Performance Measure Manipulation*

JOEL S. DEMSKI, *University of Florida*

Abstract

A two-period model in which communication restrictions preclude the usual revelation representation is analyzed, and the communication policies take on the appearance of "income smoothing." The driving force is the information content of the "smoothed" or manipulated series, relative to its counterpart were manipulation not possible. Various possibilities arise, depending on the underlying stochastic structure: performance measure manipulation might be socially efficient, or not; and when it is best to invite and motivate this manipulation, the optimal policy itself can take on a variety of forms.

Condensé

L'auteur étudie un modèle mandant-mandataire à deux périodes dans lequel les résultats globaux des activités des deux périodes sont publiquement observables, la répartition de ces résultats entre les périodes pouvant cependant être falsifiée par le mandataire ou le gestionnaire. Le modèle est volontairement construit de telle sorte que soient neutralisés les divers problèmes liés à l'aversion pour le risque et au nivellement de la consommation, susceptibles de favoriser le nivellement ou autre falsification de l'information livrée relativement aux résultats de chaque période. Cette construction permet à l'auteur de mettre l'accent sur le thème du contenu en information.

Le gestionnaire fournit un effort productif et exigeant, sur le plan personnel, au cours de la première période, et il observe ensuite en privé les résultats de cette période. De nouveau, il fournit un effort productif et exigeant, sur le plan personnel, au cours de la seconde période. À cette étape, on lui demande aussi de produire un rapport personnel sur les résultats de la première période. Ce rapport peut être falsifié dans une certaine mesure : le gestionnaire pourrait, en effet, retenir une partie des résultats de la première période dans l'intention de l'ajouter au total de la seconde période ; en revanche, le gestionnaire pourrait énergiquement revendiquer l'attribution à la première période d'une partie des résultats de la seconde période.

La possibilité pour le gestionnaire d'attribuer faussement une partie des résultats d'une période à une autre période est limitée par l'intervention d'un vérificateur. Les résultats globaux de l'une ou l'autre période sont cotés 0, 1 ou 2 unités. Au moment où le

* Accepted by Jerry Feltham. This paper was presented at the 1997 *Contemporary Accounting Research* Conference, generously supported by: the **CGA-Canada Research Foundation**, the **Canadian Institute of Chartered Accountants, CGA-Ontario**, the **Society of Management Accountants of Ontario**, and the **Ernst & Young Foundation**. Helpful comments of seminar participants at Carnegie-Mellon, Columbia, Florida, Iowa, NYU, Rochester, UCLA, and the annual CAR Conference, and especially Anwer Ahmed, Jerry Feltham, Hans Frimor, Karl Hackenbrack, Sandra Kramer, Pierre Liang, Ajay Maindiratta, Bruce Miller, Steve Rockwell, Joshua Ronen, David Sappington, Bharat Sarath, Mary Sevier, Shiva Sivaramakrishnan, Rick Young, and two anonymous referees are gratefully acknowledged.

rapport personnel doit être produit, le gestionnaire peut connaître par anticipation le résultat de la seconde période. Il peut savoir, notamment, si les résultats de la seconde période seront « élevés » (c'est-à-dire s'ils atteindront avec certitude 1 ou 2 unités) ou s'ils seront « faibles » (c'est-à-dire s'ils s'établiront avec certitude à 0 ou 1 unité). Sans cette connaissance anticipée, le vérificateur peut repérer toute tentative de falsification du rapport. Par exemple, le fait de retenir une partie des résultats de la première période peut entraîner une surestimation flagrante des résultats de la seconde, à moins que le gestionnaire ne sache à l'avance que les résultats de la seconde période seront « faibles ». Lorsqu'il possède cette connaissance anticipée, le gestionnaire est donc en mesure de falsifier le rapport de telle sorte que la falsification ne puisse en aucun cas être détectée, que ce soit par le mandant ou par un observateur expérimenté.

Prenant pour acquis que le gestionnaire possède la connaissance anticipée requise, la fonction incitative qui entre en jeu par ailleurs peut ou non motiver le gestionnaire à falsifier son rapport, selon le contenu en information des résultats. Si l'on suppose que la fonction incitative qui entre en jeu par ailleurs motive le gestionnaire à falsifier son rapport, quelle sera l'attitude optimale à prendre ? La réponse peut varier. Si les chances que le gestionnaire obtienne la connaissance anticipée requise sont faibles, la falsification du rapport sera tolérée. Autrement, l'attitude optimale pourrait consister à transformer les incitatifs de sorte que le gestionnaire ne soit plus tenté de falsifier son rapport.

En outre, le gestionnaire qui peut acquérir la connaissance anticipée requise peut favoriser le mandat lui-même ou y porter préjudice. Si cette connaissance peut invariablement être acquise, le mandat s'en trouve généralement défavorisé. L'idéal serait qu'il soit possible de refuser au gestionnaire cette connaissance anticipée. Mais si le gestionnaire ne peut acquérir cette connaissance que si son comportement est satisfaisant par ailleurs, il est possible que le mandat soit favorisé par la falsification du rapport produit par le gestionnaire. Cela s'explique du fait que la falsification du rapport appelle une structure de série chronologique particulière dans la série des résultats faisant l'objet du rapport, et que cette structure elle-même est une source de renseignements, étant donné que le vérificateur ne peut être trompé par le gestionnaire que si le comportement de ce dernier est satisfaisant par ailleurs. En d'autres termes, s'il y a covariance entre la capacité de tromper le vérificateur et le comportement productif du gestionnaire par ailleurs, il est possible, dans ce contexte, que le mandat soit mieux servi lorsque le gestionnaire falsifie son rapport personnel sur les résultats.

Ainsi l'auteur présente-t-il de multiples possibilités dans un contexte unique et simple. La falsification du rapport peut ou non être supprimée rationnellement grâce à la modification de la structure incitative sous-jacente. De façon analogue, la falsification du rapport peut ou non être une source d'efficience dans la relation mandant-mandataire. Quoi qu'il en soit, la falsification opérée ne peut en aucun cas être observée.

1. Introduction

Self-serving management of performance indicators is regarded as an unavoidable fact of organizational life. Shipments can be discreetly delayed at the end of an unusually profitable year, just as maintenance and development can be postponed during an unusually troublesome year. Inventories can be hidden during the transition to a just-in-time inventory system, just as spending can be accelerated near the end of a budget-appropriation cycle.

Accounting versions of these concerns are equally familiar: the underlying procedures can be changed. Examples include a switch in depreciation method or a well-chosen time at which to adopt a new reporting requirement.[1] The procedure in place can also transform or distort an underlying performance series in some particular fashion. Examples include variable versus full-costing procedures in the presence of inventory, and the use of deferred tax accruals.[2] Of course, at the margin, discretion is present and particular accrual adjustments can be accelerated or retarded under unusual circumstances. Restructuring charges are illustrative.

Numerous studies have focused on these phenomena, many of which examine income smoothing. Accounting accruals can be the variable of interest (e.g., Healy 1985; McNichols and Wilson 1988), as can real variables (e.g., Lambert 1984; Hand 1989; Bartov 1993). In turn, the incentives that underlie the smoothing behavior can be exogenous (e.g., Healy 1985) or endogenous (e.g., Lambert 1984).[3]

In this paper, a two-period principal-agent formulation in which the manager has an option to misreport first-period performance is presented. Any misreport must be reversed in the second period, as total output is observed at the game's conclusion. The principal, in turn, can always design the incentives so the manager will be motivated to exercise or to reject any such option, while at the same time maintaining requisite incentives for otherwise productive behavior. In this respect, any misreporting, any performance measure manipulation, is explicitly induced by the principal.

Equilibrium misreporting, however, takes a variety of forms in the model. It might be destructive in that it garbles the underlying series, but (1) is of no interest to either player; (2) is of second-order importance and, therefore, tolerated; or (3) destroys so much information the agent's incentives are designed so it does not take place. Alternatively, it might be efficiency enhancing and, therefore, encouraged. It all depends on the information content of the performance series. This wide variety of responses and implications is the central conclusion of the analysis.

The underlying tension is the joint product of aggregation and the incentive structure that would be in place were the misreporting options absent. Aggregation, within and between periods, is what fosters the misreporting options; the natural incentives follow from the information content of the periods' output totals, if they could be publicly observed (or correctly reported). Changing risk aversion, consumption smoothing, career concerns, and reliance on short-term contracts are all neutralized in the analysis. This is done to highlight the information-content theme.

The model is presented in the second section where, for benchmark purposes, the case in which the misreporting options are absent is analyzed. In the third section, a set of conditions under which the option to misreport is desirable and exercised is presented. Following this, in the fourth section, a set of conditions under which the option to misreport is undesirable and is responded to by tolerance or substantively altered incentives, making misreporting unattractive to the agent is presented. Ties to the literature are explored in the fifth section. Concluding remarks round out the paper.

The model stresses endogenous response to the manager's reporting options. The principal is fully aware of the manager's options, and designs the instructions and incentives to best serve the organization in light of these options. If it is best for the organization that the manager misreports in some particular fashion, this misreporting will be motivated. If it is not in the best interest of the organization that the manager has these options, but the manager nevertheless does, the principal must then decide whether less damage is done by condoning or by motivating rejection of the misreporting opportunities. Depending on finer details, then, equilibrium play may, or may not, reflect performance measure manipulation.

Smoothing the performance series, a particular version of misreporting, is emphasized. This reflects its significant role in the empirical literature, as well as its importance in the present model.

