A Theory of Principles-Based Classification

By: Matjaž Konvalinka, Mark Penno and Jack Stecher

August 2020

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: recurring transactions must be more persistent than nonrecurring ones. Under this principle, equilibrium firm behavior provides a new explanation for the observed relationship between income and classifications. Moreover, we find that market prices are more informative under principles-based classifications than they would be under a hypothetical, optimally chosen specific classification rule.

Keywords—classification shifting; earnings management; principles; rules; signaling

1 Introduction

Recurring revenues and expenses affect firm value more than nonrecurring gains and losses do. Yet neither US Generally Accepted Accounting Principles (GAAP) nor International Financial Reporting Standards (IFRS) specify a fixed rule for classifying transactions as recurring or nonrecurring. Our purpose is to demonstrate that a principles-based standard makes income statement classifications useful to investors—in fact, more useful than an optimal rules-based classification standard could be.

To elaborate, the principle we consider is that any transaction a firm classifies as nonrecurring must be less persistent than any it classifies as recurring. Without definitive guidance, each firm can satisfy this principle while picking its own threshold, and has incentive to do so in a way that enables it to report positive items as recurring revenues and negative ones as nonrecurring losses. We model investors as anticipating the firm's incentives and reading reports skeptically. In equilibrium, firms therefore choose their thresholds to maximize the worst-case interpretation of their reports (for detailed exposition, see Milgrom and Roberts, 1986, Okuno-Fujiwara et al., 1990, Shin, 1994, 2003).

For firms with any reasonable number of transactions, equilibrium prices are more informative under a principles-based standard than they would be under an optimal rule. We can say considerably more about the firm's reports under a principles-based standard. We provide a recursive equation that characterizes the amount of pooling that can occur under a given report. This recursion is computationally feasible, with quadratic run-time complexity. A similar recursion characterizes the number of collections of transactions that are uniquely optimal under a given report. Together, these two characterizations provide upper and lower bounds on a report's informativeness. Related expressions enable us to measure the opacity of a report, that is, the range of equilibrium firm values associated with the report.

The principle we consider, and the lack of definitive guidance, is consistent with prior theoretical models of rules versus principles (e.g., Caplan and Kirschenheiter, 2004, Gao et al., 2020, who view a rule as a bright-line) and with long-established standard-setting practice. An example of the latter is FASB (1985, CON6–24), which distinguishes between gains and revenues as follows: "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, p. 3-10 and the related discussion on ASC 605 about gains versus revenues).

Asking auditors to impose more than this principles-based standard is likely infeasible, though we discuss a hypothetical rules-based standard below. Senior audit partners in Nelson et al.'s (2002) survey indicate that determining specific accounting treatments of transactions is highly subjective, making auditors reluctant to overturn firm classification decisions. McVay (2006) and Barua et al. (2010) provide similar arguments, stating that auditors will not or cannot go as far as to impose an exact classification or specific threshold on any possible transaction.

Our results shed light on the empirical relationship between classifications and income, which has drawn considerable attention since being highlighted in McVay (2006) and in follow-up literature (Fan et al., 2010, Fan and Liu, 2017, Fan et al., 2019, Cain et al., 2020), picking up on earlier literature on negative special items and income (Gonedes, 1975, Ronen and Sadan, 1975, Barnea et al., 1976). Because firms use a principles-based standard to signal, we would not expect classification decisions to be independent of income. Instead, in all our examples and simulations, we find that firms classify more income-reducing items as losses as their net income falls. Without considering the implications of a principles-based standard, this pattern would be difficult to explain.

In addition to shedding light on our understanding of principles-based versus rules-based standards, our results contribute to the literature on classification shifting, and more generally to the literature on non-GAAP reporting and on earnings management. Studies of non-GAAP reporting include similar discussions. 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.[1]

In the earnings management literature, Breuer and Windisch (2019), Hemmer and Labro (2019), and Hiemann (2020) are similar in spirit to our paper. They focus on real actions that affect the firm's bottom line. In contrast, our setting focuses entirely on classification decisions that do not affect net income.

The structure of the rest of this paper is as follows: Section 2 introduces the model. Section 3 shows preliminary results. Section 4 then provides our main results. We interpret these and conclude in Section 5. All proofs are in an appendix.

[1]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).

2 The Model

There are two players, a representative investor and a firm. The investor's goal is to price the firm at its expected net present value, given the firm's financial report and a market discount rate $\rho \in (0, 1)$. The firm wants to maximize its share price. Both players are risk neutral and have rational expectations.

The firm's information consists of n ordered pairs, $\{(x_1, \alpha_1), \ldots, (x_n, \alpha_n)\}$. 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.

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

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

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 $\{\widetilde{x}_i\}_{i=1}^n$. We therefore can allow any value for the probability that $\widetilde{x}_i = 1$.

The firm faces a classification problem. Specifically, it publicly releases an income statement consisting of four line items: (recurring) revenues (R) and expenses (E), and (nonrecurring) gains (G) and losses (L). It classifies each x_i by aggregating it into exactly one of these four line items. In this way, the firm partitions its transactions into four general ledger accounts, sums the transactions in each ledger account, and reports each of the four account totals as the line items on its income statement.

