to “excessive” bonus payments during early periods that cannot be fully clawed back when the bias is reversed. As the discussion suggests, a bonus bank that allows the principal to explicitly claw back the first-period bonus would put an aggressive system on equal footing with a conservative system. When a new agent is hired in each period, bias is always costly; however, a (same-size) conservative bias is less costly than an aggressive bias.

**Proposition 1.**

(i) If the same agent is employed in both periods, a conservative system replicates the unbiased system and always dominates an aggressive system.

(ii) If a new agent is employed in each period, a conservative system dominates an aggressive system if and only if \( \varepsilon_c p < \frac{\varepsilon_a}{1 - \varepsilon_a} \). The unbiased system dominates any biased system.

So far, the best a biased system can do is to replicate an unbiased system. In search of a positive demand for bias, and conservatism in particular, we next explore the effect of bias on repeated agency encounters when (i) the principal can make performance-contingent firing decisions or (ii) the principal learns about the environment over time.
Performance-Contingent Firing

If the principal can commit to condition the agent’s retention on his first-period performance, the second-period rent is an additional source of first-period effort incentives. That is, the agent’s first-period incentives are derived from two sources: the first-period bonus and the second-period rent the agent will receive if and only if he retains his job (an incentive spillback). The incentive spillback created by the retention rent differs from the spillback created by the reversal of downward bias under conservatism, as being fired is costlier to the agent than performance being understated. A natural question is: how does performance-contingent firing affect the ranking of information systems? We extend the base model to answer this question.

Suppose that the initial contract between the principal and the agent specifies the situations in which the agent will be retained or replaced by a new (ex ante identical) agent in the second period. Denote by $D(y_1)$ the $y_1$-contingent decision/commitment to hire a new agent or retain the existing agent at the beginning of the second period:

$$D(\cdot) = \begin{cases} 0 & \text{if the agent is fired} \\ 1 & \text{if the agent is retained.} \end{cases}$$

If a new agent is hired, he exerts effort $a_{2N}(y_1) \in \{a_{2L}, a_{2H}\}$ and is paid $s_{2N} = \alpha_{2N}(y_1) + \beta_{2N}(y_1) \times \max\{0, y_2\}$. The new agent’s effort and contract can vary with the first-period report. The sequence of the events is presented in Figure 2.

For any system $\epsilon \in \{\epsilon_c, \epsilon_d\}$, the principal optimally chooses his firing decision $D(\cdot)$ and the optimal contract, denoted by $\{\alpha_1, \alpha_2(\cdot), \alpha_{2N}(\cdot), \beta_1, \beta_2(\cdot), \beta_{2N}(\cdot)\}$, to minimize her expected incentive cost subject to the individual rationality constraints that ensure the contracts are sufficiently attractive to the agent at the beginning of each period, the incentive compatibility
constraints that ensure the agent works in each period if retained, the incentive compatibility constraints that ensure a new agent works if hired, and the non-negativity constraints.

Figure 2: Time-line with Performance-Contingent Firing

<table>
<thead>
<tr>
<th>$t = 0$</th>
<th>$t = 1$</th>
<th>$t = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accounting system with bias $\varepsilon$ is in place.</strong> Principal and agent sign a one-period contract, which includes a commitment to a replacement decision $D(y_1)$; agent supplies $a_1$.</td>
<td>$y_1$ is reported; agent is paid $s_1(\cdot)$ with bonus $\beta_1 \times \max{0, y_1}$; principal retains/fires agent as specified by $D(y_1)$.</td>
<td>If $D(y_1) = 0$, a new agent is hired; principal and the second-period agent sign a one-period contract.</td>
</tr>
<tr>
<td>$t = 1$</td>
<td>$t = 2$</td>
<td></td>
</tr>
</tbody>
</table>
| Agent supplies $a_2(y_1)$ or $a_{2N}(y_1)$. $y_2$ is reported; old agent is paid $s_2(\cdot)$ with bonus $\beta_2(y_1) \times \max\{0, y_2\}$ or new agent is paid $s_{2N}(\cdot)$ with bonus $\beta_{2N}(y_1) \times \max\{0, y_2\}$.

**Proposition 2.** Assume the principal has the option to fire the agent at the end of the first period. Further assume $\varepsilon_c = \varepsilon_a = \varepsilon$.

(i) Under the unbiased or an aggressive system, the principal optimally commits to fire the agent whenever a low report is observed. Under a conservative system, if $\varepsilon < \frac{a}{1+q}$, the principal optimally commits to fire the agent whenever a low report is observed; if $\varepsilon \geq \frac{a}{1+q}$, the principal never fires the agent.

(ii) If $\varepsilon < \frac{a}{1+q}$ and $p + q > 1$, then an aggressive system dominates a conservative system. If $\varepsilon > \frac{pq}{(1-p)(1-q)}$, a conservative system dominates an aggressive system.

(iii) The unbiased system dominates any biased system.
Under an aggressive or an unbiased system, retaining an agent only when the first-period performance report is high is optimal because it enables the principal to use the second-period rent to reduce the cost of providing incentives in the first-period (incentive spillover created by retention rent). Under a conservative system, there is the same benefit to retaining an agent when he performs well in the first period, but there is also a cost. Namely, under a conservative system, firing an agent after a poor first-period performance removes the ability of the information system to correct (reverse) any understatement of performance in the first period. Instead, the newly hired agent will reap the benefit of the correction in the second period.

The principal is no longer able to replicate the performance of an unbiased system with a conservative one. The optimal unbiased system always relies on performance-contingent firing, while a conservative system can only replicate an unbiased one by always retaining the same agent in both periods. Because of the benefit of conditional firing under an aggressive system, there is now the possibility that an aggressive system dominates a conservative one. However, because the benefit of retaining an agent under a conservative system remains (incentive spillover created by intertemporal reversal) when the principal retains the agent, a conservative system can also dominate an aggressive system. Overall, once the possibility of performance-contingent firing is introduced, the ranking of the conservative and aggressive systems is ambiguous, but the unbiased system again dominates any biased system.

**Learning**

Suppose the agent has private (hidden) information about his own productivity or the type of project that is to be implemented. Specifically, the agent can be productive (good type) or unproductive (bad type), denoted by $\theta \in \{b, g\}$. The principal has a prior belief about the
agent’s type, denoted by $Pr(\theta)$. In each period, the technology is a function of both the agent type and the agent’s current act: $p_\theta \equiv Pr(x_t = 1|\theta, a_{tH})$ and $q_\theta \equiv Pr(x_t = 1|\theta, a_{tL})$ for $t = 1, 2$. Assume the monotone likelihood ratio property for each type: $p_\theta > q_\theta$. We also assume $p_g - q_g > p_b - q_b$, i.e., a good-type agent’s marginal productivity is greater than a bad-type agent’s marginal productivity.

The principal learns the agent’s type at the end of the first period.\(^8\) We continue to restrict attention to linear contracts: $s_1 = \alpha_1 + \beta_1 \times \max\{0, y_1\}$; and $s_2 = \alpha_2(\theta, y_1) + \beta_2(\theta, y_1) \times \max\{0, y_2\}$. Because of the agent’s risk neutrality and binary-action, binary-type setup (with the high input to be motivated for both types), there is no role for communication in the first period. The principal’s first-period contracting problem boils down to choosing a bonus sufficiently large to motivate the agent of any type to supply high input in the first period. The second-period bonus rate is conditioned on the observed agent’s type and the first-period report. Again, there is no role for the agent’s communication of his type, since the principal learns the type before contracting with the agent in the second period.

The agent’s first-period input choice can be conditioned on $\theta$, while his second-period input choice can be conditioned on both his own type and his observation of the first-period performance, $a_1(\theta) \in \{a_{1L}, a_{1H}\}$ and $a_2(\theta, y_1) \in \{a_{2L}, a_{2H}\}$. The agent’s expected compensation across the two periods is denoted by $E[s_1 + s_2|a_1(\theta), a_2(\theta, 0), a_2(\theta, 1); \varepsilon, \theta]$ and

---

\(^8\) The agent’s type in our learning model can be thought of as “fit” with the firm, which the agent knows when hired but the principal learns at the end of the first period after working with the agent for a period. If the principal’s learning were instead imperfect and based on a separate non-verifiable signal and we relax our assumption that the agent must be motivated to supply high effort in all circumstances, the result presented in this Section of the paper would continue to hold qualitatively. With imperfect learning, the principal would have to live with a bad type agent shirking when offered a contract intended for a good type.
the agent’s expected compensation in the second period is denoted by

\[ E[s_2|a_2(\theta, y_1); \epsilon, \theta, y_1, a_1]. \]

The sequence of the events is presented in Figure 3.

