A Principle of Classification

Abstract

We study a firm's decision to classify transactions as recurring or nonrecurring in a setting with no fixed classification scheme, but with the following principle: transactions classified as recurring must be more persistent than those classified as nonrecurring. This principle corresponds to existing classification standards. We find that the firm's optimal classification strategy has a simple form: maximize the product of the (absolute) total of income-reducing nonrecurring and the total income-increasing recurring items. We characterize the possible firm values consistent with a report, and provide a measure of how opaque a firm's valuation is given its classification choice.

Keywords—classification shifting; opacity; principles; signaling

1. Introduction

Income statement classifications inform investors about a firm's earnings persistence. Investors attach a higher present value factor to recurring revenues and expenses than to one-time gains or losses, restructuring charges, or other items presented as transitory.

Despite the importance of classification in determining a firm's fundamental value, neither US Generally Accepted Accounting Principles (GAAP) nor International Financial Reporting Standards (IFRS) specify a fixed rule for classifying transactions as recurring or nonrecurring. Instead, standards on classification provide a general principle. Loosely expressed, this principle is that the items considered recurring should be more persistent than those considered nonrecurring. Our purpose is to explain what a firm can credibly reveal about the persistence of its income and about its fundamental value under this classification principle.[1]

To that end, we characterize a firm's optimal classification strategy, and then describe the degree to which firms with different values pool by making identical classifications. This tells us how classifications endogenously vary with income characteristics, and how opaque optimally classified income statements are.

Without definitive guidance, any firm can privately pick its own threshold for what counts as recurring without violating the classification principle. The firm has incentive to do so in a way that enables it to report positive transactions as recurring revenues and neg-

[1]The principle we study is an abstraction, but reflects a common theme in standard-setting practice. For instance, FASB [1985, CON6–24] states "Revenues and gains are similar, and expenses and losses are similar, but some differences are significant in conveying information about an enterprise's performance. Revenues and expenses result from an entity's ongoing major or central operations. . . . In contrast, gains and losses result from incidental or peripheral transactions." Similar distinctions are not difficult to find; however, the Accounting Standards Codification (ASC) Master Glossary does not give specific definitions of these terms, leaving classification as a judgment call (see PricewaterhouseCoopers [2019, 3-10] and the related discussion on ASC 605 about gains versus revenues).

ative ones as nonrecurring losses. With so much flexibility, it is natural to wonder whether investors can reliably estimate the persistence of different line items. We argue, however, that investors can learn a considerable amount of information from the classifications.

The reason is that the classifications reveal information about the ranking of transaction persistence. Investors learn that those transactions classified as a highly persistent line item have present value factors greater than that of transactions classified as highly transitory items. To the extent that classifications reveal the persistence ranking, investors learn about present value factors from their order distribution (that is, the prior distribution of the sample order statistics of present value factors).

Our first result shows that income statement classifications could fully reveal a firm's expected value conditional on the persistence ranking of all the firm's transactions.[2] As long as the firm has at least two possible classifications, it can comply with the principle and have a distinct report for every possible valuation based on its persistence ranking. A firm with four income statement line items—say, recurring revenues, recurring expenses, nonrecurring gains, and nonrecurring losses—has a communication channel that could express as much information about firm value as one that lists every dollar it receives or spends in order of persistence.

The amount investors actually learn from classifications is more limited. As each firm can choose its own cutoff for what counts as recurring, it can classify its transactions in many ways, all of which comply with the standard. For instance, a firm for which every income-reducing transaction is more persistent than every income-increasing transaction can classify all its transactions as recurring. To make its classifications informative, a firm

[2] In the terminology of Nezlobin [2012, 261], the firm's message space is *informationally sufficient* for the firm value that investors would assign if they could rank all of the firm's transactions in order of persistence. There is a large literature on informational sufficiency; examples include Mount and Reiter [1974], Sonnenschein [1974], Calsamiglia [1977], Reichelstein and Reiter [1988], and Jordan and Xu [1999].

needs to consider whether less valuable firms can mimic it.

We address this problem by modeling investors as anticipating the firm's incentives and reading reports skeptically. Firms respond by choosing their thresholds to maximize the worst-case interpretation of their reports, as is standard in models of signaling to a skeptical receiver (for detailed exposition, see Milgrom and Roberts [1986], Okuno-Fujiwara et al. [1990], and Shin [1994, 2003]). This response to skepticism leads to our second result, in which we provide the firm's optimal classification strategy. This strategy has a remarkably simple form: choose the cutoff to maximize the product of the number of nonrecurring losses and recurring revenues.

Having established this optimal strategy, we can characterize the degree to which optimally classified income statements pool firms of different types. As our third result, we provide an explicit formula for the number of possible persistence rankings a firm would optimally report as a given income statement.[3] The formula has a convenient decomposition into the product of the number of rankings for transactions below the firm's cutoff and the number of rankings for those above the cutoff, enabling us to evaluate the nonrecurring and recurring transactions separately.

Each term in this decomposition depends on both the reported recurring revenues and nonrecurring losses. This implies that we should not expect classifications to be independent of income, or to see the persistence of recurring revenue to be independent of the value of nonrecurring losses. These results are consistent with a large empirical literature, dating back to the 1970s (examples include Gonedes [1975], Ronen and Sadan [1975], and

[3]This question is in the tradition of early work the amount of information lost through aggregation of financial reports (some examples are Ijiri [1967, in particular Appendix B], Ijiri [1968], and Lev [1968]). Other work in this stream includes Kanodia and Mukherji [1996] on the real effects of aggregation of cash flow classifications, Anctil et al. [1998] on the informational sufficiency of activity-based costing systems for identifying residual income maximizing projects, Arya et al. [2000] on the ability to recover transactions from changes to line items, and Dye and Sridhar [2004] on informational gains from aggregate reporting arising from reduced incentives to manipulate reports.

Barnea et al. [1976]), with a considerable renewed interest beginning with McVay [2006]. Although the empirical literature interprets these patterns as evidence of misreporting, our characterization shows they would arise for firms that fully comply with classification standards.[4]

Our fourth result complements this characterization by describing the effects of classifications on firm valuation (rather than on line item persistence). We provide a formula for the smallest and largest optimal firm values associated with a given report, and show that every firm value between these bounds shares the same optimal report. The range of possible values gives us a measure of the opacity of the report.

By providing this measure of classification opacity, our results contribute to ongoing policy discussions over both reporting classifications and non-GAAP reporting. The Securities and Exchange Commission has expressed concern about whether a firm's non-GAAP earnings classification choices mislead or clarify (see Donelson et al. [2020]). Prior academic empirical work on this topic provides inconclusive results. Bradshaw and Sloan [2002] and Doyle et al. [2003] argue that non-GAAP earnings conceal information from the market by conflating what is transitory and what is persistent; Brown and Sivakumar [2003], Gu and Chen [2004], and Ribeiro et al. [2019] argue instead that non-GAAP earnings inform the market. Their discussion appears difficult to settle: as shown in Abarbanell and Lehavy [2007], there is little difference in predicted behavior under either hypothesis.

The structure of the rest of this paper is as follows: Section 2 introduces the model and provides preliminary results. Section 3 shows our main results. We interpret these and conclude in Section 4. All proofs are in an appendix.

[4]For results on the interpretation of patterns often attributed to earnings manipulation, see Breuer and Windisch [2019], Hemmer and Labro [2019], Hiemann [2020]. For discussion of the difficulties of using line items to measure persistence, see Amir et al. [2013].

2. Preliminaries

2.1. Model setting

We analyze a game between two risk neutral players with rational expectations. A firm's manager (henceforth called the firm) issues a report as described below, and aims to maximize the firm's share price. A representative investor prices the firm at its expected net present value, given the firm's report and a market discount rate $\rho > 0$.

The firm bases its report on its information, consisting of n ordered pairs:

$$\{(x_i, \alpha_i)\}_{i=1}^{n}$$

Each x_i corresponds to a transaction, in the amount of 1 or -1, which the firm must include in a line item on its income statement. The α_i associated with x_i is its present value factor, in the form of an annuity due. That is, α_i includes present period income, and therefore ranges over $[1, (1+\rho)/\rho]$. As our interest is in the firm's choice of cutoff, we allow each α_i to vary continuously. To keep the focus on what the firm can communicate through its classification decisions, we assume that the firm has no credible way to reveal the α_i through a voluntary disclosure.[5]

The representative investor does not know the realized $\{\alpha_i\}_{i=1}^{n}$, but views them as random variables $\{\widetilde{\alpha}_i\}_{i=1}^{n}$. We assume a commonly known uniform prior; that is, for $i \neq j$,

[5]For discussion, see Dye and Sridhar [2008, 314] and Nezlobin [2012, 235]. Additionally, it is known that voluntary disclosure with multiple dimensions does not lead the unravelling result of Grossman [1981] and Milgrom [1981]; see, for example, Shin [2003] and Ebert et al. [2017]. Potential interactions between voluntary and mandatory disclosures are studied in Bagnoli et al. [2001] in an auditing context and in Dye [1985], Einhorn [2005], and Bertomeu et al. [2021] in a financial reporting setting.