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 $\widehat{\alpha} \in [1, (1+\rho)/\rho]$ and report its line items as follows:

$$
\begin{aligned}
R = R(\widehat{\alpha}) &= \sum_{\{i|\alpha_i \geq \widehat{\alpha}\}} \{x_i | x_i > 0\} \qquad & E = E(\widehat{\alpha}) &= \sum_{\{i|\alpha_i \geq \widehat{\alpha}\}} \{-x_i | x_i < 0\} \\
G = G(\widehat{\alpha}) &= \sum_{\{i|\alpha_i < \widehat{\alpha}\}} \{x_i | x_i > 0\} \qquad & L = L(\widehat{\alpha}) &= \sum_{\{i|\alpha_i < \widehat{\alpha}\}} \{-x_i | x_i < 0\} \qquad (2)
\end{aligned}
$$

The positive transactions are therefore recorded either as revenues or gains, the negative ones as expenses or losses. We adopt the convention of reporting E and L as positive totals, so that income is the difference between positive and negative transactions.

For convenience, we define the sum U of the firm's income-increasing transactions and the sum D of its income-reducing transactions as

$$
\begin{aligned}
U &= R + G \\
D &= E + L
\end{aligned}
$$

It is easy to see that U and D do not depend on the firm's chosen cutoff. We can write the firm's income as

$$\pi = U - D$$

and note that the investor can infer $n = U + D$. We denote the firm's report by the quadruple (D, U, L, R).

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 an item increases or reduces income.

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

$$v(D, U, L, R; \alpha_M) = E\left[\sum \tilde{\alpha}_i x_i | \alpha_M; D, U, L, R\right] \quad (3)$$

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

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

Figure 1 summarizes the timeline.

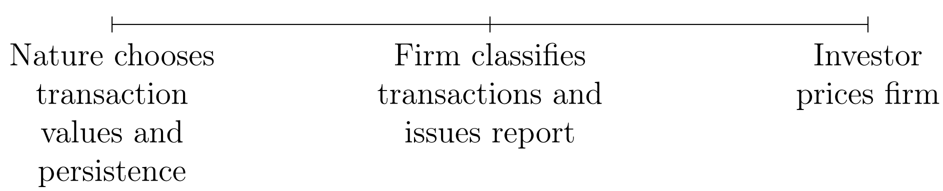

Figure 1 – Timeline.

3 Preliminary Results

3.1 Benchmark: optimal rules-based classification

We begin by considering an optimal rules-based standard. In the context of classification, a rules-based standard is a bright-line cutoff $\alpha_{bl} \in [1, (1+\rho)/\rho)]$, which a standard setter or other authority imposes. Our focus is on an idealized rules-based standard, which is costlessly and perfectly enforced. This is in contrast to prior work on enforcement issues, which emphasizes the degree to which a firm might game a rule.[2] We abstract from enforcement issues to obtain the upper bound on the informativeness of a rules-based standard.

Accordingly, let $\alpha_{bl} \in [1, (1+\rho)/\rho]$ be a given bright-line cutoff. If this cutoff is perfectly enforced, the firm's problem (4) becomes degenerate. The mean squared pricing error is

$$\eta(\alpha_{bl}) := E\left\{\left[\sum_{i=1}^{n} \widetilde{\alpha}_i \widetilde{x}_i - v(D, U, L(\alpha_{bl}), R(\alpha_{bl}); \alpha_{bl})\right]^2\right\} \tag{5}$$

Proposition 1 shows that $\eta(\cdot)$ is minimized when the cutoff is set to the ex ante mean of the $\widetilde{\alpha}_i$.[3]

Proposition 1. *The investor's mean squared error $\eta(\cdot)$ is minimized at the prior mean of the $\widetilde{\alpha}_i$, i.e., at*

$$\alpha_{bl}^* = E[\widetilde{\alpha}_i] = 1 + \frac{1}{2\rho}$$

[2]Dye (2002) discusses the shadow standard that a rule implements, adjusting for a firm's response to enforcement issues in the standard as written. Gao (2017) investigates the optimal choice of standards, knowing that the firm can structure transactions or otherwise alter evidence. From an empirical standpoint, see Dechow and Schrand (2004, p. 113) on concerns about transaction structuring as a way around a rules-based standard.

[3]We follow Shin (2003) in having the investor minimize mean squared pricing error.

At this cutoff, the mean squared error given the report is

$$\eta(\alpha^*_{bl}) = \frac{n}{48\rho^2}$$

It is clear that $\eta(\alpha^*_{bl})$ grows linearly in the number of transactions n. The reason is that, for $i \neq j$, the most informative bright-line classification of $\tilde{\alpha}_i$ is uninformative about $\tilde{\alpha}_j$. The rule cannot overcome the independence of the $\{\tilde{\alpha}_i\}_{i=1}^n$.

3.2 Maximally informative principles-based standards

Next, we consider a principles-based classification standard. Given the firm's choice of $\widehat{\alpha}$, the firm's report (D, U, L, R) reveals that there are L negative transactions and $U - R$ positive ones below the cutoff $\widehat{\alpha}$. In this way, the principles-based standard provides information on the ordering of the realized $\{\alpha_i\}_{i=1}^n$.

For convenience, for $i \in \{1, \ldots, n\}$, let

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

In other words, rank the present value factors from smallest to largest. Then $\widetilde{\alpha}_{(i)}$ is the i^{th}-smallest present value factor. Without loss of generality, the firm's chosen cutoff $\widehat{\alpha}$ can be set to $\tilde{\alpha}_{(i)}$ for some $i \in \{1, \ldots, n\}$. Because each $\tilde{\alpha}_i$ is an independent uniform draw from $[1, (1+\rho)/\rho]$, it follows that (see Arnold et al., 2008)

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

Let $x_{(i)}$ be the transaction with weight $\tilde{\alpha}_{(i)}$. It is clear that the report (D, U, L, R) under a principles-based standard reveals information about the $x_{(i)}$. A maximally informative principles-based report would reveal all of the $x_{(i)}$. With this information, the investor's value of the firm is

$$\begin{aligned} v_{os} &= \sum_{i=1}^{n} E[\tilde{\alpha}_{(i)}]x_{(i)} \\ &= \pi + \frac{1}{\rho(n+1)} \sum_{i=1}^{n} i x_{(i)} \end{aligned} \tag{6}$$

The following shows cases in which the principles-based report is maximally informative:

Proposition 2. *Suppose* $\min\{U, D\} \leq 1$. *Then there is a report* (D, U, L, R) *that is sufficient for* v_{os}.