**Figure 3: Time-line of the Learning Model**

<table>
<thead>
<tr>
<th>( t = 0 )</th>
<th>( t = 1 )</th>
<th>( t = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounting system with bias ( \epsilon ) is in place.</td>
<td>Principal and agent sign a one-period contract.</td>
<td>Principal observes ( \theta ). Principal and the agent sign a one-period contract.</td>
</tr>
<tr>
<td>Agent privately observes ( \theta ); agent supplies ( a_1 ).</td>
<td>( y_1 ) is reported; agent is paid ( s_1(\cdot) ) with bonus ( \beta_1 \times \max{0, y_1} ).</td>
<td>Agent supplies ( a_2(\theta, y_1) ). ( y_2 ) is reported; agent is paid ( s_2(\cdot) ) with bonus ( \beta_2(\theta, y_1) \times \max{0, y_2} ).</td>
</tr>
</tbody>
</table>

Parallel to Program P, Program PL describes the principal’s problem with its solution denoted by \( \{a_1, a_2(g, \cdot), a_2(b, \cdot), \beta_1, \beta_2(g, \cdot), \beta_2(b, \cdot)\} \).

**Program PL—Period One**

\[
\min_{a_1, \beta_1} \sum_{\theta} Pr(\theta) E[s_1|a_{1H}; \epsilon, \theta]
\]

Subject to

\[
E[s_1 + s_2|a_{1H}, a_{2H}; \epsilon, \theta] - a_{1H} - a_{2H} \geq U_1 + U_2 \quad \text{for all } \theta; \quad \text{(PL-IR1)}
\]

\[
E[s_1 + s_2|a_{1H}, a_{2H}, \epsilon, \theta] - a_{1H} - a_{2H} \geq E[s_1 + s_2 - a_1|a_{1i}, a_{2j}, \epsilon, \theta] - a_{1i}
\]

\[
\text{for all } \theta, \text{ and } i, j, k \in \{L, H\}; \quad \text{(PL-IC1)}
\]

\[
s_1 = a_1 + \beta_1 \times \max\{0, y_1\} \geq 0 \quad \text{for all } y_1; \quad \text{(PL-NN2)}
\]

where \( s_2 \) is the solution to Program PL—Period Two.
Program PL—Period Two

\[
\min_{a_2(\theta, y_1), \beta_2(\theta, y_1)} E[s_2|a_{2H}; \epsilon, \theta, y_1, a_{1H}]
\]

Subject to

\[
E[s_2|a_{2H}; \epsilon, \theta, y_1, a_{1H}] - a_{2H} \geq U_2 \text{ for all } \theta \text{ and } y_1; \quad \text{(PL-IR2)}
\]

\[
E[s_2|a_{2H}; \epsilon, \theta, y_1, a_{1H}] - a_{2H} \geq E[s_2|a_{2L}; \epsilon, \theta, y_1, a_{1H}] - a_{2L} \text{ for all } \theta \text{ and } y_1; \quad \text{(PL-IC2)}
\]

\[
s_2 = a_2(\theta, y_1) + \beta_2(\theta, y_1) \times \max\{0, y_2\} \geq 0 \text{ for all } \theta \text{ and } y_1, y_2. \quad \text{(PL-NN2)}
\]

Learning the agent’s type at the end of the first period mitigates the incentive problem in the second period, as the principal is able to fine-tune the second-period bonus rate to the observed agent’s type. Because the first-period agency problem is more severe (both hidden information and hidden action), the principal has to offer a “high” bonus rate to ensure the agent works regardless of his type. A conservative system reduces the probability of “overpayment” in the first period and reduces the incentive cost over two periods. As we show next, a conservative system dominates not only an aggressive system but also an unbiased system.

**Proposition 3.** In the learning model:

(i) Under a conservative system, the optimal bonus rates are: \(\beta_1 = \frac{a_H}{p_b - q_b}\) and \(\beta_2(\theta, 0) = \beta_2(\theta, 1) = \frac{a_H}{p_\theta - q_\theta}\) for any \(\theta\). The total expected incentive cost decreases in the level of downward bias and is lower than under the unbiased system.

(ii) Under an aggressive system, the optimal bonus rates are: \(\beta_1 = \frac{[p_b - q_b \epsilon_a]a_H}{p_b(p_b - q_b)(1 - \epsilon_a)}\),

\(\beta_2(\theta, 0) = \frac{a_H}{p_\theta - q_\theta}\), and \(\beta_2(\theta, 1) = \frac{[p_\theta + (1 - p_\theta) \epsilon_a]a_H}{p_\theta(p_\theta - q_\theta)}\) for any \(\theta\). The total expected
incentive cost increases in the level of upward bias and is higher than under the unbiased system.

Proposition 3 is our first finding in which downward bias improves on unbiased accounting. From the principal’s point of view, by introducing a downward bias in early periods that is reversed in later periods, conservatism allows her to wait to resolve the information asymmetry about the agent’s type before fully compensating the agent and, thus, improves contracting efficiency. To elaborate, a feature of conservatism in the learning model is that it allows for a fine-tuning not only of the second-period incentives but also of the leftover first-period incentives in play created by the bias reversal (the incentive spillback). Under unbiased or aggressive accounting, learning does not facilitate such a fine-tuning of the first-period incentives.

Risk Aversion

Three of the main results from our analysis of bias reversals with a risk neutral agent subject to limited liability constraints are: (i) a downward bias (and only a downward bias) can replicate the performance of unbiased accounting in our base model in which the same agent is retained in both periods; (ii) when the principal makes performance-contingent firing decision, a biased system is strictly dominated by unbiased accounting; and (iii) when the principal learns about the agent’s productivity over time (the learning model), a downward bias is strictly preferred to unbiased accounting.

If we introduce agent risk aversion with domain additive constant absolute risk aversion (CARA) preferences and no limited liability, either biased system (downward bias and upward
bias) can replicate the performance of the unbiased system in the base model (a revised version of Proposition 1). These preferences ensure the optimal long-term contract is memory-less, i.e., $s(j, k) = s(j) + s(k)$, (Fellingham, Newman, Suh 1985). For example, under a conservative bias, if first-period performance of 0 is followed by second-period performance of 2, the agent’s pay will be $\alpha$ in the first period and $\alpha + 2\beta$ in the second period. Suppose unbiased performance is instead 1 in the first period followed by 1 in the second period. The agent would receive the same total pay under unbiased performance measurement and the optimal long-term contract, or $\alpha + \beta$ in each period. CARA preferences also allow the principal to shift a fixed payment (salary) from the first to the second period to avoid providing the agent with incentives to “take the money and run” under an aggressive system.\footnote{With additively separable preference, shifting salaries from one period to another period would be costly to the principal.}

Proposition 2 relies on the agent’s retention rent being used to reduce the cost of providing first-period incentives. Under agent risk aversion with unlimited liability, the agent would be held to his reservation utility, eliminating any potential incentive spillback (created by a retention rent). In the learning model, our findings would remain qualitatively unchanged—a downward bias is strictly preferred to unbiased accounting or aggressive accounting (Proposition 3). That is, learning rather than the agent’s preferences or bankruptcy constraints is the driving force in the learning model.

### III. CONSERVATISM AS TIMELY RECOGNITION OF BAD NEWS

In Section II, we studied conservatism as measurement bias followed by a bias reversal (a “true-up” to actual cash flows). In that setting, a conservative bias has a multiplicative effect on
the probabilities of reporting good performance in the first period, and, consequently, the likelihood ratio associated with good performance remains intact. In this section, we study conservatism as an additive shock to the probabilities.

Suppose there are both good and bad shocks to the environment that are exogenous to the model (e.g., regulatory changes, price level changes due to changes in competition, or market demand changes due to advances in technology). The question we intend to study is how to best recognize these shocks through the choice of an accounting system: bad shocks early and good shocks late (a conservative system), good shocks early and bad shocks late (an aggressive system), or both good and bad shocks bundled in early or later periods. In contrast to the previous section, the shocks (news) recognized by the accounting systems are not related to the agent’s productive inputs. One takeaway from the comparison of these two settings is that conservatism viewed as timely recognition of bad news leads to a demand for conservatism relatively easily (e.g., without relying on learning) because the shocks shift the agent’s rents intertemporally, which has desirable incentive (spillback) effects. In contrast, the role of conservatism as bias followed by bias reversals can at best replicate unbiased accounting by relying on a repeated agency encounter (until we introduced learning).

Initially, the principal and the agent know that there are possible good news and bad news shocks over the two periods but do not directly observe these shocks. Instead, they observe accounting performance, which, by design, determines when these shocks are incorporated into performance. The principal chooses the accounting system at the beginning of the first period. The sequence of the events is the same as presented in Figure 1 (the base model). The principal again solves for the optimal contract in each period by minimizing the expected incentive cost subject to the individual rationality constraints at the beginning of each period, the incentive
compatibility constraints that ensure the agent works in each period, and the non-negativity constraints.