$\widetilde{\alpha}_i$ is independent of $\widetilde{\alpha}_j$ with a uniform marginal distribution:

$$(\forall i \in \{1, \ldots, n\}) \quad \widetilde{\alpha}_i \overset{iid}{\sim} U\left[1, \frac{1+\rho}{\rho}\right] \tag{1}$$

The investor can infer the realized $\{x_i\}_{i=1}^n$ from the firm's income statement, so their prior distribution is irrelevant unless it provides information about the present value factors. We rule this out, and assume that each $\widetilde{\alpha}_i$ is independent of the transactions $\{\widetilde{x}_i\}_{i=1}^n$.

The firm faces a classification problem. Specifically, it publicly releases an income statement, which decomposes its total income-increasing transactions (U, for *up*) into recurring revenues (R) and nonrecurring gains ($G := U - R$), and decomposes its total income-reducing transactions (D, for *down*) into nonrecurring losses (L) and recurring expenses ($E := D - L$). We represent the firm's income statement as (D, U, L, R), and adopt the convention of writing income-reducing totals as positive amounts. This lets us write the current period net income as $\pi := U - D$ and the total number of transactions n as $U + D$.

An unmodeled authority, such as an auditor or regulator, restricts the firm's classification choice, according to the following reporting principle: the firm must select a unique cutoff $\widehat{\alpha} \in [1, (1+\rho)/\rho]$, so that its classifications satisfy (2):

$$R = R(\widehat{\alpha}) = \sum_{\{i | \alpha_i \geq \widehat{\alpha}\}} \{x_i | x_i > 0\} \qquad L = L(\widehat{\alpha}) = \sum_{\{i | \alpha_i < \widehat{\alpha}\}} \{-x_i | x_i < 0\} \tag{2}$$

The values of U and D are independent of the firm's chosen classification cutoff:

$$U = \sum_{i=1}^{n} \max\{x_i, 0\} \qquad D = \sum_{i=1}^{n} \max\{-x_i, 0\}$$

The requirements in (2) serve two purposes. First, the firm's reporting must be monotone in persistence: if $\alpha_i \leq \alpha_j$, the firm cannot treat x_i as recurring and x_j as nonrecurring.

Any transactions called recurring must be more persistent than any called nonrecurring. Second, this monotonicity with respect to the present value factor α_i is independent of whether x_i is positive or negative. What counts as recurring or nonrecurring does not change based on whether a transaction increases or reduces income.[6]

The investor observes the firm's report, forms a conjecture α_M about the firm's cutoff, and updates beliefs. We write the investor's firm value v as

$$v(D,U,L,R;\alpha_M) = E\left[\sum_{i=1}^{n} \widetilde{\alpha}_i x_i | \alpha_M; D,U,L,R\right] \tag{3}$$

The firm anticipates the investor's conjecture α_M as an implicit function of its report, and optimally chooses its cutoff $\widehat{\alpha}$:

$$\max_{\widehat{\alpha}\in\left[1,\frac{1+\rho}{\rho}\right]} E\left[v(D,U,L(\widehat{\alpha}),R(\widehat{\alpha});\alpha_M(D,U,L(\widehat{\alpha}),R(\widehat{\alpha}))\right] \tag{4}$$

Figure 1 summarizes the timeline.

Figure 1 about here.

2.2. Valuation using persistence ranking

Because of the principle in (2), the firm's report reveals the total number of transactions below and above the firm's cutoff ($L+U-R$ and $R+D-L$, respectively). The investor learns the cutoff from the report, and uses the information in the report to estimate the

[6] A weaker requirement could impose conservatism, under which the firm chooses a lower cutoff for its income-reducing transactions than for its income-increasing ones (similar to Gigler et al. [2009]). We do not impose conservatism here, because conservatism would automatically generate a relationship between transaction signs and classifications, obscuring the fact that this relationship arises endogenously.

ranking of the firm's transactions in order of their persistence. Before proceeding to our main results, we make some observations on using relative persistence in valuation.

We first introduce some notation. For $i \in \{1, \ldots, n\}$, let

$$\widetilde{\alpha}_{(i)} = i^{\text{th}} \text{ order statistic of } (\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_n),$$

i.e., $\widetilde{\alpha}_{(i)}$ is the i^{th}-lowest draw among the present value factors. Let $x_{(i)}$ be the transaction that corresponds to $\widetilde{\alpha}_{(i)}$. Without loss of generality, the firm can choose its cutoff $\widehat{\alpha}$ as $\widetilde{\alpha}_{(i)}$ for some $i \in \{1, \ldots, n\}$. Because each $\widetilde{\alpha}_i$ is an independent uniform draw from $[1, (1+\rho)/\rho]$, it follows that

$$E[\widetilde{\alpha}_{(i)}] = 1 + \frac{i}{\rho(n+1)}$$

Thus, in expectation, the n ordered uniform draws chop the interval $[1, 1 + 1/\rho]$ into $n + 1$ equally long sub-intervals, each of length $1/(\rho(n + 1))$. Figure 2 illustrates the case with $n = 4$.

Figure 2 about here.

If the report reveals the persistence ranking of the transactions, then the investor prices the firm at

$$\begin{aligned} v^* = \sum_{i=1}^{n} E[\widetilde{\alpha}_{(i)}]x_{(i)} &= \sum_{i=1}^{n} x_{(i)} \left[1 + \frac{i}{\rho(n+1)}\right] \\ &= \underbrace{\sum_{i=1}^{n} x_{(i)}}_{\text{Net income}} + \frac{1}{\rho(n+1)} \sum_{i=1}^{n} i x_{(i)} \\ &= \pi + \frac{1}{\rho(n+1)} \sum_{i=1}^{n} i x_{(i)} \qquad (5) \end{aligned}$$

As n increases, the mean-squared difference between v^* and the firm's true value approaches an asymptotic bound from below. We state this as follows:

PROPOSITION 1. *If a report* (D, U, L, R) *fully reveals the persistence ranking, then the mean-squared error in the investor's estimate of the firm value, conditional on the report, is*

$$E\left[\left(v^* - \sum_{i=1}^{n} \widetilde{\alpha}_i x_i\right)^2\right] = \frac{n}{6\rho^2(n+1)} \tag{6}$$

Thus, the mean-squared error of v^* *asymptotically approaches its upper bound of* $1/(6\rho^2)$.

To understand Proposition 1, refer again to Figure 2. As n increases, the distance between $E[\widetilde{\alpha}_{(i)}]$ and $E[\widetilde{\alpha}_{(i+1)}]$ shrinks. This increase in precision offsets the additional noise associated with more transactions.

We close this section by introducing some definitions that will help us reduce the amount of notation in the analysis below. Given the persistence-ranked transactions $\{x_{(i)}\}_{i=1}^{n}$, define $y = (y_1, \ldots, y_n)$ by

$$(\forall i \in \{1, \ldots, n\}) \quad y_i := \frac{x_{(i)} + 1}{2}$$

Thus, $y_i = 1$ if $x_{(i)} = 1$ and $y_i = 0$ if $x_{(i)} = -1$.

Given y (representing the persistence-ranked transactions as a bit string), we define the sum of the positions of income-increasing transactions as

$$w(y) := \sum_{i=1}^{n} i y_i \tag{7}$$

This definition gives us an alternative way to express v^*. From the inverse transformation

$$x_{(i)} = 2y_i - 1$$

and (7), we can rewrite (5) as

$$
\begin{aligned}
v^* &= \pi + \frac{1}{\rho(n+1)} \sum_{i=1}^{n} i x_{(i)} = \pi + \frac{1}{\rho(n+1)} \sum_{i=1}^{n} i \left(2y_i - 1\right) \\
&= \pi + \frac{1}{\rho(n+1)} \left[\sum_{i=1}^{n} (-i) + 2 \sum_{i=1}^{n} i y_i \right] \\
&= \boxed{\pi - \frac{n}{2\rho} + \frac{2w(y)}{\rho(n+1)}} \qquad (8)
\end{aligned}
$$

It is clear from Equation (8) that v^* depends on the firm's classification decision only through $w(y)$, as π and n are observable directly from D and U, and the firm cannot control the market discount rate ρ.