2. Model

The underlying model is a streamlined agency setting in which accounting recognition is an issue. A manager supplies a costly input (e.g., effort), observes a first period output, and then supplies a second costly input. The parties commit to the two-period engagement. The production possibilities and stochastic output are independent across periods. This is imposed to remove any interest in evaluating the manager based on the sequence of outputs.

Preliminaries

Output in each period is the sum of two binary random variables: $x = x_1 + x_2$ denotes total output in the first period. For convenience, I scale output so $x_1, x_2 \in \{0,1\}$, implying $x \in \{0,1,2\}$. Output in the second period is modeled in the same fashion, and total second-period output is denoted $y = y_1 + y_2$, with $y_1, y_2 \in \{0,1\}$, etcetera. First-period output, as self-reported by the manager, and publically observed total output of $x + y$ provide the contracting variables. The underlying binary structure is meant to suggest aggregation; it is also important in specifying the reporting alternatives that are open to the manager.[4]

Each period, the manager is called upon to supply one of two inputs, "low" (L) or "high" (H). Let $a_t \in \{L,H\}$ denote this input in period t.

As mentioned, the output structure is repeated across the two periods, in conditionally independent fashion: $\text{prob}(x,y|a_1,a_2) = \text{prob}(x|a_1)\text{prob}(y|a_2)$. Let $\pi_t(a_t)$ denote the vector of output probabilities for period t, conditional on that period's input supply. I assume:

[A1] Output is conditionally independent across periods, along with (1) $\pi_1(H) = \pi_2(H) = \omega \in \mathbb{R}^3$; (2) $\pi_1(L) = \pi_2(L) = \hat{\omega} \in \mathbb{R}^3$; and (3) $\hat{\omega}(0)/\omega(0) \geq \hat{\omega}(1)/\omega(1) \geq \hat{\omega}(2)/\omega(2)$ (i.e., Monotone Likelihood Ratio Property [MLRP]).

Unless otherwise noted, I also assume all elements of $\pi_t(a_t)$ are strictly positive (full support).

The manager supplies input a_1, receives initial payment I^1 (following release of the first period self-report), supplies input a_2, and then receives second payment I^2 (following observation of $x + y$). The manager's preference measure is given by

[A2] $U(I^1,I^2;a_1,a_2) = -\exp[-r(I^1 + I^2 - ka_1 - ka_2)]$

where $r > 0$ is the Arrow-Pratt measure of risk aversion and $k > 0$ is the equivalent pecuniary cost per unit of input a_t. The principal's preference measure is given by

[A3] $P(x + y) - I^1 - I^2$

where $P > 0$ is the value per unit of scaled output to the principal.

Notice no discounting is present in [A2] or [A3]. Moreover, any (I^1,I^2) sequence is equivalent to $(0, \hat{I})$, where $\hat{I} = I^1 + I^2$ from either the principal's or the manager's perspective. Total compensation is what matters to both parties. The setup is purposely constructed so any interest in performance management will not have its ties in a desire for consumption smoothing or consumption timing.

We also assume that when P is sufficiently large, the principal prefers to induce input H:

[A4] input $H > L$ is motivated in each period and circumstance.

(Audited) self-reporting

As noted, the manager is called upon to self-report the first period's output. The time line, once contract terms are agreed upon, is as follows: (1) the manager supplies first-period input $a_1 \in \{L,H\}$; (2) the manager privately observes first-period output $x \in \{0,1,2\}$; (3) the manager supplies second-period output $a_2 \in \{L,H\}$; (4) after possibly and privately observing $y_1 \in \{0,1\}$, the manager self-reports first-period output of $\hat{x}$; and (5) total output of $x + y$ is publicly observed and accounts are settled.

The timing is meant to be suggestive of a typical reporting environment. The manager issues the (accounting) report. That report is somewhat delayed, giving rise to the possibility the manager already has partial knowledge of second-period output before the first-period report must be filed. In an attempt to retain arguably important features of a typical reporting environment, this mid-game self-report is audited.

At the time of the self-report, the manager privately knows first-period output, $x \in \{0,1,2\}$, whether the early read on second period output was possible, {yes, no}, and if so, what that read revealed, $y_1 \in \{0,1\}$. The manager's reporting strategy, then, can be thought of as a function that maps observables into some message space Y:

$m: \{0,1,2\}\times\{\text{yes},\text{no}\}\times\{0,1\} \rightarrow Y$
such that $m(x,\text{no},0) = m(x,\text{no},1)\ \forall x \in \{0,1,2\}$,

where the side restriction reflects the measurability restriction when y_1 is not observed.

The auditor is not explicitly modeled. Rather, two stylized features of an audit function are imposed on the reporting. One is simple. The audit function will not tolerate any self-reporting strategy that leads, with strictly positive probability, to it being common knowledge the self-report was false. This places two restrictions on the self-report:

[A5] $Y = \{0,1,2\}$; and

$(x + y - \text{m}(x,\cdot,y_1)) \in \{0,1,2\}$ for all x and $y \in \{0,1,2\}$ and $y_1 \in \{0,1\}$.

Any claim of first-period output outside of $Y = \{0,1,2\}$ would be clearly inappropriate from an audit standpoint, and, therefore, is not allowed. The auditor simply could not certify a known false output report. Likewise, any report that, with the benefit of hindsight, would be demonstrably false is ruled out. Again, the auditor simply could not be party to such fraudulent reporting.[5] These restrictions can be thought of as reflecting the judicious work of an unmodeled auditor, or the imposition of severe penalties on the manager. Regardless, the intent is to exhibit a self-reporting exercise in which the report has an unmistakable interpretation as a claim of output produced, and systematic under- or over-reporting of output is ruled out because it will, with strictly positive probability, be discovered.

It should be noted [A5] is not an idle assumption. Restricting the message space removes our ability to appeal to the revelation principle, because, once full, truthful revelation has been abandoned, it is not guaranteed some truthful coarsification is optimal.

With these reporting restrictions in place, the manager's reporting behavior is limited to some combination of honest and judicious, subtle movement of output between the periods. If y_1 is not observed before the report is due, the manager must report $\text{m}(x,\text{no},y_1) = x$. Any nontruthful report has positive probability of being detected in this event. With timely private observation of y_1, though, the manager has a variety of options. These are listed below.

[A5a] $\text{m}(0,\text{yes},1) \in \{0,1\}$;

[A5b] $\text{m}(2,\text{yes},0) \in \{1,2\}$;

[A5c] $\text{m}(1,\text{yes},0) \in \{0,1\}$; and

[A5d] $\text{m}(1,\text{yes},1) \in \{1,2\}$.

For example, when $x = 0$ and $y_1 = 1$ have been observed, the manager can report $\text{m}(0,\text{yes},1) = 0$, an honest report, or $\text{m}(0,\text{yes},1) = 1$, an inflated report.

[A5a] is interpreted as the "borrow" option. Under $x = 0$ and having seen $y_1 = 1$, the manager has the option of surreptitiously borrowing some output from the second period, aggressively recognizing some of the second-period's output. Similarly, I interpret [A5b] as the "loan" option under which some of the first period's output is loaned to the second period. Use of the options in [A5a] and [A5b] are the stylization of performance smoothing in this model. Their use clearly smooths performance.

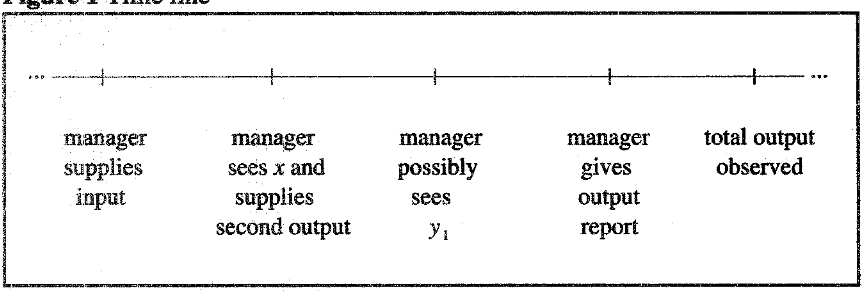
[A5c] and [A5d] confront the manager with the same type of opportunity, but have the opposite effect of "desmoothing" performance.

This leads to the second stylized feature of the audit function. Can the auditor detect use of any of the misreporting options in [A5]; stated differently, can the auditor detect any subtle misreporting behavior? I will explore two possibilities. One is the auditor has no such capability. I term this the "weak" auditor case. The other is the auditor can detect, and thus prevent, use of the de-smoothing behavior in [A5c] and [A5d]. I term this the "strong" auditor case. Thus, with the strong auditor, the manager's reporting management is confined to factual or smoothed reporting. Intuitively, the strong auditor is one who has some ability to detect subtle misreporting, but is disadvantaged in the performance extremes, where poor performance might be bolstered by aggressive accounting or superb performance might be shaved by creating reserves for the inevitable rainy day.

A final issue is whether the manager is able to observe y_1 before issuing the first period's output report. Though easiest to visualize are the extreme cases where the manager observes y_1 in timely fashion or not, it will prove insightful to parameterize on the probability this observation occurs before the report is due. In addition, the manager may be able to affect this probability. This suggests a multitask formulation, with some activity aimed at learning y_1 in a timely fashion. It is sufficient, however, to treat this probability as depending on the manager's (a_1, a_2) effort supply. For example, unusual attention to detail, such as supplying $a_1 = a_2 = H$, might lead to higher odds that y_1 is privately observed in timely fashion. To represent this possibility, let $\beta(a_1,a_2)$ denote the probability the manager privately observes output y_1 before the output report is due, conditional on input sequence (a_1,a_2).