Denote good news shock by the parameter \( \varepsilon_G > 0 \) and bad news shock by the parameter \( \varepsilon_B < 0 \). Hence, the choice of an accounting system is essentially a choice between an increasing or a decreasing pattern of expected performance. Under a conservative system, bad news is recognized early and good news is recognized late. Specifically,

\[
Pr(y_1 = 1|\varepsilon_B, a_1) = Pr(x_1 = 1|a_1) + \varepsilon_B \quad \text{and} \quad Pr(y_2 = 1|\varepsilon_G, a_2) = Pr(x_2 = 1|a_2) + \varepsilon_G, \tag{1}
\]

\[
Pr(y_t = 1|\varepsilon_G, \varepsilon_B, a_t) = Pr(x_t = 1|a_t) + \varepsilon_G + \varepsilon_B. \tag{2}
\]

where \( \varepsilon_G \leq d, \varepsilon_B \geq -d, \) and \( d = \min\{p, q, 1 - p, 1 - q\} \).

A conservative system generates an increasing pattern of expected performance. Similarly, under an aggressive system, good news is recognized early and bad news is recognized late and, therefore, expected performance is decreasing over time. \( \varepsilon_G \) and \( \varepsilon_B \) are independent and should not be interpreted as reversals of each other. A choice to recognize bad news early increases future expected performance only because the (real) news is shifted from one period to another. The additive shock captures a potential shift in the agent’s rents that does not depend on the agent’s effort. This modeling choice is made, in part, as a contrast to the earlier bias setting in which conservatism enters multiplicatively.

Another alternative is to bundle good news and bad news shocks and recognize them in the same period, so any expected effect of bad (good) news is partially offset by the expected effect of good (bad) news. If both shocks are recognized in period \( t \), we write

\[
Pr(y_t = 1|\varepsilon_G, \varepsilon_B, a_t) = Pr(x_t = 1|a_t) + \varepsilon_G + \varepsilon_B. \tag{3}
\]

In the special case of \( \varepsilon_G = -\varepsilon_B \), expected good and bad news shocks offset each other in one (reporting) period. In the other period, there would be no news shocks to recognize; such a
reporting system is news “neutral,” which is roughly the equivalent of an unbiased system in our earlier model. To make this design meaningful, we assume \( \varepsilon_G \leq d \) and \( \varepsilon_B \geq -d \), where \( d \) should satisfy \( d = \min\{p, q, 1 - p, 1 - q\} \). It is straightforward to extend the analysis to allow for good and bad news to be recognized in both periods (having some of \( \varepsilon_B \) and some of \( \varepsilon_G \) recognized in each period), giving either good or bad news a greater timeliness than the other.

In the single period version of this model, conservatism is optimal for the same reason as it is in Kwon, Newman, and Suh (2001)—conservatism results in the best (lowest) possible likelihood ratio associated with high performance \( (q + \varepsilon_B)/(p + \varepsilon_B) \). The likelihood ratio associated with low performance is irrelevant, because of the binding non-negativity constraint.

The principal always motivates high inputs in both periods. Since good and bad news shocks do not depend on the agent’s effort level, the bonus rates in equilibrium are independent of the accounting system. However, accounting affects the expected incentive cost (the probability a bonus is paid) in each period. As we show below, the principal’s total incentive cost over the two periods does not depend on the accounting system (without performance-contingent firing decisions) but depends only on the aggregate level of the shocks. Any increase or decrease in the expected incentive cost in the first period induced by accounting is exactly offset by a corresponding change in the expected incentive cost in the second period. Hence, the choice of accounting system is irrelevant in the baseline two-period model.

**Proposition 4.** In the timely news recognition model with the same agent retained in both periods:

(i) For any accounting system choice, the optimal bonus rates are \( \beta_1 = \beta_2(0) = \beta_2(1) = \frac{a_H}{p-q} \). The expected incentive cost is \( \frac{(2p+\varepsilon_B+\varepsilon_G)a_H}{p-q} \).
(ii) The choice of an accounting system is irrelevant.

If the levels of good and bad news shocks are equal ($\varepsilon_G = -\varepsilon_B$), the expected incentive cost will be $\frac{2pa_H}{p-q}$. Proposition 4 is a special case in that accounting choice can matter when retention rents from keeping his job in later periods motivate the agent in early periods (an incentive spillback), as in Proposition 2. Again, suppose the principal has the option to fire the agent at the end of the first period. Our findings on performance-contingent firing are summarized in Proposition 5.

**Proposition 5.** In the timely news recognition model with the option to fire the agent at the end of the first period:

(i) For any accounting system, the principal optimally hires a new agent if the first-period report is low and optimally retains the incumbent agent if the first-period report is high.

(ii) A conservative system that recognizes the bad news shock in the first period and the good news shock in the second period always dominates an aggressive system that recognizes the good news shock in the first period and the bad news shock in the second period.

(iii) Suppose $\varepsilon_G + \varepsilon_B < 0$. If $-\varepsilon_B < p - q$, then a conservative system is optimal. Otherwise, recognizing the bundled news shocks in the first period is optimal.

(iv) Suppose $\varepsilon_G + \varepsilon_B \geq 0$. If $\varepsilon_G < p - q$, then a conservative system is optimal. Otherwise, recognizing the bundled news shocks in the second period is optimal.
Unlike the bias model in which the principal’s optimal firing decision depends on the accounting system (Proposition 2 of Section II), the principal’s optimal firing decision here does not depend on the accounting system. Since there is no reversal, lowering the first-period expected performance is not costly. Given any accounting system, when the principal can replace the agent, she takes advantage of the incentive spillback created by retention rents in the second period to reduce the cost of motivating the agent to work in the first period. Firing the agent is optimal only when a low performance report is observed at the end of the first period.

Proposition 5 (ii) states that an aggressive system that recognizes good news shocks early is never optimal. Proposition 5 (iii) and (iv) state that whether it is optimal to recognize bad news separately from good news or instead bundle it with good news depends on the magnitude of the news. For a single fixed shock ($\varepsilon_B = 0$ or $\varepsilon_G = 0$), a bad (good) news shock is always recognized in the first (second) period and, thus, a conservative system is optimal. For two shocks ($\varepsilon_B < 0$ and $\varepsilon_G > 0$), an “interior” solution may arise in equilibrium in that news is bundled together. When the expected news is good ($\varepsilon_G + \varepsilon_B \geq 0$), the interior case arises when the good news shock is large (and the bad news shock is sufficiently small)—in this case, all news is recognized in the second period. When the expected news is bad ($\varepsilon_G + \varepsilon_B < 0$), the interior case arises when the bad news shock is large—in this case, all news is recognized in the first period. Both interior solutions can be viewed as generalized conservatism—aggregate bad (good) news is recognized in the first (second) period. So, conservatism is pervasive as a solution in this setting.

To understand the accounting system choice, shifting either good or bad news shocks from one period to the other affects the probability the second-period bonus is paid out but not the second-period bonus itself which is fixed at $\frac{aH}{p-q}$. The effect on the first period is subtler,
affecting both the probability the first-period bonus is paid out and also the first-period bonus itself. Put differently, there is a linear effect on the cost of providing second-period incentives and a non-linear effect on the cost of providing first-period incentives.

There is an important difference between the models in Section II and this section that helps explain the contrast between Propositions 2 and 5. A conservative bias reversal provides the agent with rents in the second period but only along the path in which the first-period report is low, while a system that recognizes bad news early and good news late creates extra rents for the agent in the second period along all paths.

In our earlier bias reversal model, learning and conservative accounting were optimally combined by the principal to effectively fine-tune part of the agent’s first-period compensation. This is not possible in the news model, since there is no bias to reverse. Nevertheless, it is easy to show that learning would also create a demand for conservatism in the news model. Recognizing bad news early and good news late decreases (increases) the agent’s rent in the first (second) period. Learning decreases the cost of providing incentives in the second period; conservatism maximizes the principal’s ability to extract rents from the agent by learning.10

**Risk Aversion**

Consider the impact of introducing agent risk aversion with no limited liability in the news model, assuming the news shocks are equal-sized. Unbiased accounting aggregates the news shocks so that they offset each other. Without limited liability, the likelihood ratios associated with good performance and bad performance both matter—the optimal accounting system depends on the effectiveness of rewards and penalties. As a result, recognizing good

---

10 A formal analysis of this model is available upon request.
news and bad news shocks in different periods is optimal and dominates (news) neutral accounting (in contrast to Proposition 4’s bias irrelevance result). This parallels Kwon, Newman, and Suh (2001), which is essentially a single-period version of Proposition 4, where introducing agent risk aversion and unlimited liability alters the principal’s ranking of accounting systems. Proposition 5 relies on the incentive spillback in a more critical way than Proposition 2. Since an incentive spillback would not arise under agent risk aversion in the news model, the demand for bias developed in Proposition 5 would not arise under agent risk aversion.