3. Results

3.1. Benchmark: first-best

We have seen that an investor who can infer the persistence ranking of the transactions can price the firm accurately, in the sense that the pricing error in v^* is bounded by a fixed constant regardless of n. It is also clear that the report can reveal at most the persistence ranking (along with D and U, which the report reveals regardless of the classification decision). Therefore, the most that an investor can hope to learn from the report is v^*, making v^* the appropriate first-best benchmark.

Our first main result shows that the set of possible reports (D, U, L, R) that conform to the principle (2) is large enough to enable the investor to price the firm at v^*. This result may appear surprising. Even if the firm did not have to conform to (2), it would have only 2^n possible reports, compared with $n!$ possible persistence orderings. Since $n! > 3^{n-2}$, the number of possible orderings grows much faster than the number of possible reports.

The reason a four line-item report suffices for valuing the firm at v^* is that many distinct persistence rankings generate the same firm value. From (8), we see that, given U, D,and ρ, the sum $w(y)$ of the rankings of the income-increasing transactions determines v^*. It follows that the investor can determine v^* if and only if the report (D, U, L, R) is a sufficient statistic for $w(y)$. Theorem 1 shows that, for firms conforming to the principle in (2), this is always the case.

THEOREM 1. *There are enough reports* (D, U, L, R) *satisfying* (2) *to communicate every possible value of* v^*. *In particular, given* U *positive transactions and* D *negative ones, there are*

$$D \cdot U + 1 \text{ possible values of } v^*, \text{ and}$$

$$D \cdot U + 1 + n \text{ possible reports.}$$

3.2. OPTIMAL CLASSIFICATION

If the firm could use the expressiveness of the income statement that Theorem 1 establishes, then its report would fully separate it from other firm types. This can occur only if a firm type with a lower value of v^* could not mimic the firm's report. As a preliminary step toward finding the firm's optimal classification strategy, we rule out complete separation.

PROPOSITION 2. *Let* $v^*(y)$ *be the value of* v^* *associated with a given bit string* y. *If* $\min\{D, U\} \leq 1$, *then for every possible* y *with* $\sum_{i=1}^{n} y_i = U$ *and with length equal to* $U + D$, *there exists a report* (D, U, L, R) *such that*

1. *Reporting* y *as* (D, U, L, R) *is consistent with* (2), *and*

2. *For every bit string* y' *with* $\sum_{i=1}^{n} y'_i = U$ *and length* $U + D$ *that the firm can report as* (D, U, L, R), $v^*(y) \leq v^*(y')$

If $\min\{D, U\} > 1$*, then there exists a bit string* $\widehat{y}$ *with* $\sum_{i=1}^{n} \widehat{y}_i = U$ *and with length* $U + D$ *such that, for every report* (D, U, L, R) *of* $\widehat{y}$ *that satisfies* (2)*, there exists some* y' *with* $\sum_{i=1}^{n} y'_i = U$*, length of* $U + D$*, and* $v^*(\widehat{y}) > v^*(y')$ *that the firm can also report as* (D, U, L, R)*.*

Proposition 2 implies there cannot be an equilibrium in which every firm type reports differently, that is, in which the firm necessarily reveals $v^*(y)$. Accordingly, following Shin [1994, 2003], we look for an optimal sanitization strategy for the firm (i.e., in which the firm maximizes its worst-case valuation).

The worst-case valuation is easily characterized. Given the report (D, U, L, R), the most pessimistic interpretation is that firm's least persistent transactions are its $G = U - R$ gains, followed by its L losses, then its R revenues, and finally its $E = D - L$ expenses. The bit string $\underline{y}(D, U, L, R)$ corresponding to this interpretation is

$$\underline{y}(D, U, L, R) = \left(\underbrace{1, \ldots, 1}_{U-R}, \overbrace{0, \ldots, 0}^{L}, \underbrace{1, \ldots, 1}_{R}, \overbrace{0, \ldots, 0}^{D-L} \right)$$

Our second main result, Theorem 2, provides the optimal sanitization strategy.

Theorem 2. *Let*

$$\underline{w}(D, U, L, R) = w\left(\underline{y}(D, U, L, R)\right)$$

be the investor's most pessimistic interpretation of the firm's report. Then

$$\underline{w}(D, U, L, R) = LR + \binom{U+1}{2} \tag{9}$$

and the corresponding level of v^* *is*

$$\underline{v} = v^*(\underline{y}) = \pi - \frac{n}{2\rho} + \frac{U(U+1) + 2LR}{\rho(n+1)} \tag{10}$$

From (10)*, we see that only* LR *depends on the firm's classification decision. Therefore, the firm maximizes* $\underline{v}$ *if and only if the firm chooses its cutoff* $\widehat{\alpha}$ *in* (4) *to maximize*

$$L(\widehat{\alpha}) \cdot R(\widehat{\alpha}) \tag{11}$$

As the firm's objective (11) has a simple multiplicative form, we can easily find the optimal report(s) from any given transaction sequence or its associated bit string. We use the following two auxiliary functions, which give the values of the losses and revenues that would be reported if the firm were to put its cutoff between the j^{th} and $(j+1)^{\text{st}}$ entries in y (i.e., if $\widehat{\alpha} \in (\alpha_{(j)}, \alpha_{(j+1)}]$):

$$\ell_j(y) = \begin{cases} 0, & \text{if } j < 1 \\ D, & \text{if } j \geq D+U \\ j - \sum_{i=1}^{j} i, & \text{otherwise} \end{cases} \qquad r_j(y) = \begin{cases} U, & \text{if } j \leq 1 \\ 0, & \text{if } j > D+U \\ \sum_{i=j+1}^{U+D} i, & \text{otherwise} \end{cases} \tag{12}$$

In words, $\ell_j(\cdot)$ counts the number of the j least persistent transactions that reduce income, and $r_j(\cdot)$ counts the number of the $n-j$ most persistent transactions that increase income. An increase in j slides the cutoff to the right. Example 1 illustrates:

Example 1. *Let* $y = (0,1,0,1,0,0,1)$*. Then for* $j \in \{0,\ldots,7\}$*, the firm has the following values:*

String (recurring above cutoff '|')	j	D	U	$\ell_j(y)$	$r_j(y)$	$\ell_j(y) \cdot r_j(y)$
$(\mid 0,1,0,1,0,0,1)$	0	4	3	0	3	0
$(0 \mid 1,0,1,0,0,1)$	1	4	3	1	3	3
$(0,1 \mid 0,1,0,0,1)$	2	4	3	1	2	2
$\boxed{(0,1,0 \mid 1,0,0,1)}$	3	**4**	**3**	**2**	**2**	$\boxed{4}$
$(0,1,0,1 \mid 0,0,1)$	4	4	3	2	1	2
$(0,1,0,1,0 \mid 0,1)$	5	4	3	3	1	3
$\boxed{(0,1,0,1,0,0 \mid 1)}$	6	**4**	**3**	**4**	**1**	$\boxed{4}$
$(0,1,0,1,0,0,1 \mid)$	7	4	3	4	0	0

The firm's set of optimal reports is

$$\{(4,3,2,2),(4,3,4,1)\}$$

As the example shows, L and R are those instances of ℓ_j and r_j that maximize $\ell_j \cdot r_j$, as required by Theorem 2. It is clear from Example 1 that the firm's optimal reporting strategy is a correspondence, and from Proposition 2 that different sequences can generate the same optimal report.

3.3. Classification informativeness

Our results to this point address classification from the firm's viewpoint. In this subsection, we turn to the investor's problem, addressing two aspects of the amount of information pooled in a given report. First, we characterize what the investor can learn about the persistence of the firm's line items from the report. This characterization helps clarify why line item persistence varies with income, even in the absence of any misclassification. Second, we discuss the amount of pooling in terms of valuation, rather than in terms of persistence ranking. As Proposition 2 shows, as long as the firm has at least \$2 in positive transactions and at least \$2 in negative ones, the report (D,U,L,R) is informationally insufficient for v^*. Below, we explicitly show how much variation there is in firm value given the report, that is, how opaque a given report is.

To fix terminology, call a bit string y *legal* for report if (D, U, L, R) is an optimal report of y. If (D, U, L, R) is the uniquely optimal report for y, we say that y is *strictly legal* for the report.