The various details are summarized by the time line in Figure 1.[6] Notice the role of aggregation. The manager self-reports aggregate output in the first period, and never reports whether y_1 was observed in timely fashion. In addition, the final, public report aggregates first and second period output.

Figure 1 Time line



$\beta(a_1,a_2) \equiv 0$ benchmark

The model setup is concluded by identifying a useful structure in a benchmark setting, where misreporting is not possible: $\beta(a_1,a_2) = 0$ for $a_1, a_2 \in \{L,H\}$. This implies the contracting variables are simply output x in the first period followed by (with $x + y$ publically observed) output y in the second period. As suggested above, let $\hat{I}_{xy} = I_x + I_{xy}$ denote the manager's equivalent total payment when output pair (x,y) is observed. Under [A3] and [A4] the mechanism design problem is to minimize the expected total payment, subject to individual rationality and incentive compatibility constraints.

Let $\hat{a}$: $\{0,1,2\} \to \{L,H\}$ denote a particular input supply strategy for the manager's second period input choice. $\hat{a}^H$ is the preferred strategy of $\hat{a}^H(x = 0) = \hat{a}^H(x = 1) = \hat{a}^H(x = 2) = H$. Also let $E[U(\hat{I}_{xy},\cdot)|a_1,\hat{a}]$ denote the managers expected utility when input a_1 is combined with second period strategy $\hat{a}$ in the presence of compensation arrangement $\hat{I}_{xy}$. The mechanism design program for this benchmark is:

$$F^* \equiv \text{minimum } \Sigma_{x,y} \hat{I}_{xy}\,\omega(x)\omega(y) = E[\hat{I}_{xy}|H,\hat{a}^H] \qquad [B]$$

subject to:

$$E[U(\hat{I}_{xy},\cdot)|H,\hat{a}^H] \geq U(W,0,0) \qquad [IR]$$

$$E[U(\hat{I}_{xy},\cdot)|H,\hat{a}^H] \geq E[U(\hat{I}_{xy},\cdot)|a_1,\hat{a}]\ \forall\ a_1 \in \{L,H\},\ \hat{a}: \{0,1,2\} \to \{L,H\}. \qquad [IC]$$

[IR] is the manager's individual rationality constraint, with reservation utility denoted $U(W,0,0)$. W, then, is the manager's reservation wage. [IC] is the family of incentive compatibility constraints.

Utilizing results in Fellingham, Newman, and Suh 1985 and MLRP, the solution to [B] is familiar. Compensation is increasing in either output; and the payment structure is additively separable. Output sequence carries no information.

FACT 1. *Assume [A1]–[A4]. Then the solution to program [B] can be expressed as* $\hat{I}_{xy} = J_x + J_y$, *with* $J_0 \leq J_1 \leq J_2$.

Moreover, J_0, J_1, and J_2 are the solution to the comparable one period problem with [IR] based on $U(W/2,0,0)$.[7]

The relationship among J_0, J_1, and J_2 is an important key in developing intuition in the following sections. [A1], of course, implies $J_0 \leq J_1 \leq J_2$. Beyond that, the likelihood ratio also determines whether increasing or decreasing returns to good news are present. In [A6′] the likelihood ratios imply decreasing returns to good news. In [A6″] they imply increasing returns to good news.[8]

[A6′] the solution to program [B] has $\hat{I}_{xy} = J_x + J_y$ with $2J_1 > J_2 + J_0$; and

[A6″] the solution to program [B] has $\hat{I}_{xy} = J_x + J_y$ with $2J_1 < J_2 + J_0$.

The importance of [A6] is conveyed by the following. Suppose the solution to [B] were imposed in a setting where $\beta(\cdot,\cdot) > 0$, and $x = 0$ and $y_1 = 1$ have been observed. The compensation lottery, as a function of the self-report and the yet unknown y_2 realization, is:

	$y_2 = 1$	$y_2 = 0$
m(0,yes,1) = 1	$J_1 + J_1$	$J_1 + J_0$
m(0,yes,1) = 0	$J_0 + J_2$	$J_0 + J_1$

[A6′] invites use of the borrow option, whereas [A6″] invites rejection of the borrow option. Parallel comments apply to the other possibilities. The point is, the incentive structure that would prevail if misreporting were not possible may well induce misreporting.

For later reference, suppose the underlying probability structure is Bernoulli with $a_t = H$ implying a single stage success probability of $s > \hat{s}$, where the latter is the single stage success probability under $a_t = L$. This provides $\omega(0) = (1-s)^2$, $\omega(1) = 2s(1-s)$ and $\omega(2) = s^2$, along with $\hat{\omega}(0) = (1-\hat{s})^2$, $\hat{\omega}(1) = 2\hat{s}(1-\hat{s})$, and $\hat{\omega}(2) = \hat{s}^2$. It also provides decreasing returns to good news in the benchmark setting.[9]

FACT 2. *Assume [A1]–[A4] and a Bernoulli structure. The solution to program [B] satisfies [A6′].*

The paper begins with a benchmark setting where the incentive structure is constant across the periods, where equilibrium output and compensation are independent and identically distributed across the periods, where output sequence is uninformative, and where there is a lurking, induced (natural) interest on the manager's part in judiciously manipulating the performance measures by using the noted borrow and loan options.

3. Limited misreporting case

Now examine the case where the borrow and loan options of [A5] are available only if the manager supplies input H in both periods.

[A7′] $\beta(H,H) = \alpha$; and $\beta(H,L) = \beta(L,H) = \beta(L,L) = 0$.

This emphasizes the theme that misreporting relies on an early read of second-period output being available before the first period's report must be announced, and this early read is possible only under the $a_1 = a_2 = H$ policy. For example, a side effect of diligent managerial behavior might be unusual knowledge of the customer base, or an ability to hide use of a misreporting option from the auditor. I call [A7′] the "limited misreporting" case.[10]

Let M denote the set of self-reporting strategies consistent with [A5]. (Whether the weak or strong auditor is present will be clear from the context.) Also let $E[U(\hat{I}_{xy},\cdot)|a_1,\hat{a},m]$ denote the manager's expected utility, if second-period input strategy â and reporting strategy m are used in the presence of $\hat{I}_{xy}$ and initial input a_1. Applying parallel notation to the expectation of $\hat{I}_{xy}$, the mechanism design program to best motivate [A4] and reporting strategy m ∈ M is:

F(m) = minimum $E[\hat{I}_{xy}|H,\hat{a}^H,m]$ [R] subject to:

$E[U(\hat{I}_{xy},\cdot)|H,\hat{a}^H,m] \geq U(W,0,0)$ [IR]

$E[U(\hat{I}_{xy},\cdot)|H,\hat{a}^H,m] \geq E[U(\hat{I}_{xy},\cdot)|a_1, \hat{a}, \underline{m}]$

$\forall a_1 \in \{L,H\}$, $\hat{a}: \{0,1,2\} \rightarrow \{L,H\}$, and $\underline{m} \in M$. [IC]

Efficient smoothing in the presence of a strong auditor

Now suppose the strong auditor is present. The interesting reporting strategies are truthful reporting, denoted m^T, and symmetric smoothing, denoted m^S, in which m(0,yes,1) = 1, m(2,yes,0) = 1 and no distortion is present under $x = 1$. Policy m^S, of course, is the case where the borrow and loan options are fully engaged, wherever feasible, in the presence of a strong auditor.

Whether smoothing, via m^S, is efficient is ambiguous. Smoothing adds noise to the performance statistics, but the noise is correlated across periods; in the limited misreporting case, it is possible to smooth, to add correlated noise to the performance series, only if the manager supplies high effort at each stage.

The Bernoulli setting is one case that comes down in favor of smoothing. Here, the natural incentive in the benchmark case is for the manager to misreport via policy m^S (Fact 2). To exploit this natural incentive, further suppose the single period Bernoulli version of the story exhibits $2J_1 + J_2 + J_0 < 2(W + 2kH)$: the bad news of x = 0 leads to considerable "steepness" in the incentive structure.

This steepness condition is not always present; for example, it clearly fails when $s = 0.5$. However, it is intuitive and far from vacuous.

LEMMA. *Suppose the benchmark case [B] under the Bernoulli structure, based on success probabilities of s under input H and $\hat{s} > 0$ under input L, has $2J_1 + J_2 + J_0 \geq 2(W + 2kH)$. There exists $s' \in (s,1)$ and $\hat{s}' \in (0,\hat{s})$ such that the benchmark case has $2J_1 + J_2 + J_0 < 2(W + 2kH)$ for all s" $\in [s',1]$ and $\hat{s}'' \in [0,\hat{s}']$.*

Importantly, the steepness condition provides a sufficient condition for smoothing to be efficient in the limited misreporting case. Suppose we leave the benchmark incentive structure in place. Smoothing, which is naturally invited, raises the manager's expected compensation and also lowers the compensation risk. In this case, the risk reduction is worth more to the manager than the gain in expected compensation. So social gains to smoothing are strictly positive.