Given that the firm can, in some cases, fully reveal v_{os}, we can compare the informativeness of prices under a principles-based standard to those under the optimal rules-based standard:

Theorem 3. *If a report* (D, U, L, R) *under a principles-based standard is maximally informative, then the mean-squared error in the investor's estimate of* $\widetilde{v}$*, conditional on the report, is*

$$\frac{n}{6\rho^2(n+1)} \tag{7}$$

In this case, prices under a principles-based standard are (strictly) more informative than prices under a rules-based standard if and only if the number of transactions $n > 7$ *(respectively* $n \geq 7$*).*

Table 1 compares the mean-square error in the investor's price under each regime, given that the principles-based report is maximally informative (as in Proposition 2):

Table 1 – Pricing error under rules-based vs. principles-based standard

n	Error: rules-based $n/(48\rho^2)$	Error: principles-based $n/(6\rho^2(n+1))$
5	$5/(48\rho^2)$	$5/(36\rho^2)$
6	$1/(8\rho^2)$	$1/(7\rho^2)$
7	$7/(48\rho^2)$	$7/(48\rho^2)$
8	$4/(24\rho^2)$	$4/(27\rho^2)$
9	$3/(16\rho^2)$	$3/(20\rho^2)$
10	$5/(24\rho^2)$	$5/(33\rho^2)$

Comparison of price informativeness under a maximally informative rules-based versus a principles-based standard, given that the firm's principles-based report is a sufficient statistic for the sample order statistics. With more than 7 transactions, the principles-based report is more informative.

4 Main Results

4.1 Message space adequacy of principles-based classification

Two aspects of the proof of Proposition 2 are worth mentioning. The first is that the message space is large enough to give the firm distinct reports for each possible value of v_{os}. This property is necessary if pooling does not occur for purely mechanical reasons. The second is that the firm can credibly signal its type. That is, given the firm's value based on its order statistics v_{os}, there exists a report the firm can issue that no lower-valued firm can also issue. We address the first aspect here and the next in the following subsection.

The following shows that the sufficiency of the message space generalizes to an arbitrary number of transactions:

Theorem 4. *There are enough possible distinct principles-based reports* (D, U, L, R) *to communicate every possible valuation based on the order statistics, i.e., every possible value of* v_{os}*. In particular, given* U *positive items and* D *negative ones, there are*

$$U \cdot D + 1 \text{ possible values of } v_{os}, \text{ and}$$

$$U \cdot D + 1 + n \text{ possible distinct principles-based reports.}$$

Moreover, exactly $n+1$ *distinct reports are consistent with* v_{os} *under worst possible ordering.*

The importance of Theorem 4 is that the restriction to a principles-based classification scheme reduces the number of possible reports, but leaves enough to allow a distinct report for each possible firm value. A firm with n transactions has $n!$ possible orderings of its present value factors.[4] If the firm had complete discretion in how to classify the n transactions, there would be 2^n possible reports. Since $n! \geq 3^{n-2}$, the number of possible orderings grows much faster than the number of possible reports. So it might appear that the number of reports would in general be too small, i.e., that there would necessarily be some pooling. Theorem 4 tells us that this is not correct, and the message space is always sufficient.

The reason is that, although there are n present value factors, there are far fewer possible values of v_{os}. The theorem shows that the total number of distinct firm values is

$$(\text{Total income-increasing transactions}) \cdot (\text{Total income-reducing transactions}) + 1$$

Similarly, the restriction to reports that satisfy the principles-based standard is also much

[4]As these come from an atomless continuous distribution, they are almost surely all distinct.

smaller than 2^n. The theorem shows this total is

$$(\text{Total income-increasing transactions}) \cdot (\text{Total income-reducing transactions}) + n + 1$$

In sum, the restriction to a single cutoff is much less constraining than it may appear. In fact, there are n more reports available than possible values, although all of these can be made by a firm of the lowest type. If we treat these as equivalent, Theorem 4 says that the message space is exactly as large as necessary.

4.2 Strategic considerations

We now turn to strategic considerations in reporting under principles-based classifications, that is, whether a firm can credibly reveal its type, given that the firm reports to a skeptical investor with rational expectations (as is standard since Milgrom and Roberts, 1986). In what follows, we characterize the degree to which the market's skepticism disciplines the firm's reporting choice, adding to the report's informativeness.

Proposition 2 shows that the firm can fully reveal its type if either U or D is at most 1. The following example shows that with $U = D = 2$, complete separation is no longer attainable.