Definition 1. *Given nonnegative integers* (D, U, L, R)*, a sequence* $y = (y_1, \ldots, y_{U+D})$ *is* legal *with respect to* (D, U, L, R) *if:*

1. *it contains* U 1*s and* D 0*s;*

2. *for* $i = L + U - R$*, we have* $\ell_i(y) = L$ *and* $r_i(y) = R$*;*

3. *for every* $j \in \{0, \ldots, U + D\}$*, we have* $\ell_j(y) r_j(y) \leq LR$*.*

A legal sequence is strictly legal *if we have*

3' *for every* $j \in \{0, \ldots, U + D\} \backslash \{L + U - R\}$*, we have* $\ell_j(y) r_j(y) < LR$*.*

If $D = 0$ or $U = 0$, then all classification decisions are equivalent. In these cases, all reports satisfying (2) are legal and no report is strictly legal. However, all the legal reports in this case are informationally equivalent. If D and U are both positive, then no report with $L = 0$ or $D = 0$ is legal. On the other hand, for any report (D, U, L, R) satisfying $D > 0, U > 0, L \in \{1, \ldots, D\}$, and $R \in \{1, \ldots, U\}$, the sequence $\underline{y}(D, U, L, R)$ is always strictly legal. This follows immediately from writing out $\ell_j(\underline{y})$ and $r_j(\underline{y})$. In particular, if a report with $\min\{D, U, L, R\} \geq 1$ has a unique legal sequence, it is strictly legal. Example 2 illustrates a case in which $\min\{D, U, L, R\} \geq 1$ and there is a legal sequence that is not strictly legal.

EXAMPLE 2. *Take* $D = U = 3$*,* $L = 1$*,* $R = 2$*. Among the* 20 *possible rearrangements of* 3 1*s and* 3 0*s,* 12 *satisfy* $\ell_2(y) = 1$ *and* $r_2(y) = 2$*:*

$$\begin{array}{cc} (0,1,0,0,1,1) & (1,0,0,0,1,1) \\ (0,1,0,1,0,1) & (1,0,0,1,0,1) \\ (0,1,0,1,1,0) & (1,0,0,1,1,0) \\ (0,1,1,0,0,1) & (1,0,1,0,0,1) \\ (0,1,1,0,1,0) & (1,0,1,0,1,0) \\ (0,1,1,1,0,0) & (1,0,1,1,0,0) \end{array}$$

Of those, only $(1,0,1,1,0,0)$ *and* $(1,0,1,0,1,0)$ *satisfy* $\ell_j(y)r_j(y) \leq 2$ *for* $j = 0,\ldots,6$*, so they are the only legal sequences. Furthermore, only* $(1,0,1,1,0,0)$ *is strictly legal.*

In Example 2, the classification principle (2) reduces the number of possible reports from 20 to 12 because the report indicates that there is $L = 1$ income-reducing transaction and $G = U - R = 1$ income-increasing transaction below the cutoff. This restricts the possible sequences to those that start with $(0,1)$ and those that start with $(1,0)$. Similarly, there are two 0s and two 1s above the cutoff, giving $\binom{4}{2} = 6$ permissible orderings under the principle. As the example shows, the requirement that the report maximizes $L \cdot R$ for a given sequence dramatically reduces the total number of possibilities, from 12 that are permissible to two that are optimal (*legal* in our terminology), with only one having (D,U,L,R) as its uniquely optimal report (*strictly legal* in our terminology).

An important idea in Example 2 is that we can split the task of finding sequences optimally reported as (D,U,L,R) into two subtasks. If we take the amount of revenue R as given, we can concentrate on the subsequence below the cutoff. Similarly, we can take L as given, and then focus on the subsequence above the cutoff. In the example, we see that if

y starts with $(0,1)$, the firm could shift its cutoff one position to the left, thereby increasing the number of positive transactions reported as revenues without decreasing the number of income-reducing transactions reported as losses. Therefore, we know that any sequence y that starts with $(0,1)$ cannot be a legal sequence for $(3,3,1,2)$. The subsequence to the left of the cutoff must be $(1,0)$.

Similarly, there are six possible subsequences above the cutoff:

$$\{(0,0,1,1),(0,1,0,1),(0,1,1,0),(1,0,0,1),(1,0,1,0),(1,1,0,0)\}$$

The first three all start with an income-reducing transaction. These cannot be the continuation of a legal sequence, because the firm could shift its cutoff one position to the right, increasing the number of income-reducing transactions reported as losses without changing its reported revenue.

The fourth also cannot be part of a legal sequence, but the reason is more subtle. With $L = 1$, the subsequence $(1,0,0,1)$ would give $R = 2$ and thus $LR = 2$. Shifting the cutoff three positions rightward would increase L from 1 to 3, and would decrease R from 2 to 1, making $LR = 3$. Ruling out the subsequence $(1,0,0,1)$ requires us to know L; for instance, if we had $L = 3$ and $R = 2$, we would have had $LR = 6$; increasing L by 2 and decreasing R by 1 would not be worthwhile, as it would drop the product LR from 6 to 5.

We make this argument precise in two lemmas. To fix notation, let

$$\begin{aligned}
\mathcal{G}_{L,R}(D,U) &= \{y \in \{0,1\}^n | y \text{ is legal for } (D,U,L,R)\} \\
\mathcal{G}'_{L,R}(D,U) &= \{y \in \{0,1\}^n | y \text{ is strictly legal for } (D,U,L,R)\} \\
g_{L,R}(D,U) &= |\mathcal{G}_{L,R}(D,U)| \\
g'_{L,R}(D,U) &= \left|\mathcal{G}'_{L,R}(D,U)\right| \qquad (13)
\end{aligned}$$

If $L > D$, $L < 0$, $R > U$, or $R < 0$, we write $\mathcal{G}_{L,R}(D,U) = \mathcal{G}'_{L,R}(D,U) = \varnothing$ and $g_{L,R}(D,U) = g'_{L,R}(D,U) = 0$. We state the first lemma as follows:

LEMMA 1. *For given (D,U,L,R), we have $y \in \mathcal{G}_{L,R}(D,U)$ if and only if*

$$(y_1, \ldots, y_{L+U-R}, \underbrace{1, \ldots, 1}_{R}) \in \mathcal{G}_{L,R}(L,U), \text{ and}$$

$$(\underbrace{0, \ldots, 0}_{L}, y_{L+U-R+1}, \ldots, y_{U+D}) \in \mathcal{G}_{L,R}(D,R)$$

Consequently,

$$g_{L,R}(D,U) = g_{L,R}(L,U) \cdot g_{L,R}(D,R).$$

The same statement holds for $\mathcal{G}'$ and g'.

Lemma 1 establishes that we can decompose our search for the sequences under which the report is optimal into two smaller search problems: those of finding subsequences to the left and to the right of the cutoff for which the report is optimal. Although both subsequences depend on L and R, we show that one set of subsequences is independent of U, and the other is independent of D.

Define $\gamma_{L,R}(m) := g_{L,R}(m,R)$ and $\gamma'_{L,R}(m) := g'_{L,R}(m,R)$. Then the following gives the result:

LEMMA 2. *Given (D,U,L,R), a bit string $y \in \mathcal{G}_{L,R}(D,U)$ if and only if $(1-y_{U+D}, \ldots, 1-y_1) \in \mathcal{G}_{R,L}(U,D)$. Consequently, $g_{L,R}(L,U) = g_{R,L}(U,L) = \gamma_{L,R}(U)$, and*

$$g_{L,R}(D,U) = \gamma_{L,R}(D) \cdot \gamma_{R,L}(U).$$

The same statement holds for $\mathcal{G}', g'$, and γ'.