PROPOSITION 1. *Assume [A1]–[A5] and [A7′], and the Bernoulli structure are present. Further suppose the benchmark case [B] exhibits $\hat{I}_{11} + \hat{I}_{02} = 2J_1 + J_2 + J_0 < 2(W + 2kH)$. Then $F(m^S) < F^* < F(m^T)$ for any $\beta(H,H) > 0$.*

Notice how garbled performance measures, the smoothed output reports, provide a superior contracting venue. The reason, of course, is the garbling caused by the smoothing is differentially possible across the effort levels; the statistical properties of the garbled series can be exploited to advantage. The key to the story is the fact this garbling is infeasible given other than high-effort supply at each stage, and it is relatively easy, thanks to the Bernoulli structure, to motivate the manager to follow the smoothed reporting policy. Indeed, this result holds regardless of whether the auditor is weak or strong, as it exploits the natural misreporting settings that arise in the Bernoulli story.[11]

Variations on the strong auditor, Bernoulli theme

Smoothing is desirable here because of the limited communication, the implicit smoothing incentives in the benchmark case, and the fact that a smoothing option can surface only under good behavior by the manager. Once a revelation argument has been sidestepped, less than candid reporting may be efficient, and Proposition 1 provides a setting where that is the case.[12] Several variations on this observation are noted below.

First, dropping the strong auditor opens up other interesting misreporting possibilities. But even with the Bernoulli specification, ambiguity over the most productive misreporting policy prevails. This is illustrated with a numerical example.

For this purpose, assume the Bernoulli specification is present, with success probability $s = 0.8$ under input H (= 10,000) versus $\hat{s} = 0.2$ under input L (= 5,000). So $\omega = [0.04, 0.32, 0.64]$ and $\hat{\omega} = [0.64, 0.32, 0.04]$. Remaining details include an Arrow-Pratt measure of $r = 10^{-4}$, personal-cost parameter of $k = 1$, and a reservation wage for the manager of $W = 20{,}000$.

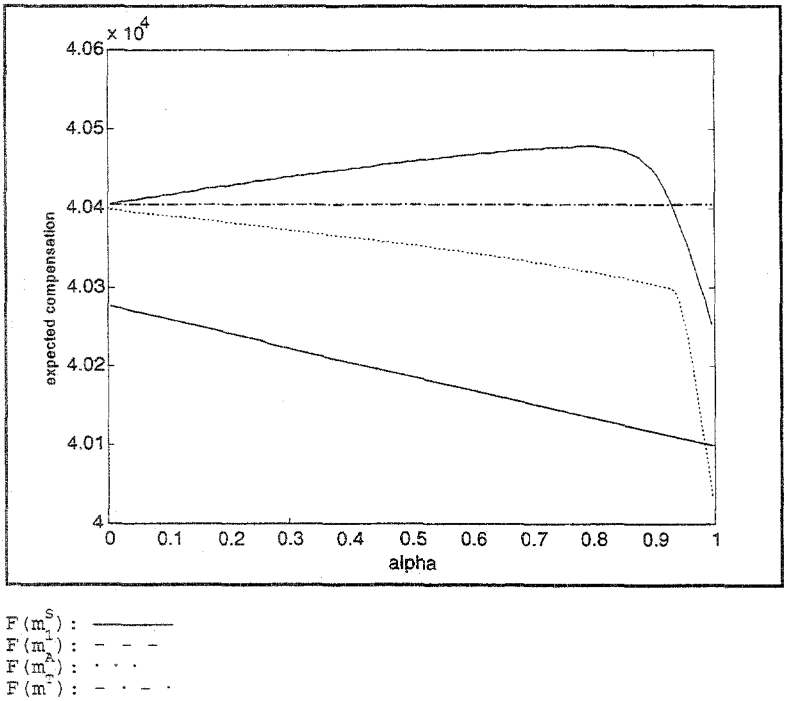
The benchmark program [B] reveals an incentive structure based on $J_0 =$ 12,929, $J_1 = 20{,}228$, and $J_2 = 20{,}545$, along with $E[\hat{I}_{xy}] = 40{,}277$.[13] In Figure 2, F(m) for $0.005 \leq \beta(H,H) = \alpha \leq 0.995$ for various reporting strategies is plotted. The solid line is the smoothing case of $F(m^S)$, and the dashed line (- - -) reflects the opposite policy of misreporting only under $x = 1$. Call this latter policy m^1. The dotted line (···) reflects an asymmetric policy of using the loan option under x = 2 and again under x = 1. Denote this latter policy m^A. Finally, the dash-dot line (-·-·) is the truthful reporting case, $F(m^T)$ (= 40,407).

The decreasing returns to good news feature of the Bernoulli structure is evident. The benchmark incentives invite use of the m^S policy. Capitalizing on this natural tendency is productive, as exhibited in the Figure and in Proposition 1. Alternative policies must dampen this natural incentive, and doing so puts excessive pressure on the control apparatus, as is apparent in Figure 2.

Second, we see for large α that $F(m^A) < F(m^S)$. This is intuitive, and best analyzed at the limit point of $\alpha = 1$. Here, with guaranteed access to the smoothing options, given good behavior by the manager, it is possible to guarantee that particular outcome sequences do not arise. For this reason it becomes relatively easy to dampen the natural misreporting incentive in the benchmark case. For example, under $\alpha = 1$, the manager can use the borrow option under $x = 0$ (and $y_1 = 1$) to guarantee the output report pair (0,2) is never observed. In turn, the optimal incentive contract can offer full insurance for the associated report pairs of (1,2)

Errata Sheets for Volume 15, Fall 1998

Figure 2 F(·) versus $\beta(H,H) = \alpha$, under [A7′]; $0.005 \leq \alpha \leq 0.995$



and (2,2). Yet, with limited borrow and lend options, at least one of the continuation points of $x \in \{0,1,2\}$ cannot be so protected. The important feature of the symmetric policy (m^S) is that it leaves $x = 1$ unprotected. In a sense, this is where the manager faces significant compensation risk.

Moving the lack of protection to the point $x = 0$, the essence of policy m^A, is intuitive now. It places the lack of insurance in the lowest odds event of x = 0 (presuming the success probability under good behavior exceeds 0.5), and this can be done without putting excessive pressure on the natural, benchmark incentives.[14]

PROPOSITION 2. *Assume [A1]–[A5] and [A7′], the Bernoulli structure with $s > 0.5$, the weak auditor and $\beta(H,H) = 1$. Then $F(m^A) < F(m^S) < F^*$.*

Again, $\beta(H,H) = 1$ is a highly structured setting, given the apparent moving support and the fact we no longer have to deal with the natural tension of implicitly motivating the m^S policy. This is why the asymmetric policy surfaces here.[15]

Third, the Bernoulli structure is also at work here in a more subtle fashion. It implies independence between the first and second units in each period as well.

This creates considerable structure in the mid-game reporting activity, because knowledge of y_1 provides no knowledge about y_2. This suggests dropping the Bernoulli structure will lead to further ambiguity, even in the limit case of $\alpha = 1$.[16]

To pursue this, slightly, let h^a denote the vector of disaggregate output probabilities, conditional on input a (i.e., prob$((x_1,x_2)|a_1)$ or prob$((y_1,y_2)|a_2)$): h^a = [prob((0,0)|a), prob((0,1)|a), prob((1,0)|a), prob((1,1)|a)]. For example, in the Bernoulli case we have $h^H = [(1-s)^2, s(1-s), s(1-s), s^2]$.

Using the same numerical setup as before, let h^H = [0.04, 0.16, 0.16, 0.64] and $\mathbf{h}^L$ = [1, 0, 0, 0]. Retain $\alpha = 1$. This, of course, is another Bernoulli story, and we know from Proposition 2 that $F(m^A) < F(m^S)$. But with $\mathbf{h}^H$ = [0.4, 0, 0.1, 0.5] we find $F(m^A) > F(m^S)$. The key is loading probability mass in the area where one policy or the other leaves the manager at significant risk.[17]

Thus, depending on the conditions under which the borrow and loan options arise, it may be efficient to motivate various self-reporting policies. In this sense, we see different aggregation policies substituting for full communication. Most important, the borrow and loan options may be a source of efficiency.[18]

4. Unlimited misreporting case

Now examine the case where the borrow and loan options of [A5] are available regardless of the manager's input supply:

[A7″] $\beta(H,H) = \beta(H,L) = \beta(L,H) = \beta(L,L) = \alpha.$

[A7″] is termed the "unlimited misreporting" case. Whether the borrow or loan options surface is completely random under [A7″]. For example, the time at which the manager observes y_1 is independent of whatever input the manager supplies. So, the same mechanism design program structure ([R]) surfaces, but the added misreporting possibilities implied by [A7″] lead to a comparable increase in the variety of the incentive compatibility constraints.

Here there is no ambiguity on the efficiency question. Indiscriminate access to the borrow and loan options, the essence of [A7″], is at best a matter of indifference. Suppose the strong auditor is present. Consider truthful reporting, the m^T policy. Under [A6″] these reporting instructions will gladly be followed when the benchmark compensation scheme is in place; that is, the solution to [B] is feasible in [R]. Conversely, if the original solution is progressive, if [A6′] holds, the solution to [B] is not feasible in [R] when we seek to implement truthful reporting.

Moreover, use of any misreporting policy, any reporting policy other than m^T, merely transforms the output vector in random fashion. From the principal's perspective, this amounts to the injection of noise into the output observation. Grossman and Hart's (1983) Proposition 13, with slight modification, is definitive. The key is the fact the garbling of the privately observed performance is indiscriminately available; its generic character does not depend on the manager's effort supply.