Example 1. *Let $U = D = 2$. Suppose*

$$x_{(1)} = x_{(3)} = -1 \qquad x_{(2)} = x_{(4)} = 1$$

Then

$$v_{os} = \frac{2}{5\rho}$$

The firm can satisfy the principles-based standard if and only if it issues one of the following reports:

$$(D, U, L, R) \in \{(2,2,0,2), (2,2,1,2), (2,2,1,1), (2,2,2,1), (2,2,2,0)\}$$

If $L = 0$ *or* $R = 0$*, then the report can be mimicked by a firm with transactions and present value factors* $\{(x'_1, \alpha'_1) \ldots, (x'_4, \alpha'_4)\}$ *satisfying*

$$x'_{(3)} = x'_{(4)} = -1 \qquad x'_{(1)} = x'_{(2)} = 1$$

which would have a lower value ($v'_{os} = -4/(5\rho)$*).*

If $L = R = 1$*, the report could be mimicked by a firm with*

$$x'_{(2)} = x'_{(4)} = -1 \qquad x'_{(1)} = x'_{(3)} = 1$$

which has a lower value of $v'_{os} = -2/(5\rho)$*.*

If $L = 2, R = 1$*, the report could be mimicked by a firm with*

$$x'_{(2)} = x'_{(3)} = -1 \qquad x'_{(1)} = x'_{(4)} = 1$$

and if instead $L = 1, R = 2$*, the report could be mimicked by a firm with*

$$x'_{(1)} = x'_{(4)} = -1 \qquad x'_{(2)} = x'_{(3)} = 1$$

In both these cases, the value of the mimicking firm is 0.

As these cases are exhaustive, the firm has no report that separates it from a worse type.

In light of Example 1, we investigate the amount of information the firm can reveal in equilibrium. The following characterizes the firm's reporting strategy:

Proposition 5. *Let $\underline{v}(D,U,L,R)$ be the market's most skeptical interpretation of the firm's report:*

$$\underline{v}(D,U,L,R) = \sum_{h=1}^{U-R} E[\tilde{\alpha}_{(h)}] - \sum_{i=U-R+1}^{U-R+L} E[\tilde{\alpha}_{(i)}] + \sum_{j=U-R+L+1}^{U-R+D} E[\tilde{\alpha}_{(j)}] - \sum_{k=U-R+D+1}^{U+D} E[\tilde{\alpha}_{(k)}] \tag{8}$$

The firm maximizes this interpretation if and only if the firm chooses its cutoff $\widehat{\alpha}$ in (4) *to maximize*

$$u_f(\widehat{\alpha}) := L(\widehat{\alpha}) \cdot R(\widehat{\alpha}) \tag{9}$$

The intuition of $\underline{v}(\cdot)$ is as follows: after receiving the principles-based report (D,U,L,R), the investor knows that L negative transactions and $U-R$ positive transactions are less persistent than the firm's cutoff. A skeptical interpretation is that the gains are less persistent than the losses, i.e., that $(x_{(1)},\ldots,x_{(U-R)})$ are all positive transactions, and the L next-least persistent items, $(x_{(U-R+1)},\ldots,x_{(U-R+L)})$, are all negative. Skepticism about the transactions above the cutoff is similar: the R revenues are treated as less persistent than the $D-L$ expenses. This puts the least weight consistent with the report on the positive transactions and the most consistent with the report on the negative ones.

Proposition 5 reduces the firm's problem to maximizing the product of total nonrecurring losses and total recurring revenues. Because $u_f(\cdot)$ has this simple form, we are able to obtain sharp bounds on price informativeness under a principles-based standard. For instance, in Example 1, the firm maximizes its value if it chooses $(L,R) \in \{(1,2),(2,1)\}$. Among the firm's feasible reports, these two produce the highest value of u_f.

4.3 Price informativeness

Understanding price informativeness requires us to consider the investor's problem. The investor observes report (D, U, L, R) and prices the firm based on the number of persistence-ordered transactions $(x_{(1)}, \ldots, x_{(n)})$ (where the investor infers $n = U + D$) for which the report is consistent with the principle (2) and maximizes (9). The amount of uncertainty the investor faces is determined by the variance in the order statistics value $v_o s$, given by (7), and by the number of sequences with distinct values that the firm might have, given the report.

To improve readability, we make a harmless change in notation: for each $i \in \{1, \ldots, n\}$, let $y_i := 2x_{(i)} - 1$. This simplifies the subscripts and replaces all -1 values with 0, and is easily seen to be invertible. Thus from here onward, unless it is necessary for interpretation, we work with sequences of bits, understanding a 0 bit to be a transaction of value -1 and understanding bit strings to correspond to transactions ranked from least to most persistent.

We first distinguish between the ordered transaction sequences that the firm could report as (D, U, L, R) and those that the firm necessarily reports as (D, U, L, R). We say a sequence is *legal* if the report is optional, and that the sequence is *strictly legal* if the report is uniquely optimal:

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 1s and D 0s;*

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

3. for every j*,* $0 \le j \le U + D$*, we have* $\lambda_j(y)r_j(y) \le LR$*.*

A legal sequence is strictly legal *if we have*

3' for every j*,* $0 \le j \le U + D$*,* $j \ne l + U - R$*, we have* $\lambda_j(y)r_j(y) < LR$*.*

In Example 1, the report $(U, D, L, R) = (2, 2, 2, 1)$ has two legal sequences: $(0, 1, 0, 1)$ and $(1, 0, 0, 1)$. The first could is also legal with respect to a different report, $(2, 2, 1, 2)$, as shown in the example. So the first is legal with respect to $(2, 2, 2, 1)$ but not strictly legal; the second sequence is strictly legal.