From Lemma 1, Lemma 2, and two technical results in the appendix, we can derive the following explicit characterization of the number of sequences pooled in a given report. Below, we use the following standard notation: for $c \in \mathbb{R}$, we denote the ceiling (least integer weakly above c) and floor (greatest integer weakly below c) as

$$\lceil c \rceil := \min\{k \in \mathbb{Z} | c \leq k\} \text{ (the \textit{ceiling} of } c)$$
$$\lfloor c \rfloor := \max\{k \in \mathbb{Z} | k \leq c\} \text{ (the \textit{floor} of } c)$$

The characterization of the amount of pooling is as follows:

THEOREM 3. *We have*

$$g_{L,R}(D,U) = \gamma_{L,R}(D) \cdot \gamma_{R,L}(U), \tag{14}$$

where function $\gamma_{L,R}$ *satisfies the following recurrence:*

$$\gamma_{L,R}(m) = \begin{cases} 0 & \textit{if } m < L \\ 1 & \textit{if } m = L \\ \sum_{j=1}^{\lceil (LR+1)/m \rceil} (-1)^{j-1} \binom{\lceil (LR+1)/m \rceil}{j} \gamma_{L,R}(m-j) & \textit{otherwise} \end{cases} \tag{15}$$

Similarly,

$$g'_{L,R}(D,U) = \gamma'_{L,R}(D) \cdot \gamma'_{R,L}(U), \tag{16}$$

where function $\gamma'_{L,R}$ *satisfies the following recurrence:*

$$\gamma'_{L,R}(m) = \begin{cases} 0 & \textit{if } m < L \\ 1 & \textit{if } m = L \\ \sum_{j=1}^{\lceil LR/m \rceil} (-1)^{j-1} \binom{\lceil LR/m \rceil}{j} \gamma'_{L,R}(m-j) & \textit{otherwise} \end{cases}. \tag{17}$$

A practical consequence of Theorem 3 is that the persistence of income-reducing line items (L in the model) should not be expected to be independent of reported recurring income. The number of possible sequences below the cutoff is at least $\gamma'_{L,R}(D)$ and at most $\gamma_{L,R}(D)$. Both depend on revenue and on total income reducing items (hence on expenses, and therefore on reported recurring revenue less recurring expenses, which the classification shifting and non-GAAP reporting literature refer to as *core* income). A systematic relationship between core income and the persistence of income-reducing items arises without any violation of the classification principle, a point that to the best of our knowledge is not found in the prior literature.

To illustrate the calculation in Theorem 3, return to Example 2, in which $(D, U, L, R) = (3, 3, 1, 2)$. Using an explicit listing, the example shows that there are two legal sequences, one of which is strictly legal. We now show that Theorem 3 gives us this result without requiring an explicit listing of the possible sequences.

EXAMPLE 3. *Let* $(D, U, L, R) = (3, 3, 1, 2)$*. From* (14)*, we have*

$$g_{1,2}(3, 3) = \gamma_{1,2}(3) \cdot \gamma_{2,1}(3)$$

From (15)*,*

$$\begin{aligned}
\gamma_{1,2}(3) &= \sum_{j=1}^{\lceil 3/3 \rceil} (-1)^{j-1} \binom{\lceil 3/3 \rceil}{j} \gamma_{1,2}(3-j) = (-1)^0 \binom{1}{1} \gamma_{1,2}(2) = \gamma_{1,2}(2) \\
&= \sum_{j=1}^{\lceil 3/2 \rceil} (-1)^{j-1} \binom{\lceil 3/2 \rceil}{j} \gamma_{1,2}(2-j) \\
&= \binom{2}{1} \gamma_{1,2}(1) - \binom{2}{2} \gamma_{1,2}(0) = 2 \cdot 1 + 1 \cdot 0 = 2
\end{aligned}$$

and

$$\gamma_{2,1}(3) = \sum_{j=1}^{\lceil 3/3 \rceil} (-1)^{j-1} \binom{\lceil 3/3 \rceil}{j} \gamma_{2,1}(3-j) = (-1)^0 \binom{1}{1} \gamma_{2,1}(2) = 1$$

giving $g_{1,2}(3,3) = 2 \cdot 1 = 2$ *legal sequences, i.e., two transaction orderings that the firm optimally reports as* (D, U, L, R).

From (16), *we have*

$$g'_{1,2}(3,3) = \gamma'_{1,2}(3) \cdot \gamma'_{2,1}(3)$$

From (17),

$$\begin{aligned} \gamma'_{1,2}(3) &= \sum_{j=1}^{\lceil 2/3 \rceil} (-1)^{j-1} \binom{\lceil 2/3 \rceil}{j} \gamma'_{1,2}(3-j) = (-1)^0 \binom{1}{1} \gamma'_{1,2}(2) = \gamma'_{1,2}(2) \\ &= \sum_{j=1}^{\lceil 2/2 \rceil} (-1)^{j-1} \binom{\lceil 2/2 \rceil}{j} \gamma'_{1,2}(2-j) = \gamma'_{1,2}(1) = 1 \end{aligned}$$

and

$$\gamma'_{2,1}(3) = \sum_{j=1}^{\lceil 2/3 \rceil} (-1)^{j-1} \binom{\lceil 2/3 \rceil}{j} \gamma'_{2,1}(3-j) = (-1)^0 \binom{1}{1} \gamma'_{2,1}(2) = 1$$

giving $g'_{1,2}(3,3) = 1 \cdot 1 = 1$ *strictly legal sequence.*

Theorem 3 allows for fast computation of the amount of pooling with a given report. For example, for the report $(D, U, L, R) = (12, 15, 5, 8)$, we can compute $\gamma_{5,8}(12) = 1,107$ and $\gamma_{8,5}(15) = 248$, giving $g_{5,8}(12, 15) = 274,356$. This provides an upper bound on the number of different sequences that can be reported as $(12, 15, 5, 8)$. A similar calculation shows $\gamma'_{5,8}(12) = 927, \gamma'_{8,5}(15) = 196$, and hence $g'_{5,8}(12, 15) = 181,692$. This tells us that roughly 2/3 of the sequences that the firm could optimally report as $(12, 15, 5, 8)$ have this as their uniquely optimal report.

It is natural to ask whether, as the number of transactions increases, the fraction of legal sequences that are also strictly legal increases or decreases. Equations (15) and (17) show that the behavior of g and g' depends on a recursion in which each step depends on a factorization problem. Solving this directly is unlikely to be attainable. However, using Theorem 3, we can conjecture that the ratio of g'/g approaches 1 as the number of transactions is scaled up.

EXAMPLE 4. *Let* $(D, U, L, R) = (120, 150, 50, 80)$. *Then*

$$g_{50,80}(120, 150) \approx 6.96 \cdot 10^{75}$$

$$g'_{50,80}(120, 150) \approx 6.92 \cdot 10^{75}$$

$$\therefore, \quad \frac{g'_{50,80}(120, 150)}{g_{50,80}(120, 150)} \approx 0.994$$

Thus, we see that for $(D, U, L, R) = (12, 15, 5, 8)$, roughly 2/3 of the legal sequences are strictly legal. Scaling all transactions by a factor of 10 increases this ratio to 99.4%. In other words, for 99.4% of the sequences that the firm optimally reports as $(120, 150, 50, 8)$, the investor knows that this was the uniquely optimal report. Figure 3 illustrates.

Figure 3 about here.

Although the report is asymptotically essentially unique, it remains for us to address how transparent the report is to the investor. In other words, we are interested in how the investor values the firm given the report.

Recall from (9) that the minimum value associated with report (D, U, L, R) is

$$LR + \binom{U+1}{2}$$

Define the *opacity* of the report as

$$o(D,U,L,R) = \max_{y \in \mathcal{G}_{L,R}(D,U)} w(y) - \min_{y \in \mathcal{G}_{L,R}(D,U)} w(y) \tag{18}$$

That is, the opacity is the range of possible valuations consistent with the firm optimally choosing a given report. Our last main result, Theorem 4, characterizes the opacity $o(\cdot)$ and the corresponding range of possible values given the report.

Theorem 4. *We have*

$$o(D,U,L,R) = \sum_{i=1}^{R} \min\left(D - L, \left\lfloor \frac{(i-1)L}{R-i+1} \right\rfloor\right) + \sum_{i=1}^{L} \min\left(U - R, \left\lfloor \frac{(i-1)R}{L-i+1} \right\rfloor\right).$$

Furthermore, the image of w on $\mathcal{G}_{L,R}(D,U)$ is all of

$$[LR + \tbinom{U+1}{2}, LR + \tbinom{U+1}{2} + o(D,U,L,R)] \cap \mathbb{Z} \tag{19}$$

Returning to Example 4, if the firm reports $(D,U,L,R) = (120,150,50,80)$, the opacity is 5,914. Theorem 4 also tells us that all values in the range (19), if possible with n bits, can occur as an optimal report. If these are distributed symmetrically, then the average value associated with the report $(120,150,50,80)$ is approximately $30 - 0.0775/\rho$.

In this case, the firm value increases in the discount rate. As ρ becomes large, the firm value asymptotically approaches its current net income π, and the distinction between recurring and non-recurring income vanishes. Even though the firm has positive net income and positive core income, the investor in this example typically interprets the negative transactions as more persistent than its positive ones. At a low enough discount rate (below 0.0026), the firm value becomes negative.

4. Discussion and conclusion

Classification standards are typically in the form of a general principle, rather than a specific rule. To some degree, this is of necessity. For example, Dye [2002] shows that a hypothetical rules-based classification standard would lead to standards creep, i.e., would inherently lack stability. Relatedly, Dechow and Schrand [2004] and McVay [2006] argue that a rule would be difficult to enforce, and that auditors cannot or will not go so far as to impose an exact classification or specific threshold on any transaction. To the best of our knowledge, this paper is the first to specify and study the informativeness of reports under a classification principle.