Errata Sheets for Volume 15, Fall 1998

PROPOSITION 3. *Assume [A1]–[A4], [A6′] or [A6″], and [A7″]. Then $F^* \leq F(m)\ \forall m \in \mathcal{M}$*

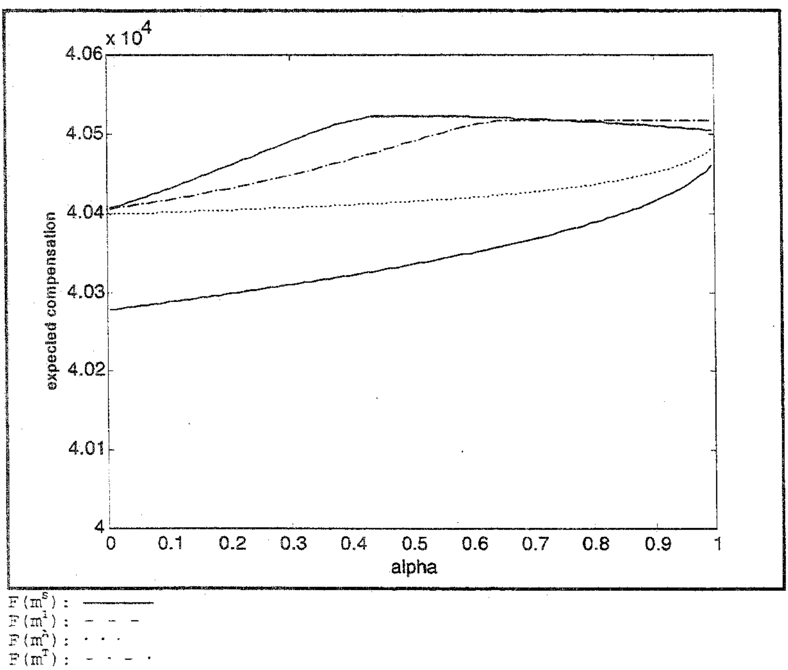
This is well illustrated by the earlier numerical illustration. Everything remains as before, except [A7″] replaces [A7′]. The usual suspects are plotted in Figure 3. The natural smoothing incentives in the benchmark solution again favor the symmetric misreporting policy; $F(m^S)$ is below its competitors. The contracting friction increases with α, reflecting the theme of indiscriminate garbling. Most important though, is the fact the option to misreport in the unlimited misreporting case is never a source of efficiency.

A side issue is how best to deal with these options, when they are present. Again, ambiguity is present. The limit point of $\alpha = 1$ in Figure 3, for example, reveals $F(m^S) > F(m^T) > F(m^1) > F(m^A)$.[19]

5. Ties to the literature

To tie the analysis to the literature, recall the key features in the model are: (1) the information content of the underlying output statistics in the benchmark setting, as this determines the "natural" misreporting tendency; (2) what the manager might know at the time of the exercise, including the intra-period information

Figure 3 F(•) versus α, under [A7″]; $0.005 \leq \alpha \leq 0.995$



content of output; and (3) how the manager's directly productive activity affects what might be known at the time of the exercise. In addition, the model relies on aggregation, as only aggregate output is reported or observed, and also relies on an auditor, whose influence limits the message space to a believable output report.

In turn, the principal conclusion is one of ambiguity. We see ambiguity in whether the misreporting might be productive, or counterproductive. We also see considerable ambiguity in how best to deal with the presence of these misreporting options. This reflects the fact the tensions exhibited in the model net in various directions, once misreporting is allowed to substitute for full reporting.

Performance misreporting, especially smoothing, has been examined from a variety of angles.[20] Recent empirical literature exploits the intuitive role of what I have termed benchmark incentives to identify plausible conditions under which performance misreporting, especially smoothing, might be practiced. The targeted conditions might be, for example, parametric conditions in publicized incentive structures (such as a minimal income level to qualify for a bonus), proxy contests, management buyouts, debt covenant violations, poor relative performance, etcetera. Alternatively, valuation implications of various treatments might be sought.[21]

A common technique is to focus on so-called discretionary accounting accruals, as measured by the residual in the Jones 1991 model. Intuitively, it is in the discretionary accruals that accounting-based performance management or misreporting is exercised; and associating these accruals with unusual events allows us to examine reporting activity in those unusual events. This has the advantage of exploiting accounting structure, via the statistical pattern in accruals.[22] It is also problematic in the presumption the misreporting is observable.

By analogy, the discretionary accruals in the model analyzed here would be actual, less reported output in each period; yet the very essence of the modeling is one in which the manager privately learns of and possibly exercises the misreporting option. Being able to observe or directly identify the discretionary component of the report removes any reason to misreport, as it returns us, presumably, to a revelation-style setting.

In turn, recent modeling literature stresses contracting frictions that result in inevitable or otherwise optimal misreporting.[23] The common theme across the models is equilibrium misreporting, implying a revelation argument is unavailable. Limited communication is the usual story, though lack of commitment also receives some play. Either way, a revelation argument cannot be relied upon.[24]

The empirical literature stresses observability of the performance management activity, while the modeling literature stresses its inherent non-observability coupled with common knowledge that the activity is present and engaged. The earlier numerical example illustrates the distinction. Suppose, for the sake of illustration, that the residual claim to this output is priced or valued at its expected value. A priori, $E[x + y] = 2(1.6) = 3.2$. Revision in this expectation, conditional on initial output being reported, is displayed in the Table 1.

The independence assumption is at work in the truthful reporting case (m^T). Observation of actual output ($x = \hat{x}$) carries no information about output y, so the conditional expectation of total output is simply $x + 1.6$.

TABLE 1
Revised conditional expectations

reporting policy	$\hat{x} = 0$	$\hat{x} = 1$	$\hat{x} = 2$
m^T	$E[x + y \mid \hat{x}] = 1.60$	$E[x + y \mid \hat{x}] = 2.60$	$E[x + y \mid \hat{x}] = 3.60$
m^S, $\alpha = 0.4$	$E[x + y \mid \hat{x}] = 1.51$	$E[x + y \mid \hat{x}] = 2.60$	$E[x + y \mid \hat{x}] = 3.67$

But, in the symmetric smoothing case, observation of output report $\hat{x}$ carries information about second-period output as well. For example, under $x = 0$ it is clear the manager was unable to misreport. Either the possibility was absent (as $\alpha = 0.4$), or it was present, but $y_1 = 0$. So $E[x + y \mid \hat{x} = 0]$ declines from 1.60 to 1.51.

Expectation revisions are larger in the presence of the misreporting policy. That is, the valuation content of the initial output report depends on which reporting policy is present. Yet when policy m^S is present, the actual misreporting is not visible; rather, the policy in place is common knowledge. The information question is then not one of decoding an observed or discernible misreporting of performance; it is one of decoding an observed performance report, knowing that unobservable misreporting may be influencing this report. In this fashion, the initial output report carries, or does not carry, information about future performance.[25]

This implies that, in the empirical counterpart to this style of model we should not be able to detect misreporting behavior per se. In principle, though, a firm that employs a particular misreporting strategy would exhibit a statistical series distinguishable from that of another firm that employs a different misreporting strategy.[26] Naturally, and as exhibited here, there is also no reason to presume the misreporting policy applied to the underlying real series is constant across a sample of firms, even after controlling for growth and other market conditions faced by the firm. This creates additional difficulty in pooling various misreporting policies for empirical purposes.

6. Concluding remarks

The model is designed to highlight several features of the performance management game: (1) an accounting procedure aspect that is suggestive of accrual recognition; (2) an interest in possibly misreporting the performance series that is divorced from consumption smoothing, signaling, or cross-period dependencies; and (3) the fact any such misreporting must pass through an audit filter. This leads to an ambiguous view of whether such misreporting might be either tempting or efficient. The closer link to accounting structure is reflected in the additive nature of output, the aggregation of output statistics, the possibility of shifting recognition across periods, and the presence of an auditing function. Aggregation is a hallmark of accounting procedure, and is central to the model. Consumption smoothing, signaling, and the importance of output sequence are de-emphasized to highlight the information content side of the issue. Misreporting, therefore, surfaces simply because of aggregation in both the accounting records and communication between the players. In turn, whether the ability to misreport is desirable is linked to whether and how this ability covaries with otherwise desirable behavior by the manager. When desirable, motivating active performance management illustrates the possibility of designing a performance report to carry higher information content for control purposes, at the implicit cost of carrying lower information content for valuation purposes.

The net result is a setting suggestive of a sample of firms in which misreporting may or may not be present in the observed equilibrium behavior, and in which the efficiency implications may well vary across the sample. Where present, the misreporting may reflect a variety of policies.

As with other models of misreporting, however, it is not possible to observe the misreporting per se. The misreporting possibility and organizational overlay of instructions and incentives are common knowledge in the model, but the actual misreporting is strictly private. By implication, the interim report carries different information, depending on whether the underlying misreporting possibility is present and motivated. This creates different valuation implications of the first period report, depending on the equilibrium misreporting resolution.

Continuing the analogy, the empirical distinction between discretionary and nondiscretionary accruals is moot in this setting. Any misreporting is discretionary and done privately; it is unobservable to an outsider. (Otherwise, smoothing would be observable, which gets us close to presuming the misreporting option is observable.) Continuing, accruals cannot be factored into discretionary and nondiscretionary components in the setting. Rather, potential smoothers in the sample would be distinguished from non-smoothers by examining the time-series behavior of the performance reports.[27] This returns us to the fact that misreporting alters the equilibrium information content of the information release. So, in principle, the use of a misreporting policy would be empirically identifiable, but not the report manipulation itself.