A slightly more complex example is as follows:

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* $\lambda_2(y) = 1$ *and* $r_2(y) = 2$*:* 010011*,* 010101*,* 010110*,* 011001*,* 011010*,* 011100*,* 100011*,* 100101*,* 100110*,* 101001*,* 101010*, and* 101100*. Of those, only* 101100 *and* 101010 *satisfy* $\lambda_j(y)r_j(y) \le 2$ *for* $j = 0, \ldots, 6$*, so they are the only legal sequences. Furthermore, only* 101100 *is strictly legal.*

To characterize the set of legal sequences and the number of legal sequences associated with a report, we proceed as follows: denote by $\mathcal{G}_{L,R}(U, D)$ (resp. $\mathcal{G}'_{L,R}(U, D)$) the set of all legal (resp. strictly legal) sequences with respect to (U, D, L, R), and write $g_{L,R}(U, D) = |\mathcal{G}_{L,R}(U, D)|$, $g'_{L,R}(U, D) = |\mathcal{G}'_{L,R}(U, D)|$. If $L > D$, $L < 0$, $R > U$ or $R < 0$, we write $\mathcal{G}_{L,R}(U, D) = \mathcal{G}'_{L,R}(U, D) = \varnothing$ and $g_{L,R}(U, D) = g'_{L,R}(U, D) = 0$.

Then we have the following theorem:

Theorem 6. *We have*

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

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

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

Similarly,

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

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

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

Theorem 6 allows for fast computation of $g_{L,R}(U,D)$ and $g'_{L,R}(U,D)$. For example, we can compute $\gamma_{5,8}(12) = 1107$, $\gamma_{8,5}(15) = 248$, $\gamma'_{5,8}(12) = 927$, $\gamma'_{8,5}(15) = 196$ and therefore $g_{5,8}(15,12) = 274536$ and $g_{5,8}(15,12) = 181692$.

We see from the theorem that the formulas for γ and γ' allow for separately computing the nonrecurring portion of the sequence and the recurring portion; that is, the portions above and below the cutoff can be chosen independently. This is a consequence of the following lemma:

Lemma 7. *For given U, D, L, R, we have $(y_1, \ldots, y_{U+D}) \in \mathcal{G}_{L,R}(U, D)$ if and only if*

$$(y_1, \ldots, y_{L+U-R}, \underbrace{1, \ldots, 1}_{R}) \in \mathcal{G}_{L,R}(U, L) \quad \& \quad (\underbrace{0, \ldots, 0}_{l}, y_{L+U-R+1}, \ldots, y_{U+D}) \in \mathcal{G}_{L,R}(R, D).$$

Consequently,

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

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

Finally, given that we can characterize the legal sequences associated with a report, we can discuss the degree to which the report is opaque.

We first observe that the ratio of the number of strictly legal sequences to the number of legal sequences appears to converge to one as D, U, L, R become large. In this sense, the amount of pooling becomes small asymptotically.

To say more about the informativeness of a report, we define the opaqueness of a sequence as follows: For a sequence $y = (y_1, \ldots, y_n) \in \{0, 1\}^n$, define

$$w(y) = \sum_{i=1}^{n} i w_i,$$

i.e. $w(a)$ is the sum of the positions of 1's in a. Define the *opaqueness*

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

Theorem 8. *We have*

$$\min_{y \in \mathcal{G}_{L,R}(U,D)} w(y) = LR + \binom{U+1}{2}$$

and

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

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

$$[LR + \tbinom{U+1}{2}, LR + \tbinom{U+1}{2} + o(U,D,L,R)].$$

5 Discussion and conclusion

Classifications are designed to provide information to the end users of financial reports. It may seem surprising that standard setters do not generally provide definitive guidance on classifying transactions as recurring or transitory. We show, however, that a principles-based classification scheme makes prices more informative than an optimally chosen rules-based scheme.

References

Abarbanell, J. S. and R. Lehavy (2007). Letting the "tail wag the dog": The debate over GAAP versus street earnings revisited. *Contemporary Accounting Research 24* (3), 675–723.

Arnold, B. C., N. Balakrishnan, and H. N. Nagaraja (2008). *A First Course in Order*

Statistics, Volume 54 of *Classics In Applied Mathematics.* Philadelphia: Society for Industrial and Applied Mathematics.

Barnea, A., J. Ronen, and J. Sadan (1976). Classificatory smoothing of income with extraordinary items. *The Accounting Review 51* (1), 110–122.

Barua, A., S. Lin, and A. M. Sbaraglia (2010). Earnings management using discontinued operations. *The Accounting Review 85* (5), 1485–1509.

Bradshaw, M. T. and R. G. Sloan (2002). GAAP versus the street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research 40* (1), 41–66.

Breuer, M. and D. Windisch (2019). Investment dynamics and earnings-return properties: A structural approach. *Journal of Accounting Research 57* (3), 639–674.

Brown, L. D. and K. Sivakumar (2003). Comparing the value relevance of two operating income measures. *Review of Accounting Studies 8* (4), 561–572.

Cain, C. A., K. Kolev, and S. E. McVay (2020). Detecting opportunistic special items. *Management Science 66* (5), 1783–1801.

Caplan, D. and M. Kirschenheiter (2004). A model of auditing under bright-line accounting standards. *Journal of Accounting, Auditing & Finance 19* (4), 523–559.

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

Donelson, D. C., A. Kartapanis, and C. Koutney (2020). SEC non-GAAP comment letters and firm disclosures. Working paper, University of Texas at Austin, Texas A&M University, and George Mason University.

Doyle, J. T., R. J. Lundholm, and M. T. Soliman (2003). The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies 8*(2–3), 145–174.

Dye, R. A. (2002). Classifications manipulation and Nash accounting standards. *Journal of Accounting Research 40*(4), 1125–1162.

Fan, Y., A. Barua, W. M. Cready, and W. B. Thomas (2010). Managing earnings using classification shifting: Evidence from quarterly special items. *The Accounting Review 85*(4), 1303–1323.