IV. MULTI-TASK SETTING: OVERLAPPING PROJECTS

In this section, we extend the news model from Section III to include multiple projects, each to be implemented over two periods with the first (second) project begun in the first (second) period. Our focus is on the overlapping projects in the second period.

Suppose the agent is hired to implement two identical long-term projects over three periods. To avoid confusion between projects and periods, denote the project begun in the first period as Project $E$ and the project begun in the second period as Project $F$. For each project $j \in \{E,F\}$, the agent sequentially exerts two productive inputs to implement the project, denoted by $a_{jt} \in \{a_{jLt}, a_{jLtH}\}$, for $j = E, t = 1, 2$; and $j = F, t = 2, 3$. The agent’s cost of effort is denoted by $a_{jLt} = a_L$ and $a_{jLtH} = a_H$. In the second period, the agent’s cost of effort is additive and written as $a_{E2} + a_{F2}$. Each project produces an output, high ($x_{jt} = 1$) or low ($x_{jt} = 0$), in each period of its lifetime. The production technology that may differ across periods is identical for both projects and for each project independent across the two periods. We describe the technology for period $t$ by $p_t \equiv Pr(x_t = 1|a_{jLtH})$ and $q_t \equiv Pr(x_t = 1|a_{jLtL})$ for any project.
The contract between the principal and the agent depends on the aggregate accounting report in each period: \( y_1 \in \{0, 1\} \) for the first period, \( y_2 \in \{0, 1, 2\} \) for the second period, and \( y_3 \in \{0, 1\} \) for the third period. This is meant to capture a more realistic assumption that managerial contracts are written on aggregate performance measures and also to introduce task interdependency—to create a multi-task problem in the spirit of Demski (1994, 573-79), whose model is close to ours, and Holmstrom and Milgrom (1991). The reporting technology is described below, where \( \epsilon_{E1} \) and \( \epsilon_{E2} \) are the news shocks recognized for project \( E \) in the first and the second period, respectively, and \( \epsilon_{F2} \) and \( \epsilon_{F3} \) are the news shocks recognized for project \( F \) in the second and the third period, respectively.

\[
P_r(y_1 = 1|a_{E1}) = Pr(x_1 = 1|a_{E1}) + \epsilon_{E1}; \quad (4)
\]

\[
P_r(y_2 = y_{E2} + y_{F2}|a_{E2}, a_{F2}) = [Pr(x_{E2}|a_{E2}) + \epsilon_{E2}][Pr(x_{F2}|a_{F2}) + \epsilon_{F2}]; \text{ and} \quad (5)
\]

\[
P_r(y_3 = 1|a_{F3}) = Pr(x_3 = 1|a_{F3}) + \epsilon_{F3}. \quad (6)
\]

The setup here is close to the news model in Section III with two exceptions. First, instead of linear contracts, we solve for optimal contracts. This is not a significant change from Section III in that linear contracts are optimal in the news model but not optimal in the multi-task setting. Second, to demonstrate that a multi-task setting can make an interior accounting system optimal, we allow for the continuous choice of \( \epsilon_{E1} = \epsilon_{F2} = \epsilon \in [-d, +d] \) for \( d > 0 \). The news shock recognized in the subsequent period is \( \epsilon_{E2} = \epsilon_{F3} = -\epsilon \). Allowing for a continuous choice of \( \epsilon \) can be thought of as bundling news. For example, if all good and bad news is bundled into a single period, then the system is a neutral one with no aggregate shock. The principal might also partially bundle—combine some good news with all of the bad news or vice versa and recognize the unbundled component in the next period (for \( |\epsilon| < d \)). Lastly, the principal might not bundle
the news at all so that good news shock is recognized in one period and bad news shock is
recognized in another (for \( \varepsilon = d \) or \( \varepsilon = -d \)).

Because high inputs are always motivated, there is no benefit to long-term contracting.
That is, the optimal long-term contract can be represented as three short-term contracts:
\[ s(y_1, y_2, y_3) = s(y_1) + s(y_2) + s(y_3). \]
The agent’s overall payoff is: 
\[ s(y_1) + s(y_2) + s(y_3) - a_{E_1} - a_{E_2} - a_{F_2} - a_{F_3}. \]
The sequence of the events is presented in Figure 4.

**Figure 4: Time-line of the Model with Overlapping Projects**

<table>
<thead>
<tr>
<th>t=0</th>
<th>t=1</th>
<th>t=2</th>
<th>t=3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounting system with bias ( \varepsilon ) is in place. Principal and agent sign a one-period contract.</td>
<td>Agent supplies ( a_{E_1} ) to manage Project ( E ). ( y_1 ) is reported.</td>
<td>Agent is paid ( s_1(y_1) ). Principal and agent sign a one-period contract.</td>
<td>Agent supplies ( a_{E_2}(y_1) ) and ( a_{F_2}(y_1) ) to manage Projects ( E ) and ( F ). ( y_2 ) is reported.</td>
</tr>
</tbody>
</table>

The principal’s problem is to minimize the expected incentive cost subject to the
following constraints. The agent’s individual rationality and incentive compatibility constraints
are as in the news model (of Section III), except that a third period is added and the second
period has the agent choosing how much effort to devote to each project. To study the effect of
multiple tasks, we focus on the cost of providing incentives in the second period (the only period
with multiple tasks). This can also be thought of as a choice to focus on the steady state. In the
case in which the projects are productively identical, solving the second-period problem is the
same as solving the overall three-period problem.
Overlapping Identical Projects

We first consider a case in which the project has identical technology across periods, i.e., \( p_1 = p_2 = p_3 = p \) and \( q_1 = q_2 = q_3 = q \). The productive homogeneity of the projects makes them well suited for grouping together and assigning to a single agent (Holmstrom and Milgrom 1991). However, the accounting system can introduce heterogeneity, since it affects old and new projects in opposite directions. When, if at all, is it optimal to introduce heterogeneity via the accounting system? Proposition 6 answers this question.

**Proposition 6.** Assume \( d > \frac{p-q}{2} \). In the overlapping projects model with identical two-period projects:

(i) The conservative and aggressive accounting systems are equivalent.

(ii) The optimal accounting system sets \( \epsilon^* = \pm \frac{p-q}{2} \).

Proposition 6 (i) follows directly from the productive homogeneity of the two projects. Part (ii) tells us that the principal optimally chooses to partially separate good and bad news shocks (partial bundling). The reason that partial bundling is optimal is that, from the starting point of neutral accounting \( (\epsilon = 0) \), partial bundling initially reduces the probability that a bonus payment is made without impacting the size of the required bonus payment. To be more precise, the agent’s marginal productivity is a function of \( (p + \epsilon)(p - \epsilon) - (q + \epsilon)(q - \epsilon) = p^2 - q^2 \), which determines the size of the required bonus and does not depend on how news is recognized. The probability that the bonus is paid out is \( (p - \epsilon)(p + \epsilon) \), which is maximized when the news shocks are bundled (so that \( \epsilon = 0 \)) and is decreasing in \( |\epsilon| \). Hence, as long as the aggregate
shock in the bundled news is not so large that it changes which incentive constraints bind, increasing $|\varepsilon|$ is optimal.

The optimal magnitude of the aggregate shock in the bundled news ($|\varepsilon|$) is determined by the underlying incentive problem. If $|\varepsilon|$ is less than $(p - q)/2$, the binding constraint has the agent working on both projects rather than shirking on both projects (neither project is a bottleneck). If $|\varepsilon|$ is greater than $(p - q)/2$, the binding constraint has the agent working on both projects rather than shirking on only one of them (the bottleneck project). For example, if $\varepsilon > (p - q)/2$, the aggregate shock in the bundled news is positive and therefore Project F is the bottleneck (the agent more likely shirks on implementing Project F). The problem becomes essentially a single-task problem whose cost is minimized by reducing the likelihood ratio $(q + \varepsilon)/(p + \varepsilon)$. The way to make the likelihood ratio and the expected cost of incentives small is to reduce $\varepsilon$ (as the likelihood ratio increases in $\varepsilon$). So, the optimal choice of bundling the news shock is to choose $\varepsilon$ just large enough so that we do not make one of the two projects a bottleneck, or $\varepsilon = (p - q)/2$.

As we discussed in the previous section, the optimal accounting system is the maximally conservative one in the single-period/single-task version of this model (as in Kwon, Newman, and Suh 2001). However, in Proposition 6 with multiple overlapping tasks, an interior accounting system can be optimal. The driving force in Proposition 6 is that an extreme accounting system (a corner solution of $\varepsilon = \pm d$) changes the nature of the incentive problem (which incentive constraint binds), even though the two projects are ex ante identical. In other words, extreme accounting (conservative or aggressive accounting) creates an incentive problem similar to the one created by grouping easy and difficult tasks together, as studied in Holmstrom and Milgrom (1991). The problem is created by the realistic assumption that there will be
overlapping projects—some new and some old—and that accounting system affects them in opposite ways.