Our findings provide a simple, multiplicative characterization of how firms optimally choose their persistence threshold for what counts as recurring. As is widely appreciated, firms face a trade-off between setting a higher persistence threshold, in order to report more income-reducing transactions as nonrecurring, and setting a lower one, in order to report more income-increasing transactions as recurring. The multiplicative characterization of the firm's objective arises as a result of the firm's response to investor skepticism.

A consequence for empirical studies on classification is that a firm's reported income-reducing items is not independent of the firm's recurring revenues and other recurring income-increasing items. The two are selected simultaneously, and empirical work may need to adjust for this source of endogeneity.

In a more positive direction, we provide explicit formulas for characterizing the amount that an optimally classified report pools firms of different types. Our first formula describes the degree that firms report the same way despite having different persistence of their income-statement line items. Our second formula describes the amount that a given report pools firms with distinct fundamental values.

A. Proofs

Proof of Proposition 1. For each $i \in \{1, \ldots, n\}$, $\widetilde{\alpha}_i \sim U[1, (1+\rho)/\rho]$, so

$$\rho\left(\widetilde{\alpha}_i - 1\right) \sim U[0,1]$$

and $\rho(\widetilde{\alpha} - 1)_{(i)} = \rho\left(\widetilde{\alpha}_{(i)} - 1\right)$, i.e., the order is the same as the ordering of the $\widetilde{\alpha}_{(i)}$. From the fact that the i^{th} sample order statistic of n independent $U[0,1]$ draws is distributed $\text{Beta}(i, n-i+1)$ (see Arnold et al. [2008]), we have

$$\rho\left(\widetilde{\alpha}_{(i)} - 1\right) \sim \text{Beta}(i, n-i+1)$$

Using the fact that the a $\text{Beta}(a,b)$-distributed random variable has mean $a/(a+b)$ and variance $ab/[(a+b)^2(a+b+1)]$, it follows that

$$\begin{aligned} E[\widetilde{\alpha}_{(i)}] &= \frac{i}{\rho(n+1)} + 1 \\ \text{Var}[\widetilde{\alpha}_{(i)}] &= \frac{i(n-i+1)}{\rho^2(n+1)^2(n+2)} \end{aligned}$$

The investor's mean-squared error, given that the report reveals the persistence ordering, is therefore

$$\begin{aligned} \sum_{i=1}^{n} \frac{i(n-i+1)}{\rho^2(n+1)^2(n+2)} &= \frac{1}{\rho^2(n+1)^2(n+2)}\left[(n+1)\sum_{i=1}^{n} i - \sum_{i=1}^{n} i^2\right] \\ &= \frac{n(n+1)^2/2 - n(n+1)(2n+1)/6}{\rho^2(n+1)^2(n+2)} \\ &= \frac{n}{6\rho^2(n+1)} \end{aligned} \tag{20}$$

As $n \to \infty$, expression (20) is easily seen to converge to $1/(6\rho^2)$ from below. $\square$

Proof of Theorem 1. Fix U and D. From (8), $w(y)$ is a sufficient statistic for v^*. The smallest value of $w(y)$ consistent with (D, U) is obtained if the 1s in y occur in positions $\{1, \ldots, U\}$. The largest value of $w(y)$ consistent with (D, U) is obtained if the 1s in y occur in positions $\{D+1, \ldots, D+U\}$. Call these values $\underline{w}$ and $\overline{w}$, respectively. Then

$$\underline{w} = \sum_{i=1}^{U} i = \frac{U(U+1)}{2}$$
$$\overline{w} = \sum_{i=D+1}^{D+U} i = \frac{(D+U)(D+U+1)}{2} - \frac{D(D+1)}{2} = D \cdot U + \frac{U(U+1)}{2}$$

For any y, interchanging a 0 with a 1 that is j positions ahead of (respectively behind) the 0 decreases (respectively increases) the sum by j. Thus, every integer sum between $\underline{w}$ and $\overline{w}$ is attainable, so there are

$$\overline{w} - \underline{w} + 1 = D \cdot U + 1$$

possible distinct values of $w(y)$, and hence $D \cdot U + 1$ possible values of v^*.

For any report that satisfies (2),

$$L \in \{0, \ldots, D\} \text{ and } R \in \{0, \ldots, U\},$$

giving $D+1$ possible values for L and $U+1$ possible values for U. The number of possible reports is therefore

$$(D+1)(U+1) = D \cdot U + \overbrace{D+U}^{n} + 1 = D \cdot U + n + 1$$

$\square$

Proof of Proposition 2. We proceed by cases. First, suppose $\min\{D, U\} \leq 1$. If

$$\max\{i \in \mathbb{Z}_{++} | y_i = 1\} < \min\{i \in \mathbb{Z}_{++} | y_i = 0\}$$

set $L = 0$ and $R = U$. Otherwise, let i be the position in y of the 0 before a 1. By hypothesis, there is exactly one $\{0, 1\}$ subsequence in y, so i is uniquely determined. Set

$$L = i - \sum_{j=1}^{i} y_i \text{ and } R = \sum_{j=i+1}^{U+D} y_j$$

A direct examination of subcases shows no firm can be mimicked by a type with lower v^*.

For the second case, assume that $\min\{D, U\} > 1$. Fix $i \in \{1, \ldots, U + D - 3\}$, and let

$$\widehat{y} = (y_1, \ldots, y_{i-1}, 0, 1, 0, 1, y_{i+4}, \ldots, y_{U+D})$$
$$y' = (y_1, \ldots, y_{i-1}, 0, 1, 1, 0, y_{i+4}, \ldots, y_{U+D})$$
$$y'' = (y_1, \ldots, y_{i-1}, 1, 0, 0, 1, y_{i+4}, \ldots, y_{U+D})$$

Then $w(\widehat{y}) = w(y') + 1 = w(y'') + 1$, so both y' and y'' are firm types with strictly lower values of v^* than that of $\widehat{y}$. However,

$$(\widehat{y}_1, \ldots, \widehat{y}_{i+1}) = \left(y'_1, \ldots, y'_{i+1}\right), \text{ and}$$
$$(\widehat{y}_{i+2}, \ldots, \widehat{y}_{U+D}) = \left(y''_{i+2}, \ldots, y''_{U+D}\right)$$

If a firm of type $\widehat{y}$ sets its cutoff at or below the $(i + 1)^{\text{st}}$ position, it pools with a firm of type y'. If it sets its cutoff higher, it pools with a firm of y''. Either way, the firm pools with a type with a lower value of v^*. $\square$

Proof of Theorem 2. We have

$$
\begin{aligned}
w(\underline{y}(D,U,L,R)) &= \sum_{i=1}^{U-R} i + \sum_{\substack{j=U-R\\+L+1}}^{U+L} i \\
&= \frac{(U-R)(U-R+1)}{2} + \frac{(U+L)(U+L+1)}{2} \\
&- \frac{(U-R+L)(U-R+L+1)}{2} \\
&= \frac{1}{2}\left(U^2+U+2LR\right) = LR + \binom{U+1}{2}
\end{aligned}
$$

Substituting into (8),

$$
\begin{aligned}
v^*(\underline{y}(D,U,L,R)) &= \pi - \frac{n}{2\rho} + \frac{2w(y(D,U,L,R))}{\rho(n+1)} \\
&= \pi - \frac{n}{2\rho} + \left[\frac{2}{\rho(n+1)}\right]\left[\frac{U(U+1)}{2} + LR\right] \\
&= \pi - \frac{n}{2\rho} + \frac{U(U+1)}{\rho(n+1)} + \left[\frac{2}{\rho(n+1)}\right] LR
\end{aligned}
$$

As $\rho > 0$ and $n > 0$, the most skeptical interpretation of the report is linear and increasing in the product $L \cdot R$. $\square$

Proof of Lemma 1. Write

$$
\begin{aligned}
y &= (y_1, \ldots, y_{U+D}) \\
y' &= (y_1, \ldots, y_{L+U-R}, \underbrace{1, \ldots, 1}_{R}) \\
y'' &= (\underbrace{0, \ldots, 0}_{L}, y_{L+U-R+1}, \ldots, y_{U+D})
\end{aligned}
$$

Then $\ell_j(y) r_j(y) > LR$ for $j < L+U-R$ if and only if $\ell_j(y') r_j(y') > LR$, and $\ell_j(y) r_j(y) >$

LR for $j > L+U-R$ if and only if $\ell_{j-u+r}(y'')r_{j-U+R}(y'') > LR$. The equivalence follows. The proof for $\mathcal{G}'$ and g' is the same. □

Proof of Lemma 2. Immediate from interchanging the roles of D and U and of L and R. □

Proof of Theorem 3. Lemmas 1 and 2 establish that $g_{L,R}(D,U) = \gamma_{L,R}(D) \cdot \gamma_{R,L}(U)$. It remains to prove the recursive formula for $\gamma_{L,R}(m)$ and $\gamma'_{L,R}(m)$. It is clear that $\gamma_{L,R}(m) = \gamma'_{L,R}(m) = 0$ if $m < L$, and that $\gamma_{L,R}(L) = \gamma'_{L,R}(L) = 1$. The following characterizes $\mathcal{G}_{L,R}(R,m)$ and $\mathcal{G}'_{L,R}(R,m)$.