Fan, Y. and X. Liu (2017). Misclassifying core expenses as special items: Cost of goods sold or selling, general, and administrative expenses? *Contemporary Accounting Research 34*(1), 400–426.

Fan, Y., W. B. Thomas, and X. Yu (2019). The impact of financial covenants in private loan contracts on classification shifting. *Management Science 65*(8), 3637–3653.

FASB (1985). Elements of financial statements. Statement of Financial Accounting Concepts No. 6, Financial Accounting Standards Board.

Gao, P. (2017). Optimal thresholds in accounting recognition standards. Working paper, University of Chicago.

Gao, P., H. Sapra, and H. Xue (2020). A model of principles-based vs. rules-based standards. Working paper, University of Chicago and Duke University.

Gonedes, N. J. (1975). Risk, information, and the effects of special accounting items on capital market equilibrium. *Journal of Accounting Research 13*(2), 220–256.

Gu, Z. and T. Chen (2004). Analysts' treatment of nonrecurring items in street earnings. *Journal of Accounting and Economics 38*, 129–170.

Hemmer, T. and E. Labro (2019). Management by the numbers: A formal approach to deriving informational and distributional properties of "unmanaged" earnings. *Journal of Accounting Research 57*(1), 5–51.

Hiemann, M. (2020). Earnings and firm value in the presence of real options. *The Accounting Review Forthcoming*.

McVay, S. E. (2006). Earnings management using classification shifting: An examination of core earnings and special items. *The Accounting Review 81*(3), 501–531.

Milgrom, P. and J. Roberts (1986). Relying on the information of interested parties. *RAND Journal of Economics 17*(1), 18–32.

Nelson, M. W., J. A. Elliott, and R. L. Tarpley (2002). Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review 77*(Supplement), 175–202.

Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura (1990). Strategic information revelation. *Review of Economic Studies 57*(1), 25–47.

PricewaterhouseCoopers (2019). Financial statement presentation.

Ribeiro, A., Y. Shan, and S. Taylor (2019). Non-GAAP earnings and the earnings quality trade-off. *Abacus 55*(1), 6–41.

Ronen, J. and S. Sadan (1975). Classificatory smoothing: Alternative income models. *Journal of Accounting Research 13*(1), 133–149.

Shin, H. S. (1994). News management and the value of firms. *RAND Journal of Economics 25*(1), 58–71.

Shin, H. S. (2003). Disclosure and asset returns. *Econometrica 71*(1), 105–133.

A Proofs

Proof of Proposition 1. By independence, it is enough to focus on the case where $n = 1$. Let $p \in (0, 1)$ be the probability that $\widetilde{x} = 1$. The cumulative distribution of the firm value $\widetilde{v} = \widetilde{x} \cdot \widetilde{\alpha}$ is easily verified to be

$$F(v) = \begin{cases} (1-p)[1+\rho(v+1)], & v \in \left[-\frac{1+\rho}{\rho}, -1\right] \\ 1-p, & v \in (-1, 1) \\ 1-p+p\rho(v-1), & v \in \left[1, \frac{1+\rho}{\rho}\right] \end{cases} \quad (10)$$

Given bright-line cutoff α_{bl}, probability that $\widetilde{x}$ is classified as recurring revenue R is

$$Pr(\widetilde{x} = 1 \text{ and } \widetilde{\alpha} \geq \alpha_{bl}) = Pr\left(\widetilde{v} \in \left[\alpha_{bl}, \frac{1+\rho}{\rho}\right]\right) = 1 - F(\alpha_{bl}) = p[1 - \rho(\alpha_{bl} - 1)]$$

Similarly, the probability that $\widetilde{x}$ is recorded as recurring expenses E is

$$Pr(\widetilde{x} = -1 \text{ and } \widetilde{\alpha} \geq \alpha_{bl}) = Pr\left(\widetilde{v} \in \left[-\frac{1+\rho}{\rho}, -\alpha_{bl}\right]\right) = F(-\alpha_{bl}) = (1-p)[1 - \rho(\alpha_{bl} - 1)]$$

In each of these cases, the investor's squared error is identical. The investor's estimate is the midpoint of the interval, and the intervals are of the same length. The report reveals whether $x = 1$ or $x = -1$, so the distribution of the firm value given the subinterval inherits the uniform distribution from $\widetilde{\alpha}$. In sum, conditional on the transaction being classified as recurring (whether as R or as E), the investor's squared error is the variance in the subinterval, which equals

$$\frac{\left(\frac{1+\rho}{\rho} - \alpha_{bl}\right)^2}{12}$$

By an analogous argument, conditional on the transaction being classified as nonrecurring (whether as G or as L), the investor's squared error is

$$\frac{(\alpha_{bl} - 1)^2}{12}$$

Therefore, the mean squared error is

$$\begin{aligned}
\eta(\alpha_{bl}) &= [1 - \rho(\alpha_{bl} - 1)] \cdot \frac{\left(\frac{1+\rho}{\rho} - \alpha_{bl}\right)^2}{12} + \rho(\alpha_{bl} - 1) \cdot \frac{(\alpha_{bl} - 1)^2}{12} \\
&= [1 - \rho(\alpha_{bl} - 1)] \cdot \frac{\left(\frac{1+\rho}{\rho} - \alpha_{bl}\right)^2}{12} + \frac{(\alpha_{bl} - 1)^3}{12}
\end{aligned}$$

Differentiating with respect to α_{bl} and setting to zero gives a unique critical point at

$$\alpha_{bl}^* = \frac{1 + 2\rho}{2\rho} = 1 + \frac{1}{2\rho}$$

The second derivative is easily checked to equal 1/2, so α_{bl}^* is the unique global minimum.