**Overlapping Heterogeneous Projects**

We next study the case of projects that are heterogeneous to begin with. This could be referred to as heterogeneity in the projects themselves or an intertemporal heterogeneity that makes old and new projects differ, which is the approach we take. In particular, each project’s first-period productivity is different from the second-period productivity, denoted by $p_2 > p_1$. For simplicity, we maintain the assumption that $q_1 = q_2 = q$. A natural interpretation is that the agent learns to be more efficient in implementing the project over time.

We continue to analyze the cost of providing incentives in the second period to focus on the steady state and the impact of overlapping projects. However, unlike the case with homogeneous projects, the effects of accounting on the first and third periods are no longer offsetting. The following proposition characterizes the optimal accounting system that minimizes the expected incentive costs over three periods.

**Proposition 7.** Assume $d > \frac{p_1 p_2 + q^2 - 2p_1 q}{p_1 + p_2 - 2q}$, $p_2 > p_1$, and $\frac{p_2 - p_1}{p_1 - q} < 1$. The optimal accounting system is conservative with $\epsilon^* = -\frac{p_1 p_2 + q^2 - 2p_1 q}{p_1 + p_2 - 2q} < 0$.

Proposition 7 differs from Proposition 6 in that the optimal accounting system is uniquely determined with overlapping heterogeneous projects. The assumption $\frac{p_2 - p_1}{p_1 - q} < 1$ ensures that the
marginal effect of the choice of \(\varepsilon\) is small so that its choice is primarily determined by its effect on the second-period expected pay.

Similar to Proposition 6, it is optimal to introduce as extreme of a system as we can without making one of the projects a bottleneck—to use accounting to introduce as much heterogeneity as possible without going beyond the switching point at which one of the projects becomes a bottleneck. Since \(p_2 > p_1\), the optimal accounting system brings us to the switching point when the old project, Project \(E\), just starts to become a bottleneck.

If the boundary condition imposed by \(d\) does not hold, then either a conservative system or an aggressive system may be optimal depending on the underlying agency problem for each project and the magnitude of \(d\). For example, if \(p_2\) is sufficiently large so that initially Project \(F\) is the bottleneck under unbiased accounting and the size of \(d\) is not too constraining, then a conservative system that recognizes a negative aggregate shock in the first period is optimal to remove the bottleneck by making the projects more homogenous after accounting measurement. The accounting system uses a judicious (partial) bundling of good and bad news to create just the right level of the heterogeneity/homogeneity.

**V. CONCLUDING REMARKS**

Much of the criticism directed toward accounting conservatism is multi-period in nature—that understating performance today means that some future period’s performance will be overstated. Yet, the information economics-based theoretical literature on accounting conservatism has concentrated on single-period models. We study two models of conservatism: (i) bias with bias reversals and (ii) timely bad news recognition. Without learning, the best that bias reversals can do is to replicate the performance of unbiased accounting. Without
performance-contingent firing, downward bias is preferred to upward bias because the subsequent bias reversal allows the principal to correct early underpayments but not to fully correct early overpayments. With performance-contingent firing, upward bias is sometimes preferred to downward bias because performance-contingent agent replacement eliminates conservatism’s ability to correct early bonus underpayments. Learning creates a demand for downward bias even over unbiased accounting. Learning allows the principal to effectively fine-tune part of the first-period bonus with the benefit of hindsight (learning). In this case, conservatism’s comingling of the first- and second-period performance is beneficial.

When conservatism is viewed as early bad news recognition (late good news recognition), performance-contingent firing creates a strict preference for conservatism over aggressive or unbiased accounting. Conservatism shifts rents from the first to the second period, which reduces the cost of providing first-period incentives. Since this is model of (effort-independent) news timing rather than bias, there is no role for retaining the agent to correct a first-period bonus underpayment. Hence, the cost of conservatism that performance-contingent firing introduces with bias reversals is not present in the timely news model. We also studied an extension of the news model with multiple overlapping projects, a multi-task problem. If old and new projects are sufficiently homogenous that neither of them is a bottleneck when news shocks are bundled and recognized in the same period, then maximally conservative or maximally aggressive accounting can make one of the projects a costly bottleneck. Hence, a non-neutral interior accounting system can be optimal.

Throughout our analysis, we assumed the principal and agent have the same information about performance. Suppose the agent observes the underlying (unbiased) performance. The agent would “take the money and run” when performance is overstated under an aggressive bias,
while an understatement under a conservative bias would only strengthen the agent’s resolve to remain in the firm’s employ. Accounting conservatism seems to have a broader role in fostering long-term relationships. Relatedly, new insights about the interaction between conservatism and earnings management seem likely to arise in dynamic models (e.g., a multi-period version of Bertomeu, Darrough, and Xue, 2017). When might conservatism lead to less manipulation?

Recent theoretical research on accounting conservatism, including our paper, seems to be overly focused on conservatism as bias. Consider perhaps the most often discussed example of accounting conservatism: lower of cost or market. Writing the asset down but not up to reflect market prices can be thought of as devoting additional attention to bad states. This additional attention comes in the form of resources devoted to detecting these bad states and the granularity of disclosure. When the market price is higher than cost, we are told no more. When the market price is lower than cost, we are given more information (or can recover it)—we know both the historical cost and the write-down. In this sense, traditional historical cost accounting subject to write-downs is honed on bad news, since the communication channels are expanded when there is bad news (similar to the treatment of a sick patient by his doctor).11 Instead, recent models of accounting conservatism mostly focus on the information content of a single aggregate performance measure such as net income. The broader question is what is the comparative advantage of accounting over other information sources, and does that comparative advantage involve conservatism? Public financial statements are themselves developed sequentially by multiple economic agents who each have their own information, economic incentives, and expertise (for example, in making various one-sided corrections, as in Arya and Glover 2008).

APPENDIX

Proof of Lemma 1

Positive bonus rates, non-negative performance measures, and non-negativity constraints ensure the optimal salaries (intercepts) are zero, \( \alpha_1 = \alpha_2 (\cdot) = 0 \). This observation applies to all the programs in the appendix.

We solve for the optimal bonus rates: \( \beta_1 \), \( \beta_2 (0) \), and \( \beta_2 (1) \). In the second period, for each \( y_1 \), there is a single incentive constraint (IC2), which holds as in equality. The second-period contract varies with \( y_1 \) but not with the agent to whom the contract is offered. If \( y_1 = 1 \) is reported, then the first-period output must be \( x_1 = 1 \) so that there is no reversal. The optimal bonus rate that motivates high input in the second period is \( \beta_2 (1) = \frac{a_H}{p-q} \). If \( y_1 = 0 \) is reported, then the first-period output can be either \( x_1 = 1 \) or \( x_1 = 0 \) so that a reversal would occur with probability \( \frac{pe_c}{pe_c+1-p} \). If a reversal occurs, the second-period report is either \( y_2 = 2 \) or \( y_2 = 1 \).

\[
\frac{pe_c}{pe_c + 1 - p} [p\beta_2 (0) \times 2 + (1-p)\beta_2 (0) \times 1] + \frac{1-p}{pe_c + 1 - p} [p\beta_2 (0) \times 1] - a_H
\]

\[
= \frac{pe_c}{pe_c + 1 - p} [q\beta_2 (0) \times 2 + (1-q)\beta_2 (0) \times 1] + \frac{1-p}{pe_c + 1 - p} [q\beta_2 (0) \times 1] \]

\[\Rightarrow \beta_2 (0) = \frac{a_H}{p-q} \quad (A1)\]

The first-period contract varies depending on whether the first-period agent is employed (retained) in the second period. Suppose the first-period agent is retained in the second period. In this case, there are seven (IC1) constraints. The first-period bonus rate is determined as the smallest \( \beta_1 \) that satisfies all seven (IC1) constraints. For a conservative system, the binding constraint is (IC1) – \( (a_{1L}, a_{2H}, a_{2H}) \), i.e., the constraint that ensures the agent prefers working in both periods to shirking in the first period and working in the second period. The first-period bonus rate satisfies:

\[
p (1 - \epsilon_c) [\beta_1 \times 1 + p\beta_2 (1) \times 1] + p\epsilon_c [p\beta_2 (0) \times 2 + (1-p)\beta_2 (0) \times 1] + (1-p)[p\beta_2 (0) \times 1] - 2a_H = q (1 - \epsilon_c) [\beta_1 \times 1 + p\beta_2 (1) \times 1] + q\epsilon_c [p\beta_2 (0) \times 2 + (1-p)\beta_2 (0) \times 1] + (1-q)[p\beta_2 (0) \times 1] - a_H; \]

\[\Rightarrow (1 - \epsilon_c) [\beta_1 + p\beta_2 (1)] + \epsilon_c [(1 + p)\beta_2 (0)] - [p\beta_2 (0)] = \frac{a_H}{p-q} \quad (A2)\]
Plugging $\beta_2 (\cdot) = \frac{a_H}{p-q}$ in (A2) yields $\beta_1 = \frac{a_H}{p-q}$. The expected incentive cost over two periods is determined as $E S^c = \frac{2pa_H}{p-q}$, which is independent of $\varepsilon_c$.