The model's ability to exhibit a variety of equilibrium conclusions does, of course, come at a cost. The additive, binary outcome assumption adds structure and also limits the misreporting possibilities (as claiming output outside the support of the underlying random variable would be rather transparent misbehavior). The timing presumption interacts with the binary outcome assumption as well; that is, the mechanical interpretation of misreporting relies on the manager knowing something about second-period results before announcing first-period results. Otherwise, the reallocation between periods might be caught by the auditor. Of course, institutionally we know accounting reports are not timely; and that feature is used here to open the door to misreporting behavior.[28] The two-period horizon is also at work, as it removes any substantive concern over the best time to exercise a misreporting option.

Appendix

PROOF OF LEMMA. Consider the one period (Bernoullii structure) version of benchmark program [B], as used in the proof of Fact 1:

minimize $\Sigma_x J_x \omega(x)$
subject to: $\Sigma_x \mathrm{U}(J_x,H)\omega(x) \geq \mathrm{U}(W/2,0)$
$\Sigma_x \mathrm{U}(J_x,H)\omega(x) \geq \Sigma_x \mathrm{U}(J_x,L)\hat{\omega}(x)$.

In the extreme case of $\hat{s} = 0$, $\mathrm{U}(J_0,L) = \mathrm{U}(W/2,0)$, as $\hat{\omega}(0) = 1$. Moreover, $J = J_1 = J_2 > J_0$, and $(1-s)^2\mathrm{U}(J_0,H) + (1-(1-s)^2)\mathrm{U}(J,H) = \mathrm{U}(W/2,0)$.
With $\mathrm{U}(J_0,L) = \mathrm{U}(W/2,0)$, J_0 is a constant. Totally differentiating the above expression provides $2(1-s)[-\mathrm{U}(J_0,H) + \mathrm{U}(J,H)]ds + (1-(1-s)^2)\mathrm{U}'(J,H)dJ = 0$. $J > J_0$ implies the term in brackets is positive; hence, $dJ/ds < 0$.

This provides some $s' > s$ for which $2J_1 + J_2 + J_0 < 2(W + 2kH)$. Continuity implies this holds in a neighborhood of $\hat{s} = 0$ as well. ■

Proof of Proposition 1. Begin with the solution to [B], the $\beta(H,H) = \alpha = 0$ case. Denote the expected payment to the manager $\mathrm{E}[\hat{I}_{xy} \mid \alpha = 0]$. It is routine to verify this solution is feasible in the present case, as smoothing is strictly motivated and possible only under $a_1 = a_2 = H$. (In particular, supplying $a_1 = a_2 = H$ and smoothing whenever possible offers expected utility in excess of $\mathrm{U}(W,0,0)$ and this can be achieved with no other policy.) With $\hat{I}_{01} = \hat{I}_{10}$ and $\hat{I}_{02} = \hat{I}_{20}$, and the noted Bernoulli structure, the expected payment to the manager is $\mathrm{E}[\hat{I}_{xy} \mid \alpha = 0] + 2\alpha s^2(1-s)^2 (\hat{I}_{11} - \hat{I}_{02}) \equiv \mathrm{E}[\hat{I}_{xy} \mid \alpha = 0] + \theta(\hat{I}_{11} - \hat{I}_{02})$.

Now define a perturbed payment function by $\acute{I}_{xy} = \hat{I}_{xy} - \delta$, where $\delta = \theta(\hat{I}_{11} - \hat{I}_{02})$. This implies an expected payment of $\mathrm{E}[\hat{I}_{xy} \mid \alpha = 0]$. With constant absolute risk aversion, this is clearly incentive compatible in the present mechanism design program. The only question is whether it satisfies the individual rationality constraint. Dropping (without loss of generality) the $2kH$ personal cost term, the manager's equilibrium expected utility is $\mathrm{U}(W+2kH-\delta,0,0) + \theta[\mathrm{U}(\hat{I}_{11}-\delta,0,0) - \mathrm{U}(\hat{I}_{02}-\delta,0,0)]$.

Expanding $\mathrm{U}(\cdot,0,0)$ in a Taylor series about the point $W = W + 2kH$ and dropping higher order terms, the manager's equilibrium expected utility is

$$\mathrm{U}(W,0,0) + 0.5r^2\mathrm{U}(W,0,0)[\delta^2 + \theta(\hat{I}_{11} - \hat{I}_{02})(\hat{I}_{11} + \hat{I}_{02}) - 2\theta(\delta + W)(\hat{I}_{11} - \hat{I}_{02})] = \mathrm{U}(W,0,0) + 0.5r^2\delta\mathrm{U}(W,0,0)[\hat{I}_{02} + \hat{I}_{11} - \delta - 2W].$$

But $[\hat{I}_{02} + \hat{I}_{11} - \delta - 2W] < 0$ implies the manager's equilibrium expected utility strictly exceeds $\mathrm{U}(W,0,0)$. This implies a slight lowering of δ will strictly improve the principal relative to $\mathrm{E}[\hat{I}_{xy} \mid \alpha = 0]$, and remain feasible in program [R] for policy m^{S}.

Fact 2 implies $F^* < \mathrm{F}(\mathrm{m}^{\mathrm{T}})$. ■

Proof of Proposition 2. Consider the mechanism design program for $\mathrm{F}(\mathrm{m}^{\mathrm{S}})$, but with all incentive compatibility (IC) constraints removed except for the continuation input choice following $x = 1$ (i.e., $\mathrm{E}[\mathrm{U} \mid x = 1, a_2 = H, \mathrm{m}^{\mathrm{S}}] \geq \mathrm{E}[\mathrm{U} \mid x = 1, a_2 = L, \mathrm{m}^{\mathrm{S}}]$). For later reference, notice that, under the symmetric smoothing policy, the equilibrium probabilities for selected output reports are:

$\gamma_0 \equiv \text{prob}(1,0) = 3s(1-s)^3$;
$\gamma_1 \equiv \text{prob}(1,1) = 6s^2(1-s)^2$; and
$\gamma_2 \equiv \text{prob}(1,2) = 3s^3(1-s)$.

With $\gamma_0 + \gamma_1 + \gamma_2 = 3s(1-s) \equiv \tau$, let $\gamma = [\gamma_0/\tau, \gamma_1/\tau, \gamma_2/\tau]$.

With moving support, it is readily verified the solution to this modified program is feasible (and therefore optimal) in the parent problem. Moreover, the multiplier on the remaining IC constraint is strictly positive.

Next, the first order conditions reveal $I_{00} = I_{01} = I_{21} = I_{22} = W + 2kH$. Moreover, $(\gamma_0/\tau)U(I_{10},H,H) + (\gamma_1/\tau)U(I_{11},H,H) + (\gamma_2/\tau)U(I_{12},H,H) = U(W,0,0)$. The latter fact implies the risky payments, I_{1j}, $j = 0,1,2$, can be located by solving for the minimal risk premium (conditional on not being in a fully insured event), subject to $(\gamma_0/\tau)U(I_{10},H,H) + (\gamma_1/\tau)U(I_{11},H,H) + (\gamma_2/\tau)U(I_{12},H,H) = U(W,0,0)$ and the single IC constraint of $E[U|x = 1, a_2 = H, m^S] \geq E[U|x = 1, a_2 = L, m^S]$.

Now turn to the asymmetric policy, m^A. Here, the risky payments are associated with $x = 0$. And we have the following equilibrium probabilities for selected output reports:

$\upsilon_0 \equiv \text{prob}(0,0) = (1-s)^4$;
$\upsilon_1 \equiv \text{prob}(0,1) = 4s(1-s)^3$; and
$\upsilon_2 \equiv \text{prob}(0,2) = 3s^2(1-s)^2$.

Also, $\upsilon_0 + \upsilon_1 + \upsilon_2 = (1+2s)(1-s)^2 \equiv \zeta < \tau$. Denote $\upsilon = [\upsilon/\zeta_0\ \upsilon/\zeta_1\ \upsilon/\zeta_2]$. Moreover, the same structure emerges from the first order conditions. This implies the risky payments, I_{0j}, $j = 0,1,2$, can be located by solving for the minimal risk premium (conditional on not being in a fully insured event), subject to $(\upsilon_0/\zeta)U(I_{10},H,H) + (\upsilon_1/\zeta)U(I_{11},H,H) + (\upsilon_2/\zeta)U(I_{12},H,H) = U(W,0,0)$ and the appropriate single IC constraint. However, γ is readily shown to be a garbling of υ. So any compensation vector that is feasible for the symmetric case is also feasible (when appropriately permuted) for the asymmetric case, and carries a lower risk premium. Coupled with the fact $\zeta < \tau$ (as $s > 0.5$) we conclude $F(m^A) < F(m^S)$.

Finally, notice that $\gamma = [(1-s)^2, 2s(1-s), s^2]$. As this is the benchmark equilibrium vector, we conclude $F(m^S) < F^*$. ■

PROOF OF PROPOSITION 3. Assume [A6′] is present. If $\alpha > 0$ the solution to [B] is clearly infeasible in the $F(m^T)$ program. This implies contracting with constraints to remove any misreporting incentives results in an expected wage larger than F^*.