Finally, the variance of a uniform draw over $[1, 1 + 1/(2\rho)]$ (or over $[1 + 1/(2\rho), 1 + 1/\rho]$) is $1/(12(2\rho)^2) = 1/(48\rho^2)$. Over n independent draws, the total variance is $n/(48\rho^2)$. □

Proof of Proposition 2. Let $D = 0$, so that $U = \pi = n$. Then each $x_{(i)} = 1$, and

$$\begin{aligned} v_{os} &= n + \frac{1}{\rho(n+1)} \sum_{i=1}^{n} i \\ &= n + \frac{1}{\rho(n+1)} \cdot \frac{n(n+1)}{2} = n + \frac{n}{2\rho} \\ &= n \left(\frac{1+2\rho}{2\rho} \right) \end{aligned}$$

Analogously, if $U = 0$, then

$$v_{os} = -n \left(\frac{1+2\rho}{2\rho} \right)$$

If $U = 1$, let k be the persistence ranking of the lone positive transaction, so that $x_{(k)} = 1$ and, for $i \neq k$, $x_{(i)} = -1$. Therefore, the firm can report

$$(D, U, L, R) = (D, 1, k-1, 1)$$

The firm can issue this report only if the lone positive transaction is at least the k^{th}-least persistent item, so it cannot be mimicked by a firm with the same values of (D, U) with a lower value of v_{os}. Similarly, if the investor believes the firm reports according to this strategy, then the firm cannot deviate and mimic a firm with a higher value of v_{os}. Thus, the investor knows that

$$\begin{aligned} v_{os} &= 1 - D + \frac{1}{\rho(n+1)} \left[-\sum_{i=1}^{k-1} i + k - \sum_{j=k+1}^{n} j \right] \\ &= 1 - D + \frac{1}{\rho(n+1)} \left[-\sum_{i=1}^{n} i + 2k \right] \\ &= \pi - \frac{n}{2\rho} + \frac{2k}{\rho(n+1)} \end{aligned}$$

The case of $D = 1$ is analogous. $\square$

Proof of Theorem 3. We begin by the following useful lemma on the distribution of the $\widetilde{\alpha}_{(i)}$, which is a trivial modification of a standard result in the theory of order statistics (see Arnold et al., 2008). We provide an elemantary proof here.

Lemma 9. *For each $i \in \{1, \dots, n\}$,*

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

Consequently,

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

Proof of Lemma 9. For each i, $\widetilde{\alpha}_i \sim U[1, (1+\rho)/\rho]$, so

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

and $\rho(\widetilde{\alpha} - 1)_{(i)} = \rho(\widetilde{\alpha}_{(i)} - 1)$, i.e., the order is the same as the ordering of the $\widetilde{\alpha}_{(i)}$. The pdf $f_{(i)}(\cdot)$ of the i^{th} sample order statistic of n iid random variables with distribution $F(\cdot)$ and density $f(\cdot)$ is

$$f_{(i)}(t) = \frac{n!}{(i-1)!(n-i)!} F^{i-1}(t)[1 - F(t)]^{n-i} f(t)$$

For a $U[0, 1]$ variable, $F(t) = t, 1 - F(t) = 1 - t$, and $f(t) = 1$. Therefore, using $\Gamma(k) =$

$(k-1)!$ for positive integer k,

$$f_{(i)}(t) = \frac{\Gamma(n+1)}{\Gamma(i)\Gamma(n-i+1)} t^{i-1}(1-t)^{n-i}$$

which is the pdf of a Beta$(i, n-i+1)$-distributed random variable.

The mean and variance of a Beta(a, b)-distributed random variable are $a/(a+b)$ and $ab/[(a+b)^2(a+b+1)]$. Therefore,

$$E[\rho(\widetilde{\alpha}_{(i)} - 1)] = \frac{i}{n+1} \quad \Rightarrow \quad E[\widetilde{\alpha}_{(i)}] = 1 + \frac{i}{\rho(n+1)}$$
$$Var[\rho(\widetilde{\alpha}_{(i)} - 1] = \frac{i(n-i+1)}{\rho^2(n+1)^2(n+2)}$$

■

From Lemma 9, the investor's mean-squared error in estimating $\widetilde{v}$ given a report that reveals the order statistics is

$$\begin{aligned}\sum_{i=1}^{n} \frac{i(n-i+1)}{\rho^2(n+1)^2(n+2)} &= \frac{1}{\rho^2(n+1)^2(n+2)}\left[(n+1)\sum_{i=1}^{n} i - \sum_{i=1}^{n} i^2\right] \\ &= \frac{n(n+1)^2/2 - n(n+1)(2n+1)/6}{\rho^2(n+1)^2(n+2)} \\ &= \frac{n}{6\rho^2(n+1)} \qquad (11)\end{aligned}$$

From Proposition 1, the investor's mean-squared error in estimating $\widetilde{v}$ from a fully revealed bright-line classification standard is $n/(48\rho^2)$. This error is weakly greater than that of learning the order statistics (11) if and only if

$$\frac{n}{48\rho^2} \geq \frac{n}{6\rho^2(n+1)}$$
$$\Leftrightarrow n \geq 7$$

□

Proof of Theorem 4. An immediate corollary of Lemma 9 is that, for $i \in \{1, \ldots, n-1\}$,

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

i.e., the expected weights are evenly spaced. The number of possible values of the firm is therefore equivalent to the number of possible ways to sum U integers chosen from $\{1, \ldots, n\}$, i.e., to determine the possible weighted sums of U positive items. This total plus knowledge of U and n are sufficient for determining the sum of the negative items and therefore of the firm's present value.