Suppose a new agent is hired at the beginning of the second period. The first-period bonus rate satisfies the single (IC1) constraint as an equality: $\beta_1 = \frac{a_H}{(p-q)(1-\varepsilon_c)}$. The expected cost is determined as $E S^c = \frac{(2+\varepsilon_c)pa_H}{p-q}$, which increases in $\varepsilon_c$ ($\frac{dE S^c}{d\varepsilon_c} = \frac{pa_H}{p-q} > 0$).

Hiring a new agent in the second period is costlier because downward bias reversal increases the likelihood of high report in the second period to the extent that the new agent earns additional rent, written as $(p\varepsilon_c + 1 - p) \left[ \frac{p\varepsilon_c}{p\varepsilon_c + 1 - p} \right] \frac{a_H}{p-q} = \frac{p\varepsilon_c a_H}{p-q}$. ■

**Proof of Lemma 2**

Parallel to the proof of Lemma 1, we first solve the second-period contract. If $y_1 = 0$ is reported, then there is no reversal and the optimal bonus rate is $\beta_2 (0) = \frac{a_H}{p-q}$. If $y_1 = 1$ is reported, a reversal occurs with probability $\frac{(1-p)e_a}{p+(1-p)e_a}$. With a reversal, the second-period report is $y_2 = 0$ or $y_2 = -1$, and the agent is paid zero. The optimal rate that satisfies the constraint (IC2) is $\beta_2 (1) = \frac{[p+(1-p)e_a]a_H}{p(p-q)}$.

Suppose the same agent is employed in both periods. The binding (IC1) constraint is $(IC1 - (a_{1L}, a_{2H}, a_{2L}))$, i.e., the constraint that ensures the agent prefers working in both periods to shirking in the first period and shirking in the second period only when a high report ($y_1 = 1$) is observed. The first-period bonus rate satisfies:

$$p[\beta_1 + p\beta_2 (1)] + (1 - p)e_a \beta_1 + (1 - p)(1 - e_a)p\beta_2 (0) - 2a_H$$

$$= q [\beta_1 + q\beta_2 (1)] + (1 - q)e_a \beta_1 + (1 - q)(1 - e_a)[p\beta_2 (0) - a_H]$$

$$\Rightarrow \beta_1 = \frac{[p-qe_a]a_H}{p(p-q)(1-e_a)}. \quad (A3)$$

The expected incentive cost is written as $E S^a = \left[ \frac{(p+(1-p)e_a)(p-qe_a)}{p(1-e_a)} + p \right] \frac{a_H}{p-q}$, which increases in $e_a$ ($\frac{dE S^a}{d\varepsilon_a} > 0$).
Suppose a new agent is employed in each period. The first-period bonus rate ensures that the agent prefers working to shirking in the first period and is determined as 

$$\beta_1 = \frac{a_H}{(p-q)(1-\varepsilon_a)}.$$ 

The expected incentive cost is 

$$ES^a = \left[\frac{p+(1-p)e_a}{1-\varepsilon_a} + p\right] \frac{a_H}{p-q},$$ 

which increases in 

$$\varepsilon_a \left(\frac{dES^a_{\varepsilon_a}}{d\varepsilon_a} > 0\right).$$

Comparing the incentive costs in parts (i) and (ii) shows that hiring a new agent in the second period is costlier. □

**Proof of Proposition 1**

The preferred system generates the lowest expected incentive cost. Part (i) is immediate from Lemma 1 (i) and Lemma 2 (i). To prove part (ii), the incentive costs in Lemma 1 (ii) and Lemma 2 (ii) both increase in the level of bias so that any biased system is dominated by unbiased system. Comparing these two incentive costs, if 

$$\varepsilon_c p > \frac{\varepsilon_a}{1-\varepsilon_a},$$

then an aggressive system dominates a conservative system. If 

$$\varepsilon_c p = \frac{\varepsilon_a}{1-\varepsilon_a},$$

then the two systems are equivalent. If 

$$\varepsilon_a \varepsilon_c,$$

then a conservative system always dominates an aggressive system. □

**Proof of Proposition 2**

The principal has four strategies: always fire the agent, fire the agent only when a low report is observed, fire the agent only when a high report is observed, and never fire the agent. We first solve for the optimal firing decision under each information system and then compare systems.

*Case 1: Suppose a conservative system is in place.*

Lemma 1 implies the principal prefers never firing to always firing. Firing the agent when a high report is observed requires higher bonus rate to motivate the agent to work in the first period. The second-period bonus rates offered to the incumbent agent and the new agent are not affected by the principal’s firing decision: 

$$\beta_{2N}() = \beta_2(\cdot) = \frac{a_H}{p-q}.$$ 

Never firing dominates firing the agent when a high report is observed. If the principal fires the agent only when a low report is observed, a higher first-period bonus rate again is required to motivate the agent to work:

$$p(1 - \varepsilon_c)[\beta_1 \times 1 + p\beta_2(1) \times 1 - a_H] - a_H = q(1 - \varepsilon_c)[\beta_1 \times 1 + p\beta_2(1) \times 1 - a_H]$$
The principal prefers never firing to firing only when a low report is observed if and only if \( \varepsilon_c \geq \frac{q}{1+q} \).

In sum, if \( \varepsilon_c < \frac{q}{1+q} \), the optimal strategy is to fire the agent only when a low report is observed. The principal’s expected incentive cost is written as

\[
ES^c(D(0) = 0, D(1) = 1) = \frac{p(2-q+\varepsilon_c+q\varepsilon_c)a_H}{(p-q)}.
\]

(A5)

Otherwise, if \( \varepsilon_c \geq \frac{q}{1+q} \), then the optimal strategy is to never fire. The principal’s expected incentive cost is written as

\[
ES^c(D(0) = 1, D(1) = 1) = \frac{2pa_H}{(p-q)}.
\]

(A6)

**Case 2: Suppose an aggressive system is in place.**

Lemma 2 implies the principal prefers never firing to always firing. If the principal fires the agent only when a high report is observed, a higher bonus rate is required to motivate the agent to work in the first period as

\[
\beta_1 = \frac{[1+q(1-\varepsilon_a)]a_H}{(p-q)(1-\varepsilon_a)} > \frac{[p-q\varepsilon_a]a_H}{p(p-q)(1-\varepsilon_a)}.
\]

The second-period bonus rates are \( \beta_{2N}(0) = \beta_2(0) = \frac{a_H}{p-q} \) and \( \beta_{2N}(1) = \beta_2(1) = \frac{[p+(1-p)\varepsilon_a]a_H}{p(p-q)} \), which are not affected by the principal’s firing decision. Firing when a high report is observed is costlier than never firing. If the principal fires the agent only when a low report is observed, the first-period bonus rate ensures that the agent prefers always working to always shirking:

\[
p[\beta_1\times1 + p\beta_2(1)\times1 - a_H] + (1-p)\varepsilon_a[\beta_1\times1 - a_H] - a_H
\]

\[
= q[\beta_1\times1 + q\beta_2(1)\times1] + (1-q)\varepsilon_a[\beta_1\times1];
\]

\[
\Rightarrow \beta_1 = \frac{p-q[p+(1-p)\varepsilon_a][a_H]}{(p-q)(1-\varepsilon_a)} < \frac{[p-q\varepsilon_a]a_H}{p(p-q)(1-\varepsilon_a)}.
\]

(A7)

The principal prefers firing only when a low report is observed to never firing. Consequently, the optimal strategy is to fire the agent only when a low report is observed. The principal’s expected cost is written as

\[
ES^a(D(0) = 0, D(1) = 1) = \frac{p+(1-p)\varepsilon_a}[p-q[p+(1-p)\varepsilon_a]]a_H + \frac{pa_H}{(p-q)}.
\]

(A8)

**Case 3: Suppose an unbiased system is in place.**

The optimal firing decision is determined by applying \( \varepsilon_c = 0 \) in Case 1 or \( \varepsilon_a = 0 \) in Case 2, and it is optimal to fire the agent only when a low report is observed.
To prove part (ii), if $\varepsilon_a = \varepsilon_c < \frac{q}{1+q}$ and $p + q > 1$, $(A8) < (A5)$, i.e., an aggressive system dominates a conservative system. If $\varepsilon_a = \varepsilon_c > \frac{pq}{(1-p)(1-q)}$, then $(A8) > (A6)$, i.e., a conservative system dominates an aggressive one.