Contracting on smoothed output, policy m^S, is also relatively undesirable. Consider the design program [R] for policy m^S, but with all IC constraints removed except those that force $a_1 = a_2 = H$ in the presence of policy m^S. Let $z = [x, y_1, y_2]$ denote the somewhat condensed output vector. Smoothing policy m^S, now, transforms $z = [0,1,0]$ into $[1,0,0]$ with probability α, $z = [0,1,1]$ into $[1,0,1]$, with probability α, $z = [2,0,0]$ into $[1,1,0]$ with probability α, and $z = [2,0,1]$ into $[1,1,1]$

with probability α, regardless of the manager's effort supply policy. Let $\Phi(z)$ denote the output probability mass induced by some specific effort supply policy, and $\underline{\Phi}(z)$ that induced by the same policy coupled with policy m^S. With smoothing odds independent of the policy itself, we are able to write $\underline{\Phi}(z) = \sum_{\hat{z}} g(z|\hat{z})\Phi(\hat{z})$, where the garbling, $g(z|\hat{z})$, is independent of the effort policy. From here invoke the garbling argument in Proposition 13 of Grossman and Hart (1983). Thus a less constrained version of the m^S design program is an inferior contracting venue to that of benchmark [B].

A parallel argument applies to the remaining policies. ■

Endnotes

1. See, for example, Archibald 1967, Holthausen 1981, and Suh 1990.
2. Sunder (1976), for example, demonstrates successful efforts costing in the petroleum industry results in higher income variance, while full costing has higher serial correlation in the reported income series. Ryan (1995) highlights the book to market implications of this smoothing phenomenon.
3. Likewise, the presumed pattern can entail something as qualitative as the "big bath" or "window dressing," or be based on an explicit model of selected accruals, as tested in Dechow, Sloan, and Sweeney (1995) for example.
4. In particular, the manager will eventually have the option of moving some, but not all, of the output across periods. This is meant to be suggestive of exercising discretion as to when output is claimed, but only at the margin. The additive, binary structure is a convenient way to model this. In addition, it allows us to readily model the agent knowing some of the second-period output at the time of the first-period report. For example, early sales returns in the current year may be known at the time the prior year's financial report is finalized.
5. The manager does not have the ability to hide or consume excess output.
6. An alternative characterization is to envision period t's managerial act consisting of $a_t \in \{L,H\}$ and $b_t \in \{0,1\}$. $b_t = 1$ denotes managerial effort aimed at learning enough to possibly distort the self-report, while ignoring such activity amounts to setting $b_t = 0$. Motivating misreporting, then, reduces to motivating $b_1 = b_2 = 1$, etcetera. This also presumes, consistent with the explanation in the text, that the personal cost of $b_t = 1$ versus $b_t = 0$ is trivial. A richer model would, of course, entail a less modest trade off of effort aimed at the two activities.
7. I assume throughout that the probability structure in [A1] admits to a solution to this and subsequent design programs.
8. Let $L(x) = \hat{\omega}(x)/\omega(x)$ and consider the single period version of [B] where $\omega(x)$ has full support. Grossman and Hart's (1983) Proposition 9 ensures that in this setting $L(1) - L(0) \leq$ (resp. $\geq$) $L(2) - L(1)$ is equivalent to $J_1 - J_0 \geq$ (resp. $\leq$) $J_2 - J_1$.
9. The Bernoulli specification is important because it has a natural interpretation, and it also neutralizes the information content of y_1 from the manager's perspective (since y_1 carries no information about y_2, given knowledge of a_1). To verify the claim, consider a single period version of benchmark program [B]:
   minimize $\sum_x J_x \omega(x)$

   subject to: $\sum_x U(J_x,H)\omega(x) \geq U(W/2,0)$ and
   $\sum_x U(J_x,H)\omega(x) \geq \sum_x U(J_x,L)\hat{\omega}(x)$.

With λ and μ denoting the (strictly positive) multipliers on the two constraints, the first order condition for payment J_x is:

$$-(1/rU(J_x,H)) = \lambda + \mu - \mu\exp[rk(L - H)]\hat{\omega}(x)/\omega(x).$$

Rescaling provides $\exp[rJ_x] = \gamma - \rho\hat{\omega}(x)/\omega(x)$. From here, the following inequality emerges, as $(s - \hat{s})^2 > 0$:

$$\exp[rJ_0 + rJ_2] = [\gamma - \rho(1 - \hat{s})^2/(1 - s)^2][\gamma - \rho\hat{s}^2/s^2] < \exp[2rJ_1] = [\gamma - \rho\hat{s}(1 - \hat{s})/s(1 - s)]^2.$$

10. An equivalent interpretation, suggested in an oral communication from Bharat Sarath and Shiva Sivaramakrishnan, stresses a multitask view and the idea that, with the proper supply of inputs, the manager can produce a pair of output random variables that has, for example, a smoothed appearance. The story stressed in the text is a mechanical version of just this story. Also see DeFond and Park 1997.
11. Further note, as becomes clear in the proof, the argument identifies a feasible incentive structure improvement, based on well-motivated smoothing; it does not rest on identifying the optimal incentive structure. Rather, the noted compensation insurance effect is exploited to identify an improvement. Moreover, the steepness condition itself is not necessary for a contracting gain. For example, using $s = 0.5$ in the forthcoming numerical illustration provides a setting where $2J_1 + J_2 + J_0 < 2(W + 2kH)$ fails, yet smoothing with probability $\alpha > 0$ can be profitably encouraged.
12. An alternative reporting strategy in the presence of the strong auditor is to use the (borrow) option only under $x = 0$ or the (loan) option only under $x = 2$. Of course, in the Bernoulli setting, this merely serves to lessen the statistical difference between misreported and off-equilibrium observables.
13. Conversely, [A6″] surfaces if we (drop the Bernoulli assumption and) set $\omega = [0.1, 0.1, 0.8]$ and $\hat{\omega} = [0.5, 0.4, 0.1]$. The comparable one-period program has $J_0 = 13{,}445$, $J_1 = 15{,}885$, and $J_2 = 21{,}977$.
14. In particular, if m^A is motivated, the only binding IC constraint will be for input $a_2 = H$, following $x = 0$; likewise, under policy m^S the only binding IC constraint will be for input $a_2 = H$, following $x = 1$.
15. It is also worth noting the importance of communication restrictions. It is evident, especially in the setting of Proposition 2, that full communication is dominating. To see this, stay with the asymmetric policy and suppose, under obedient behavior, the manager sees $x = 0$ followed by $y_1 = 0$. The manager can now communicate with an enlarged message space the claim that $x = 0$ and $y \leq 1$. Such a claim is not sustainable under disobedient behavior, and this opens the door to improved risk sharing.
16. Equally clear, the usual continuity argument suggests the Proposition 1 result does not rest exclusively on the Bernoulli structure.
17. Indeed, it is also possible, under [A1], to have the second period $\boldsymbol{h}^a$ differ between $(x_1,x_2) = (1,0)$ and $(0,1)$. To illustrate, let $\omega = [0.1, 0.1, 0.8]$ and $\hat{\omega} = [0.5, 0.4, 0.1]$. Also set the continuation probabilities at $\boldsymbol{h}^H = [0.1, 0, 0.1, 0.8]$ under $x = 0$ and $\boldsymbol{h}^H = [0.1, 0.1, 0, 0.8]$ under $x = 1$. So, any misreporting possibility removes considerable ambiguity as to the remaining output, an admittedly extreme case. [A1] and [A6′] are readily verified. Under [A7′], now, the desirability of smoothing in conjunction with the strong auditor is ambiguous. Small α favors motivating m^T over m^S, while a larger α favors the opposite. In particular, under $\alpha = 0.02$ I find $F(m^S) = 41{,}036$, and under $\alpha = 0.10$ I find $F(m^S) = 40{,}996$, while $F(m^T) = 41{,}029$.

18. To reinforce this theme, notice that policy m^S might also be used to hide poor performance, say via $\beta(L,L) = \alpha$ and $\beta(H,L) = \beta(L,H) = \beta(H,H) = 0$. In this case, the principal strictly prefers to avoid any access to the m^S technology on the part of the manager.
19. Respective numerical values are 40,530, 40,518, 40,505, and 40,504. The usual continuity argument can be used to show $F(m^S) < F(m^T)$ for small α in the unlimited misreporting case. In a related vein, Evans and Sridhar (1996; 1997) provide a model in which the manager's ability to manipulate the interim performance measure is linked to the internal control system. The principal, in turn, prefers to allow manipulation when the odds that the internal control system is functioning are high, a result that parallels the fact $F(m^S) < F(m^T)$ for small α. The intuition is the same across the settings. The model here relies on the ability to smooth being linked to the manager's input supply.
20. Hepworth (1953) and Gonedes (1972), for example, impose (exogenous) smoothing preferences and then explore the resulting smoothing behavior. More broadly, Ronen and Sadan (1981) stress classificatory and intertemporal smoothing, in the setting of an abstract accounting technique applied to a stochastic process with an exogenous incentive structure.
21. Healy (1985), McNichols and Wilson (1988), DeFond and Jiambalvo (1994), Perry and Williams (1994), Gaver, Gaver, and Austin (1995), Holthausen, Larcker, and Sloan (1995) and DeFond and Park (1997) illustrate various approaches. Dechow, Sloan, and Sweeney (1995) examine the power of these types of tests across a variety of hypothesized processes and environments; Kang and Sivaramakrishnan (1995) also examine these tests, and propose an instrumental variables approach. Ahmed (1996), Beaver and Engel (1996), Hunt, Moyer, and Shevlin (1996), and Subramanyam (1996) examine the earnings multiple or implied pricing of discretionary accruals.
22. This approach also raises questions about the efficacy of the Jones model in these settings. Guay, Kothari, and Watts (1996) and Kang and Sivaramakrishnan (1995), for example, document empirical difficulty in isolating discretionary accruals. Also see Bernard and Skinner (1996). Nevertheless, we should remember the trade off between micro and aggregate specification when the micro relationships are not well understood (e.g., Grunfeld and Griliches 1960).
23. For example, Trueman and Titman (1988) provide a one-period signaling model in which the privately informed firm incurs an additional cost associated with the variance of its income. In equilibrium, smoothing (when possible) is motivated, as it lowers the debt-holders' assessment that the firm is a high variance type. Verrecchia (1986;1990) emphasizes the discretionary aspect of smoothing in a model where the manager knows economic earnings but can issue a report within a privately known range of economic earnings. Dye (1988) stresses limited communication, a forced sale in the capital market (due to intergenerational transfers) and incomplete markets. He distinguishes internal (labor market) from external (capital market) demands for earnings management. Consumption smoothing also interacts with communication, as in Christensen and Feltham 1993. Evans and Sridhar (1996) continue this theme, absent the intergenerational sale, in a setting where the manager privately learns whether earnings management is possible. The possibility no such option is present, and is also independent of the underlying activity or performance, provides structure that can be used to identify conditions where earnings management is preferred to motivating rejection of earnings management.