The smallest possible sum of U integers from $\{1, \ldots, n\}$ is

$$\sum_{i=1}^{U} i = \frac{U(U+1)}{2}$$

The largest possible sum of U integers from $\{1, \ldots, n\}$ is

$$\sum_{i=n-U+1}^{n} i = \sum_{i=1}^{n} i - \sum_{j=i}^{n-U} j = \frac{n(n+1)}{2} - \frac{(n-U)(n-U+1)}{2}$$
$$= \frac{U(2n-U+1)}{2}$$

Because the integers in $\{1,\ldots,n\}$ are consecutive, any integer sum between the smallest and largest possible sum is attainable. Therefore, the total number of possible sums is one plus the difference between the largest and smallest possible sum. Noting that $D = n - U$, this total is (after some substitutions)

$$U \cdot D + 1$$

With a binary classification system, a total of U positive items can be split into two nonnegative subtotals, R and $U - R$, in $U + 1$ possible ways. Similarly, a total of D negative items can be split into two nonnegative subtotals in $D+1$ ways. Therefore, the firm's total number of possible ways to classify U positive and D negative transactions is

$$(U+1)(D+1) = U \cdot D + n + 1$$

Finally, among the $U \cdot D + n + 1$ possible reports, the firm has $U+1$ with no recurring revenue reported (i.e., with R set to 0) and $D + 1$ with all negative items reported as recurring. Exactly one report is consistent with both, so the firm has $n+1$ possible classifications that satisfy the principles-based standard if all income-increasing items are less persistent than all income-reducing items. $\square$

Proof of Proposition 5. From (8),

$$\begin{aligned}\underline{v}(D,U,L,R) = &\sum_{h=1}^{U-R} E[\tilde{\alpha}_{(h)}] - \sum_{i=U-R+1}^{U-R+L} E[\tilde{\alpha}_{(i)}] \\ &+ \sum_{j=U-R+L+1}^{U-R+D} E[\tilde{\alpha}_{(j)}] - \sum_{k=U-R+D+1}^{U+D} E[\tilde{\alpha}_{(k)}]\end{aligned}$$

After repeated use of the identity that $\sum_{i=j+1}^{m} i = (m(m+1) - j(j+1))/2$, rearranging, and substituting $n = U + D$ and $\pi = U - D$, this reduces to

$$\underline{v}(D,U,L,R) = \overbrace{\pi - \frac{n}{2\rho} + \frac{U(U+1)}{\rho(n+1)}}^{\text{constant}} + \underbrace{\frac{2}{\rho(n+1)}}_{\text{positive constant}} LR$$

Thus, the investor's most skeptical interpretation of the report (D,U,L,R) is strictly monotone in LR. Maximizing this interpretation is therefore equivalent to maximizing the Cobb-Douglas utility $u_f(\widehat{\alpha})$ given in (9). □

Proof of Lemma 7. Write

$$y = (y_1, \ldots, y_{U+D}), \quad y' = (y_1, \ldots, y_{L+U-R}, \underbrace{1, \ldots, 1}_{R}), \quad y'' = (\underbrace{0, \ldots, 0}_{L}, y_{L+U-R+1}, \ldots, y_{U+D}).$$

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

Proof of Theorem 6. Theorem 6 follows from Lemma 7 and the three additional lemmas. First, define $\gamma_{L,R}(m) = g_{L,R}(R,m)$ and $\gamma'_{L,R}(m) = g'_{L,R}(R,m)$. The following is obvious:

Lemma 10. *For given (U,D,L,R), a sequence $(y_1, \ldots, y_{U+D})$ is in $\mathcal{G}_{L,R}(U,D)$ if and only if $(1-y_{U+D}, \ldots, 1-y_1)$ is in $\mathcal{G}_{R,L}(D,U)$. Consequently, $g_{L,R}(U,L) = g_{R,L}(L,U) = \gamma_{R,L}(U)$ and*

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

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

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 $\gamma_{L,R}(L) = \gamma'_{L,R}(L) = L$. The following characterizes $\mathcal{G}_{L,R}(R,m)$ and $\mathcal{G}'_{L,R}(R,m)$.

Lemma 11. *For given m, L, R, we have $y = (y_1, \ldots, y_{R+m}) \in \mathcal{G}_{L,R}(R,m)$ if and only if:*

1. *$y_1 = \ldots = y_L = 0$;*
2. *there are exactly R 1's in $(y_{L+1}, \ldots, a_{R+m})$;*
3. *for $i = 1, \ldots, R$, the number of 0's 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 $\mathcal{G}'_{L,R}(R,m)$, with (3) replaced by

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

Proof. First we introduce some notation: for integer ℓ, let $[\ell] := \{1, \ldots, \ell\}$. Assume first that $y \in \mathcal{G}_{L,R}(R,m)$. Then all 1's 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 $\lambda_{j_i-1}(y) r_{j_i-1}(y) \leq LR = L(R-i+1) + (i-1)L$, so $\lambda_{j_i-1}(y) \leq L + (i-1)L/(R-i+1)$.

Conversely, suppose that the three conditions are satisfied. Clearly, there are m 0's and R 1's in y, and $\lambda_L(y) r_L(y) = LR$. Again, denote by j_i, $i \in [R]$, 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 $\lambda_j(y) \leq \lambda_{j_i}(y) \leq L + (i-1)L/(R-i+1)$ and

$r_j(y) = R - i + 1$, so $\lambda_j(y) r_j(y) \leq LR$.

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

□