To prove part (iii), because $(A8)$ is increasing in the level of $\varepsilon_a$, an unbiased system strictly dominates a system with upward bias. Similarly, because $(A5)$ is increasing in the level of $\varepsilon_c$ while $(A6)$ does not vary with the level of $\varepsilon_c$, an unbiased system (weakly) dominates a system with downward bias.

**Proof of Proposition 3**

Suppose a conservative system is in place. For any project $\theta$, the constraint (PL-IC2) binds so the second-period bonus rate is $\beta_2(\theta, y_1) = \frac{a_H}{p_\theta - q_\theta}$. The agent privately observes his own type before exerting effort. For any agent type $\theta$, the agent has incentive to work as long as $\beta_1 \geq \frac{a_H}{p_\theta - q_\theta}$. Because $p_g - q_g > p_b - q_b$, the optimal first-period bonus rate is $\beta_1 = \frac{a_H}{p_b - q_b}$. The expected incentive cost under a conservative system is written as:

$$E_S^c = [2Pr(b)p_b + Pr(g)p_g(1 - \varepsilon_c)]\frac{a_H}{p_b - q_b} + [Pr(g)p_g(1 + \varepsilon_c)]\frac{a_H}{p_g - q_g},$$

which is decreasing in $\varepsilon_c$ ($\frac{dE_S^c}{d\varepsilon_c} = [Pr(g)p_g]\left[\frac{a_H}{p_g - q_g} - \frac{a_H}{p_b - q_b}\right] < 0$). A conservative system dominates an unbiased system.

Suppose an aggressive system is in place. For any project $\theta$, if $y_1 = 0$ is reported, then there is no reversal and the second-period bonus rate is $\beta_2(\theta, 0) = \frac{a_H}{p_\theta - q_\theta}$. If $y_1 = 1$ is reported, then a reversal occurs in the second period with probability $\frac{(1-p_\theta)\varepsilon_a}{p_\theta + (1-p_\theta)\varepsilon_a}$. The second-period bonus rate is $\beta_2(\theta, 1) = \frac{[p_\theta + (1-p_\theta)\varepsilon_a]a_H}{p_\theta (p_b - q_b)}$. In the first period, the agent of either type is motivated to work and the bonus rate is determined as $\beta_1 = \frac{[p_b - q_b\varepsilon_a]a_H}{p_b (p_b - q_b)(1-\varepsilon_a)}$. The expected incentive cost under an aggressive system is written as:

$$E_S^a = \{Pr(g)\left[p_g + (1-p_g)\varepsilon_a\right] + Pr(b)\left[p_b + (1-p_b)\varepsilon_a\right]\}\frac{[p_b - q_b\varepsilon_a]a_H}{p_b (p_b - q_b)(1-\varepsilon_a)} +$$

$$Pr(g)\frac{p_ga_H}{(p_g - q_g)} + Pr(b)\frac{p_ba_H}{(p_b - q_b)}.$$ (A10)
Note that $\frac{d\beta_1}{d\varepsilon_a} > 0$ and $\frac{dE_S^a}{d\varepsilon_a} > 0$, so an aggressive system is dominated by an unbiased system.

Note that $E_S^c = E_S^a$ for $\varepsilon_c = \varepsilon_a = 0$. For any $\varepsilon_c > 0, \varepsilon_a > 0$, a conservative system dominates an aggressive system. ■

**Proof of Proposition 4**

Denote the first- and second-period news recognition by $\varepsilon_1$ and $\varepsilon_2$ respectively. For any $y_1$, the second-period incentive compatibility constraint binds and the bonus rate is $\beta_2(y_1) = \frac{a_H}{p-q}$. The first-period incentive compatibility constraint also binds and the bonus rate is $\beta_1 = \frac{a_H}{p-q}$. The expected incentive cost is $E_S = \frac{(2p+\varepsilon_1+\varepsilon_2)a_H}{p-q}$, which depends on the aggregate level of news shocks ($\varepsilon_1 + \varepsilon_2 = \varepsilon_B + \varepsilon_G$) but not on the timing of recognition. ■

**Proof of Proposition 5**

We first characterize the optimal firing decision for any given $\varepsilon_1$ and $\varepsilon_2$ and then solve for the optimal pair of $\varepsilon_1$ and $\varepsilon_2$.

Step 1: for any $\varepsilon_1$ and $\varepsilon_2$, determine the optimal firing decision.

The principal’s firing decision does not affect the second-period incentive problem as $\beta_{2N}(y_1) = \beta_2(y_1) = \frac{a_H}{p-q}$. If the principal always fires the agent, the first-period bonus rate is $\beta_1 = \frac{a_H}{p-q}$. Proposition 4 characterizes the solution in which the principal never fires the agent (that is, $\beta_1 = \frac{a_H}{p-q}$). The principal is indifferent between never firing and always firing the agent.

If the principal fires the agent only when a high report is observed, a higher bonus rate is required to motivate the agent to work in the first period, determined as:

$$
(p + \varepsilon_1)[\beta_1 \times 1] + (1 - p - \varepsilon_1)(p + \varepsilon_2)[\beta_2(0) \times 1] - a_H - (1 - p - \varepsilon_1)a_H
$$

$$
= (q + \varepsilon_1)[\beta_1 \times 1] + (1 - q - \varepsilon_1)(p + \varepsilon_2)[\beta_2(0) \times 1] - (1 - q - \varepsilon_1)a_H
$$

$$
\Rightarrow \beta_1 = \frac{(1+q+\varepsilon_2)a_H}{p-q} > \frac{a_H}{p-q}.
$$

(A11)

Firing the agent when a high report is observed is costlier than never firing. If the principal fires the agent only when a low report is observed, the incentive spillback created by the second-period rent reduces the first-period bonus rate,

$$
(p + \varepsilon_1)[\beta_1 \times 1] + (p + \varepsilon_1)(p + \varepsilon_2)[\beta_2(1) \times 1] - a_H - (p + \varepsilon_1)a_H
$$
\[
= (q + \varepsilon_1)[\beta_1 \times 1] + (q + \varepsilon_1)(p + \varepsilon_2)[\beta_2(1) \times 1] - (q + \varepsilon_1)a_H
\]

\[
\Rightarrow \beta_1 = \frac{(1-q-\varepsilon_2)a_H}{p-q} < \frac{a_H}{p-q}. \quad \text{(A12)}
\]

Firing the agent only when a low report is observed is less costly than never firing. For any given \(\varepsilon_1\) and \(\varepsilon_2\), the optimal firing decision is to fire the agent when a low report is observed. The expected incentive cost is:

\[
ES = (p + \varepsilon_1)\left[\frac{(1-q-\varepsilon_2)a_H}{p-q}\right] + (p + \varepsilon_2)\left[\frac{a_H}{p-q}\right]
\]

\[
= \frac{[p(2-q)+(1-q)\varepsilon_1+(1-p)\varepsilon_2-\varepsilon_1\varepsilon_2]a_H}{p-q}. \quad \text{(A13)}
\]

Step 2: solve for the optimal \(\varepsilon_1\) and \(\varepsilon_2\).

Consider four choices of \(\varepsilon_1 \in \{\varepsilon_B, \varepsilon_G, \varepsilon_B + \varepsilon_G, 0\}\) with \(\varepsilon_2 = \varepsilon_B + \varepsilon_G - \varepsilon_1\) by definition. The optimal choice of \(\varepsilon_1\) minimizes the expected incentive cost (A13). By design, a smaller \(\varepsilon_1\) produces a larger \(\varepsilon_2\), which means the probability of the agent being paid in the second period is higher and, thus, the second-period expected rent is higher (given the second-period bonus rate remains constant). A smaller \(\varepsilon_1\) has two effects on the first-period rent: (i) the probability of the agent being paid in the first period is lower, and (ii) the first-period bonus rate is smaller due to the incentive spillover created by retention rent. Both effects lead to a lower first-period expected rent. The principal chooses \(\varepsilon_1\) to trade off its positive effect on the first-period rent and its negative effect on the second-period rent.

Suppose the good news shock and the bad news shock are recognized in different periods so that the choice is between \(\varepsilon_1 = \varepsilon_B\) and \(\varepsilon_1 = \varepsilon_G\). The preferred choice is \(\varepsilon_1 = \varepsilon_B\). A conservative system dominates an aggressive system, since (A13) is smaller when \(\varepsilon_1 = \varepsilon_B\) than when \(\varepsilon_1 = \varepsilon_G\). This is because the effect of \(\varepsilon_1\) on the first-period rent is more pronounced than the effect on the second-period rent.