Boylan and Villadsen (1996) offer an infinitely repeated platform, where smoothing is driven by consumption smoothing and bankruptcy concerns. Evans and Sridhar (1997) emphasize an exogenous probability, linked to control-system failure, that the manager will be able to manipulate an interim report, and interact this possibility with insider trading. Chaney and Lewis (1995) analyze a two-period setting with private-type information, tax frictions, and linear contracts. Arya, Glover, and Sunder (1998) emphasize lack of commitment, and the possibility misreporting may be efficient behavior. Fudenberg and Tirole (1995) also emphasize lack of commitment, linking smoothing incentives to a manager who receives incumbency rents, a lack of long-term contracts, and information "decay."

24. Stated differently, the central feature is lack of a revelation representation. This is stressed by Dye 1988 and Schipper 1989.
25. Naturally, presuming statistical dependence between the two outputs in the absence of smoothing would imply $\hat{x}$ carries information about y in both cases. But the basic point that what information is carried will depend on whether smoothing is possibly present, in equilibrium, will remain. By analogy, empirical work exhibits a dependency between variance and the valuation implications of the earnings report (e.g., Ahmed 1996 or Hunt, Moyer, and Shevlin 1996). The same dependency occurs here. The reason is not that output is of higher or lower quality, as, for example, in the case of Chaney and Lewis 1995. Rather, it is the fact the information content of the initial report is being affected by the possibility of smoothing behavior. While smoothing lowers the quality of the valuation information in our case, it also creates a statistical dependence across the periods.
26. Depending on additional details, performance may, or may not, be smoothed. To the extent we have observed a sample from such a population, we would be mixing cases where misreporting is present and beneficial, misreporting is present and harmful, misreporting is present but not motivated, and any misreporting option is absent or irrelevant. But in all cases, without enlarged communication possibilities, the presence and use of the misreporting option would remain unobservable. Rather, the distinguishing feature would be different stochastic properties of the firms' performance series. This explanation provides a more subtle view of the empirical counterpart than implied by, for example, Kang and Sivaramakrishnan's (1995) simulation experiment.
27. Sappington and Weisman (1996) examine difficulties, including performance misreporting, in empirically identifying the effects of incentive-based regulation in the telecommunications industry.
28. Thus, one variation of the model would argue for reporting delay just as another would argue for no delay. For example, if the time at which the first-period report is due were subject to choice, early reporting would be advantageous in the setting of Proposition 3, just as delayed reporting would be advantageous in the setting of Proposition 1.

References

Ahmed, A. 1996. Discretionary accruals, earnings management and the valuation of earnings. Working paper, Syracuse University.

Archibald, T. R. 1967. The return to straight-line depreciation: An analysis of a change in accounting method. *Journal of Accounting Research* 5 (Supplement): 164–80.

Arya, A., J. Glover, and S. Sunder. 1998. Earnings management and the revelation principle. Working paper, Ohio State.

Bartov, E. 1993. The timing of asset sales and earnings manipulation. *Accounting Review* 68 (October): 840–55.

Beaver, W., and E. Engel. 1996. Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. *Journal of Accounting and Economics* 22 (August–December): 177–206.

Bernard, V., and K. Skinner. 1996. What motivates managers' choice of discretionary accruals? *Journal of Accounting and Economics* 22 (August–December): 313–25.

Boylan, R., and B. Villadsen. 1996. An algorithmic approach to contracting in an infinite agency model. Working paper, Iowa State.

Chaney, P., and C. Lewis. 1995. Earnings management and firm valuation under asymmetric information. *Journal of Corporate Finance* 1: 319–45.

Christensen, P., and G. Feltham. 1993. Communication in multiperiod agencies with production and financial decisions. *Contemporary Accounting Research* 9 (Spring): 706–44.

Dechow, P., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *Accounting Review* 70 (April): 193–225.

DeFond, M., and J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics* 17 (January): 145–76.

DeFond, M., and C. Park. 1997. Smoothing income in anticipation of future earnings. *Journal of Accounting & Economics* 23 (July): 115–39.

Dye, R. 1988. Earnings management in an overlapping generations model. *Journal of Accounting Research* 26 (Autumn): 193–235.

Evans, J., and S. Sridhar. 1996. Multiple control systems, accrual accounting, and earnings management. *Journal of Accounting Research* 34 (Spring): 45–65.

———. 1997. Firm characteristics, management of performance measures and market price efficiency. Working paper, Pittsburgh State.

Fellingham, J., D. Newman, and Y. Suh. 1985. Contracts without memory in multiperiod agency models. *Journal of Economic Theory* 37 (December): 340–55.

Fudenberg, D., and J. Tirole. 1995. A theory of income and dividend smoothing based on incumbency rents. *Journal of Political Economy* 103 (1): 75–93.

Gaver, J., K. Gaver, and J. Austin. 1995. Additional evidence on the association between income management and earnings-based bonus plans. *Journal of Accounting & Economics* 19 (February): 3–28.

Gonedes, N. 1972. Income-smoothing behavior under selected stochastic processes. *Journal of Business* 45 (4): 570–84.

Grossman, S., and O. Hart. 1983. An analysis of the principal-agent problem. *Econometrica* 51 (January): 7–46.

Grunfeld, Y., and Z. Griliches. 1960. Is aggregation necessarily bad? *Review of Economics and Statistics* 42 (February): 1–13.

Guay, W., S. Kothari, and R. Watts. 1996. A market-based evaluation of discretionary-accrual models. *Journal of Accounting Research* 34 (Supplement): 83–115.

Hand, J. 1989. Did firms undertake debt-equity swaps for an accounting paper profit or true financial gain? *Accounting Review* 64 (October): 587–623.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7 (April): 85–107.

Hepworth, S. 1953. Smoothing periodic income. *Accounting Review* 28 (January): 32–9.

Holthausen, R. 1981. Evidence on the effect of bond covenants and management compensation contracts on the choice of accounting techniques: The case of the depreciation switch-back. *Journal of Accounting and Economics* 3 (March): 73–109.

Holthausen, R., D. Larcker, and R. Sloan. 1995. Annual bonus schemes and the manipulation of earnings. *Journal of Accounting and Economics* 19 (February): 29–74.

Hunt, A., S. Moyer, and T. Shevlin. 1996. Earnings volatility, earnings management, and equity value. Working paper, University of Washington.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (Autumn): 193–228.

Kang, S., and K. Sivaramakrishnan. 1995. Issues in testing earnings management and an instrumental variable approach. *Journal of Accounting Research* 33 (Autumn): 353–67.

Lambert, R. 1984. Income smoothing as rational equilibrium behavior. *Accounting Review* 59 (October): 604–18.

McNichols, M., and G. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

Perry, S., and T. Williams. 1994. Earnings management preceding management buyout offers. *Journal of Accounting and Economics* 18 (September): 157–79.

Ronen, J., and S. Sadan. 1981. *Smoothing income numbers: Objectives, means, and implications*. Reading, MA: Addison-Wesley.

Ryan, S. 1995. A model of accrual measurement with implications for the evolution of the book-to-market ratio. *Journal of Accounting Research* 33 (Spring): 95–112.

Sappington, D., and D. Weisman. 1996. Potential pitfalls in empirical investigations of the effects of incentive regulation plans in the telecommunications industry. *Information Economics and Policy* 8 (June): 125–40.

Schipper, K. 1989. Commentary on earnings management. *Accounting Horizons* 3 (December): 91–102.

Subramanyam, K. 1996. The pricing of discretionary accruals. *Journal of Accounting and Economics* 22 (August–December): 249–81.

Suh, Y. 1990. Communication and income smoothing through accounting method choice. *Management Science* (June): 704–23.

Sunder, S. 1976. Properties of accounting numbers under full costing and successful efforts costing in the petroleum industry. *Accounting Review* 51 (January): 1–18.

Trueman, B., and S. Titman. 1988. An explanation for accounting income smoothing. *Journal of Accounting Research* 26 (Supplement): 127–39.

Verrecchia, R. 1990. Information quality and discretionary disclosure. *Journal of Accounting and Economics* 12 (March): 365–80.

———. 1986. Managerial discretion in the choice among financial reporting alternatives. *Journal of Accounting and Economics* 8 (October): 179–95.