LEMMA 3. *For given m, L, R, we have $y = (y_1, \ldots, y_{R+m}) \in \mathcal{G}_{L,R}(m,R)$ if and only if:*

1. *$y_1 = \ldots = y_L = 0$;*
2. *there are exactly R 1s in $(y_{L+1}, \ldots, a_{R+m})$;*
3. *for $i = 1, \ldots, R$, the number of 0s in $(y_{l+1}, \ldots, y_{r+m})$ before the i^{th} 1 is at most $(i-1)L/(R-i+1)$.*

A similar statement holds for $\mathcal{G}'_{L,R}(m,R)$, with (3) replaced by

(3') *for $i = 1, \ldots, R$, the number of 0s in $(y_{L+1}, \ldots, y_{R+m})$ before the i^{th} 1 is less than $(i-1)L/(R-i+1)$.*

Proof. We use the following standard notation: for $k \in \mathbb{Z}_{++}$, let $[k] := \{1, \ldots, k\}$. Assume first that $y \in \mathcal{G}_{L,R}(m,R)$. Then all 1s have to be to the right of position L, so the first two conditions are obvious. Furthermore, if $i \in [R]$ and j_i is the position of the i^{th} 1, then $r_{j_i-1}(y) = R-i+1$ and $\ell j_i - 1(y) r_{j_i-1}(y) \leq LR = L(R-i+1) + (i-1)L$, so $\ell_{j_i-1}(y) \leq L + (i-1)L/(R-i+1)$.

Conversely, suppose that the three conditions are satisfied. Clearly, there are m 0s and R 1s in y, and $\ell_L(y)r_L(y) = LR$. Again, for $i \in [R]$, denote by j_i the position of the i^{th}

1 in y, and also write $j_0 = 0$, $j_{R+1} = R + m + 1$. Now take $j \in [0, R + m]$, and pick (the unique) $i \in [R + 1]$ so that $j_{i-1} \leq j < j_i$. Then $\ell_j(y) \leq \ell_{j_i}(y) \leq L + (i - 1)L/(R - i + 1)$ and $r_j(y) = R - i + 1$, so $\ell_j(y)r_j(y) \leq LR$.

The proof for $\mathcal{G}'_{L,R}(R, m)$ is analogous. ■

For the last step in establishing the recursions (15) and (17), we define the following auxiliary functions. Given $m, L, R \in \mathbb{Z}_+$, write $k = \lceil (LR + 1)/m \rceil$ and $k' = \lceil LR/m \rceil$. Define the functions φ and φ', where

$$\bigcup_{j=0}^{k} \binom{[k]}{j} \times \mathcal{G}_{L,R}(m - j, R) \xrightarrow{\varphi_{m,L,R}} \bigcup_{j=0}^{k} \binom{[k]}{j} \times \mathcal{G}_{L,R}(m - j, R)$$
$$\bigcup_{j=0}^{k'} \binom{[k']}{j} \times \mathcal{G}_{L,R}(m - j, R) \xrightarrow{\varphi'_{m,L,R}} \bigcup_{j=0}^{k'} \binom{[k']}{j} \times \mathcal{G}_{L,R}(m - j, R),$$

as follows. Choose $S \in [k]$ and a sequence $(y_1, \ldots, y_{R+m-|S|}) \in \mathcal{G}_{L,R}(m - |S|, R)$. Let

$$\begin{aligned} i = \max\{\iota \in [k] | \iota \in S, \text{ or there is at least one 0 between the} \\ (R - k + \iota)^{\text{th}} \text{ 1 and the } (R - k + \iota + 1)^{\text{st}} \text{ 1}\} \\ = \max\{\iota \in [k] | \iota \in S \text{ or } j_{R-k+\iota} \leq j_{R-k+\iota+1} - 2\}, \end{aligned}$$

where for $h \in [R]$, j_h is as defined in the proof of Lemma 3.

For $\iota = k$, we interpret this as meaning there is at least one 0 after the last 1. If $i \in S$, define $S' = S \backslash \{i\}$, and let y' be the sequence obtained from y by inserting a 0 between the $(R - k + i)^{\text{th}}$ and the $(R - k + i + 1)^{\text{st}}$ 1. If $i \notin S$, define $S' = S \cup \{i\}$, and let y' be the sequence obtained from y by deleting one of the zeroes between the $(R - k + i)^{\text{th}}$ and the $(R - k + i + 1)^{\text{st}}$ 1. Then let $\varphi_{m,L,R}(S, y) = (S', y')$. Define $\varphi'_{m,L,R}$ analogously, with k' in place of k throughout.

Lemma 4. *For all $m, L, R \in \mathbb{Z}_+$, the maps $\varphi_{m,L,R}$ and $\varphi'_{m,L,R}$ are well-defined involutions which change the parity of the size of the first argument, and therefore prove*

$$\sum_{j=0}^{k} \binom{k}{j} \gamma_{L,R}(m-j) = 0$$

and

$$\sum_{j=0}^{k'} \binom{k'}{j} \gamma'_{L,R}(m-j) = 0.$$

Proof. Take $k = \lceil (LR+1)/m \rceil$, $S \subseteq [k]$, $y \in \mathcal{G}_{L,R}(m-|S|, R)$. First, let us prove that the set $\{\iota \in [k] | \iota \in S \text{ or } j_{R-k+\iota} \leq j_{R-k+\iota+1} - 2\}$ is non-empty. If $S \neq \varnothing$, this is obvious. Assume that $S = \varnothing$ (which implies $y \in \mathcal{G}_{L,R}(m, R)$) and that $j_{R-k+\iota} = j_{R-k+\iota+1} - 1$ for $\iota = 1, \ldots, k$. Note that this means that the last k terms in y are 1. The number of 0s in $(y_{L+1}, \ldots, y_{R+m})$ before the $(R-k+1)^{\text{st}}$ 1 is therefore $m - L$. On the other hand, by Lemma 3, this must be at most $(R-k)L/k$. We get $k \leq LR/m$, which contradicts $k = \lceil (LR+1)/m \rceil$.

It is clear that S' defined in the procedure lies in $[k]$ and differs by S by exactly one element, and also that the sequence y' is well-defined (we erase a 0 only when there is at least one 0). Let us prove that $y' \in \mathcal{G}_{L,R}(m - |S'|, R)$. The first two conditions of Lemma 3 are obviously satisfied; we have to prove the same for the third condition. When $i \notin S$, this is obvious (as there are fewer 0s then before). When $i \in S$, we have to see that $m - |S| - L + 1$, the number of 0s before the $(R-k+i+1)^{\text{st}}$ 1 in y', is at most $(R-k+i)L/(k-i)$ (if $k = i$, there is nothing to prove). But $|S| \geq 1$, so it is enough to prove that $(m-L)(k-i) \leq (R-k+i)L$. But $(m-L)(k-i) \geq (R-k+i)L + 1$ would imply $mk \geq LR + 1 + mi$ and $k \geq \frac{LR+1}{m} + i$, which contradicts $k = \lceil (LR+1)/m \rceil$.

It is clear that the map is an involution.

The proof for $\varphi'_{m,L,R}$ is almost exactly the same. ■

The theorem then follows from Lemmas 1–4. □

Proof of Theorem 4. We note first that $\underline{y}(D,U,L,R) \in \mathcal{G}_{L,R}(D,U)$. Begin by considering the $D-L+R$ positions to the right of the cutoff. If the subsequence $\{1,0\}$ occurs, we can replace it with $\{0,1\}$, and thereby increase increase the opacity by 1. We can proceed sequentially, first pushing the leftmost 0 as far to the left as possible through repeatedly replacing a $\{1,0\}$ subsequence with a $\{0,1\}$ subsequence, then pushing the next leftmost 0 as far to the left as possible, and so forth. The largest number of 0s before the i^{th} 1 cannot exceed $D-L$, as this is the total number of available 0s. By Lemma 3, the largest number of 0s before the i^{th} 1 also cannot exceed $\lceil (i-1)L/(R-i+1)$. With this procedure, we have increased the opacity by 1 in each step, and by

$$\sum_{i=1}^{R} \min\left(D-L, \left\lceil \frac{(i-1)L}{R-i+1} \right\rceil\right)$$

in total.