Suppose both good and bad news shocks are recognized in the same period so that the choice is between \(\varepsilon_1 = \varepsilon_B + \varepsilon_G\) and \(\varepsilon_1 = 0\). The preferred choice is \(\varepsilon_1 = \varepsilon_B + \varepsilon_G\) as long as \(\varepsilon_B + \varepsilon_G < 0\). A bundled news shock is always preferred to recognizing a good news shock by itself in the first period, since (A13) is less when \(\varepsilon_1 = \varepsilon_B + \varepsilon_G\) or \(\varepsilon_1 = 0\) than when \(\varepsilon_1 = \varepsilon_G\).

Suppose \(\varepsilon_B + \varepsilon_G < 0\). Compare the choices between \(\varepsilon_1 = \varepsilon_B\) and \(\varepsilon_1 = \varepsilon_B + \varepsilon_G\). The incentive cost (A13) is less when \(\varepsilon_1 = \varepsilon_B\) than when \(\varepsilon_1 = \varepsilon_B + \varepsilon_G\) as long as \(-\varepsilon_B < p - q\). That is, if \(-\varepsilon_B < p - q\), then \(\varepsilon_1 = \varepsilon_B\) is optimal; otherwise, \(\varepsilon_1 = \varepsilon_B + \varepsilon_G\) is optimal. Suppose \(\varepsilon_B + \varepsilon_G < 0\) and \(\varepsilon_B < 0\).
$\varepsilon_G \geq 0$. Then we compare the choices between $\varepsilon_1 = \varepsilon_B$ and $\varepsilon_1 = 0$. (A13) is less when $\varepsilon_1 = \varepsilon_B$ than when $\varepsilon_1 = 0$ as long as $\varepsilon_G < p - q$. That is, if $\varepsilon_G < p - q$, then $\varepsilon_1 = \varepsilon_B$ is optimal; otherwise, $\varepsilon_1 = 0$ is optimal. ■

**Proof of Proposition 6**

Since the two projects are identical, we drop the subscripts $E$ and $F$. The accounting choice is described by the parameter $\varepsilon$. The news shock recognition does not affect the first- and the third-period incentive pay: $s_1(1) = s_3(1) = \frac{a_H}{p-q}$. For any $\varepsilon$, the reduction (increase) in the first-period expected pay is exactly offset by an increase (reduction) in the third-period expected pay. That is, $(p + \varepsilon)\frac{a_H}{p-q} + (p - \varepsilon)\frac{a_H}{p-q} = \frac{2p a_H}{p-q}$. Only the second-period incentive problem is relevant to the choice of $\varepsilon$.

In the second period, the following constraints ensure the agent prefers working on both projects than working on only Project E (constraint (A14)), or working on only Project F (constraint (A15)), or not working on any project (constraint (A16)).

$$s_2(2) \geq \frac{a_H}{(p-\varepsilon)(p+\varepsilon)-(p-\varepsilon)(q+\varepsilon)} \quad \text{(A14)}$$

$$s_2(2) \geq \frac{a_H}{(p-\varepsilon)(p+\varepsilon)-(q-\varepsilon)(p+\varepsilon)} \quad \text{(A15)}$$

$$s_2(2) \geq \frac{2a_H}{(p-\varepsilon)(p+\varepsilon)-(q-\varepsilon)(q+\varepsilon)} \quad \text{(A16)}$$

It is easy to verify that: (A14) is more constraining than (A15) if and only if $\varepsilon > 0$; (A14) is more constraining than (A16) if and only if $\varepsilon \geq (p - q)/2$; and (A15) is more constraining than (A16) if and only if $\varepsilon \leq -(p - q)/2$.

The second-period expected pay depends on which of the three constraints, (A14) - (A16), is more constraining. There are three regions. For $\varepsilon \geq (p - q)/2$, (A14) binds and the second-period expected pay is $E[s_2] = \frac{(p-\varepsilon)(p+\varepsilon) a_H}{(p-\varepsilon)(p+\varepsilon)-(q-\varepsilon)(q+\varepsilon)} = \frac{(p+\varepsilon) a_H}{p-q}$, which is increasing in $\varepsilon$. The optimal news shock recognition in this region is $\varepsilon^* = (p - q)/2$.

For $\varepsilon \leq -(p - q)/2$, (A15) binds and the second-period expected pay is $E[s_2] = \frac{(p-\varepsilon)(p+\varepsilon) a_H}{(p-\varepsilon)(p+\varepsilon)-(q-\varepsilon)(p+\varepsilon)} = \frac{(p-\varepsilon) a_H}{p-q}$, which is decreasing in $\varepsilon$. The optimal news shock recognition in this region is $\varepsilon^* = -(p - q)/2$. 


For $-\frac{(p-q)}{2} < \varepsilon < (p-q)/2$, (A16) binds and the second-period expected pay is
\[ E[s_2] = \frac{2(p+\varepsilon)(p-\varepsilon)a_H}{(p-\varepsilon)(p+\varepsilon) - (q-\varepsilon)(q+\varepsilon)} = \frac{2(p^2-\varepsilon^2)a_H}{p^2-q^2}. \]
The optimal news shock recognition in this region is $\varepsilon^* = \pm (p-q)/2$.

Because $\varepsilon^*$ is nonzero, partially separating good and bad news shocks is always optimal, and unbiased accounting ($\varepsilon^* = 0$) is dominated. In all three regions, once plugging in the respective optimal $\varepsilon^*$, the second-period expected pay is $\frac{(3p-q)a_H}{2(p-q)}$. ■

**Proof of Proposition 7**

The proof is analogous to the proof of Proposition 6. The expected pay for the first and the third periods is written as
\[ (p_1 + \varepsilon) \frac{a_H}{p_1-q} + (p_2 - \varepsilon) \frac{a_H}{p_2-q} + \frac{p_1a_H}{p_1-q} + \frac{p_2a_H}{p_2-q} + \frac{\varepsilon(p_2-p_1)a_H}{(p_1-q)(p_2-q)}, \] which is increasing in $\varepsilon$.

The following constraints ensure the agent is motivated to work on both projects rather than working only on Project E (condition (A18)), or working only on Project F (condition (A19)), or not working on any project (condition (A20)).

\[ s_2(2) \geq \frac{a_H}{(p_2-\varepsilon)(p_1-q)}, \]
\[ s_2(2) \geq \frac{a_H}{(p_1+\varepsilon)(p_2-q)}, \]
\[ s_2(2) \geq \frac{2a_H}{(p_2-\varepsilon)(p_1+\varepsilon)-(q-\varepsilon)(q+\varepsilon)}. \]

Constraint (A18) is more stringent than (A19) if and only if $\varepsilon > \frac{(p_1-p_2)q}{p_1+p_2-2q} \equiv B_0$; (A18) is more stringent than (A20) if and only if $\varepsilon > \frac{p_1p_2+q^2-2p_2q}{p_1+p_2-2q} \equiv B_1$; and (A19) is more stringent than (A20) if and only if $\varepsilon < \frac{-p_1p_2+q^2-2p_1q}{p_1+p_2-2q} \equiv B_2$. Clearly, $B_2 < B_0 < 0 < B_1$. In addition, the assumption $d > \frac{p_1p_2+q^2-2p_1q}{p_1+p_2-2q}$ implies $B_2 > -d$ and $B_1 < d$.

There are three parameter regions to consider. For $\varepsilon \in [B_1, d]$, constraint (A18) binds, and the second-period expected pay is $E s_2 = \frac{(p_2-\varepsilon)(p_1+\varepsilon)a_H}{(p_2-\varepsilon)(p_1-q)} = \frac{(p_1+\varepsilon)a_H}{p_1-q}$. The choice of $\varepsilon$ has the same directional effect on the expected pay across three periods and, thus, the optimal news shock in this region is $\varepsilon^* = B_1$. 

For $\varepsilon \in [-d, B_2]$, constraint (A19) binds and the second-period expected pay is $\frac{(p_2-\varepsilon)a_H}{p_2-q}$.

The choice of $\varepsilon$ has opposite effects on (i) the second-period expected pay and (ii) the first- and the third-period expected pay. The assumption $p_2 - p_1 < p_1 - q$ ensures the impact on the second-period pay is dominant and, therefore, $\varepsilon^* = B_2$.

For $\varepsilon \in [B_2, B_1]$, constraint (A20) binds and the second-period expected pay is $\frac{2(p_2-\varepsilon)(p_1+\varepsilon)a_H}{(p_2-\varepsilon)(p_1+\varepsilon)-(q-\varepsilon)(q+\varepsilon)}$. The choice of news shock boils down to the choice between $\varepsilon = B_1$ and $\varepsilon = B_2$. It is straightforward to check that in this region $\varepsilon^* = B_2$ is optimal. ■
REFERENCES