Next, consider the $U-R+L$ positions to the left of the cutoff. We can keep pushing the 1s to the right by an analogous procedure, while still keeping the sequence legal. In this way, we increase the opacity in each step, and by

$$\sum_{i=1}^{R} \min\left(U-R, \left\lceil \frac{(i-1)R}{L-i+1} \right\rceil\right)$$

in total. This finishes the proof. □

References

ABARBANELL, J. S., AND R. LEHAVY. "Letting the 'Tail Wag the Dog': The Debate over GAAP versus Street Earnings Revisited." *Contemporary Accounting Research* 24 (2007): 675–723.

AMIR, E.; E. EINHORN; AND I. KAMA. "Extracting Sustainable Earnings from Profit Margins." *European Accounting Review* 22 (2013): 685–718.

ANCTIL, R. M.; J. S. JORDAN; AND A. MUKHERJI. "Activity-Based Costing for Economic Value Added.®" *Review of Accounting Studies* 2 (1998): 231–64.

ARNOLD, B. C.; N. BALAKRISHNAN; AND H. N. NAGARAJA. *A First Course in Order Statistics.* Vol. 54 of *Classics In Applied Mathematics.* Philadelphia: Society for Industrial and Applied Mathematics, 2008.

ARYA, A.; J. C. FELLINGHAM; J. C. GLOVER; D. A. SCHROEDER; AND G. STRANG. "Inferring Transactions from Financial Statements." *Contemporary Accounting Research* 17 (2000): 365–85.

BAGNOLI, M.; M. PENNO; AND S. G. WATTS. "Auditing in the Presence of Outside Sources of Information." *Journal of Accounting Research* 39 (2001): 435–47.

BARNEA, A.; J. RONEN; AND J. SADAN. "Classificatory Smoothing of Income with Extraordinary Items." *The Accounting Review* 51 (1976): 110–22.

BERTOMEU, J.; I. VAYSMAN; AND W. XUE. "Voluntary versus Mandatory Disclosure." *Review of Accounting Studies* (2021): Forthcoming.

Bradshaw, M. T., and R. G. Sloan. "GAAP versus The Street: An Empirical Assessment of Two Alternative Definitions of Earnings." *Journal of Accounting Research* 40 (2002): 41–66.

Breuer, M., and D. Windisch. "Investment Dynamics and Earnings-Return Properties: A Structural Approach." *Journal of Accounting Research* 57 (2019): 639–74.

Brown, L. D., and K. Sivakumar. "Comparing the Value Relevance of Two Operating Income Measures." *Review of Accounting Studies* 8 (2003): 561–72.

Calsamiglia, X. "Decentralized Resource Allocation and Increasing Returns." *Journal of Economic Theory* 14 (1977): 263–83.

Dechow, P. M., and C. M. Schrand. *Earnings Quality.* Charlottesville, VA: The Research Foundation of CFA Institute, 2004.

Donelson, D. C.; A. Kartapanis; and C. Koutney. "SEC non-GAAP Comment Letters and Firm Disclosures." Working paper, University of Texas at Austin, Texas A&M University, and George Mason University, 2020.

Doyle, J. T.; R. J. Lundholm; and M. T. Soliman. "The Predictive Value of Expenses Excluded from Pro Forma Earnings." *Review of Accounting Studies* 8 (2003): 145–74.

Dye, R. A. "Strategic Accounting Choice and the Effects of Alternative Financial Reporting Requirements." *Journal of Accounting Research* 23 (1985): 544–74.

——— "Classifications Manipulation and Nash Accounting Standards." *Journal of Accounting Research* 40 (2002): 1125–62.

DYE, R. A., AND S. S. SRIDHAR. "Reliability-relevance Tradeoffs and the Efficiency of Aggregation." *Journal of Accounting Research* 42 (2004): 51–88.

——— "A Positive Theory of Flexibility in Accounting Standards." *Journal of Accounting and Economics* 46 (2008): 312–33.

EBERT, M.; D. SIMONS; AND J. D. STECHER. "Discretionary Aggregation." *The Accounting Review* 92 (2017): 73–91.

EINHORN, E. "The Nature of the Interaction between Mandatory and Voluntary Disclosures." *Journal of Accounting Research* 43 (2005): 593–621.

FASB. *Elements of Financial Statements.* Statement of Financial Accounting Concepts No. 6, Financial Accounting Standards Board, 1985.

GIGLER, F.; C. KANODIA; H. SAPRA; AND R. VENUGOPALAN. "Accounting Conservatism and the Efficiency of Debt Contracts." *Journal of Accounting Research* 47 (2009): 767–97.

GONEDES, N. J. "Risk, Information, and the Effects of Special Accounting Items on Capital Market Equilibrium." *Journal of Accounting Research* 13 (1975): 220–56.

GROSSMAN, S. J. "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics* 24 (1981): 461–83.

GU, Z., AND T. CHEN. "Analysts' Treatment of Nonrecurring Items in Street Earnings." *Journal of Accounting and Economics* 38 (2004): 129–70.

HEMMER, T., AND E. LABRO. "Management by the Numbers: A Formal Approach to Deriving Informational and Distributional Properties of 'Unmanaged' Earnings." *Journal of Accounting Research* 57 (2019): 5–51.

Hiemann, M. “Earnings and Firm Value in the Presence of Real Options.” *The Accounting Review* 95 (2020): 263–89.

Ijiri, Y. *The Foundations of Accounting Measurement: a Mathematical, Economic, and Behavioral Inquiry.* Prentice Hall, Inc., 1967.

——— “The Linear Aggregation Coefficient as the Dual of the Linear Correlation Coefficient.” *Econometrica* 36 (1968): 252–59.

Jordan, J. S., and D. Xu. “On the Communication Complexity of Expected-Profit Maximization.” *Journal of Economic Theory* 86 (1999): 185–202.

Kanodia, C., and A. Mukherji. “Real Effects of Separating Investment and Operating Cash Flows.” *Review of Accounting Studies* 1 (1996): 51–71.

Lev, B. “The Aggregation Problem in Financial Statements: An Informational Approach.” *Journal of Accounting Research* 6 (1968): 247–61.

McVay, S. E. “Earnings Management Using Classification Shifting: An Examination of Core Earnings and Special Items.” *The Accounting Review* 81 (2006): 501–31.

Milgrom, P., and J. Roberts. “Relying on the Information of Interested Parties.” *RAND Journal of Economics* 17 (1986): 18–32.

Milgrom, P. R. “Good News and Bad News: Representation Theorems and Applications.” *Bell Journal of Economics* 12 (1981): 380–91.

Mount, K. R., and S. Reiter. “The Informational Size of Message Spaces.” *Journal of Economic Theory* 8 (1974): 161–92.

NEZLOBIN, A. "Accrual Accounting, Informational Sufficiency, and Equity Valuation." *Journal of Accounting Research* 50 (2012): 233–73.

OKUNO-FUJIWARA, M.; A. POSTLEWAITE; AND K. SUZUMURA. "Strategic Information Revelation." *Review of Economic Studies* 57 (1990): 25–47.

PRICEWATERHOUSECOOPERS. "Financial Statement Presentation," 2019. Guide available at `https://www.pwc.com/us/en/`.

REICHELSTEIN, S. J., AND S. REITER. "Game Forms with Minimal Message Spaces." *Econometrica* 56 (1988): 661–92.

RIBEIRO, A.; Y. SHAN; AND S. TAYLOR. "Non-GAAP Earnings and the Earnings Quality Trade-off." *Abacus* 55 (2019): 6–41.

RONEN, J., AND S. SADAN. "Classificatory Smoothing: Alternative Income Models." *Journal of Accounting Research* 13 (1975): 133–49.

SHIN, H. S. "News Management and the Value of Firms." *RAND Journal of Economics* 25 (1994): 58–71.

——— "Disclosure and Asset Returns." *Econometrica* 71 (2003): 105–33.

SONNENSCHEIN, H. "An Axiomatic Characterization of the Price Mechanism." *Econometrica* 42 (1974): 425–34.