

A Theory of Classification Shifting

Abstract

This article demonstrates that managers can influence the market by classifying some items as core earnings and others as non-core. Investors react to classifications because managers have incomplete discretion over how to classify results. Managers optimally use their discretion to pool good news with items mandatorily classified as core earnings and bad news with items mandatorily classified as non-core. Aggregation reduces this temptation to classify strategically, provided managers also have incomplete discretion over how to aggregate. That is, managers can use aggregation to separate from strategic classifiers. Our results provide empirical implications for the cross-sectional properties of financial reports.

Key words: Aggregation; Classification Shifting; Simpson's Paradox

I. Introduction

Financial disclosures contain both quantitative and qualitative information. While quantitative information is subject to *uncertainty*, the qualitative information inherent in classifications is subject to *vagueness*. For example, transitory vs. permanent, operating vs. nonoperating, or expense vs. capitalization classifications are not necessarily based on numerical thresholds – because generally accepted measures of ‘permanent-ness,’ etc., do not typically exist. Yet these classifications serve to organize balance sheets, income statements and the non-GAAP disclosures which are becoming more common. While certain items may be confidently classified – this is why classifications are useful in everyday discourse – an accounting standard which fixes a limited number of categories for describing transactions will inevitably result in gray areas.¹

We develop a model of opportunistic classification which gives a firm discretion over how to classify those items (transactions or events) which fall into a gray area. To the extent that the firm has more than one item to classify, we assume that classifications must be consistent— if an item is classified (not classified) as a member of a category, then items with stronger (weaker) membership must also be included (not included) in the classification.² We also examine the effect of aggregation – it is common practice to aggregate the quantitative assignments (i.e., sum them up) for similarly classified items before disclosing them. We show that consistency and aggregation can both increase the informativeness of disclosure by limiting the benefit to the firm from exercising opportunistic discretion. Noting that consistency requires that more than one item be classified, the former result suggests that adding an item to be disclosed can improve the overall informativeness of the disclosure. If classifying many (versus few) items suggests the evolution of an accounting routine, then the result also suggests that firms which have been able to establish accounting routines may have better disclosures. The latter result on aggregation contrasts more

¹ For example, it may be easy to point out a person in a crowd if we can identify that characteristic for which the individual is an extreme example (color of clothing, hair style, accessories, etc.). The task becomes problematic, however, if, we forced to use a single category (e.g., tall/short). See Penno (2008).

² This is commonly referred to as *a fortiori* reasoning.

conventional views of aggregation (that ignore opportunistic behavior) which hold that aggregation destroys information.³

The paper is related to the economic literature which recognizes the significance of classification. They include Mullainathan et al. (2008) who study how individuals group situations into categories and apply the same model of inference to all situations within a category (coarse thinking), Barberis and Shleifer (2003) who study the pervasive classification which occurs in financial markets (style investing), and Barberis et al. (1998) who consider investors who consider only limited number of ways to classify stock price movement. In terms of the accounting literature, our results on aggregation expand the research initiated by Dye and Sridhar (2004) who examine the efficiency of aggregating soft and hard information. Our Section II result on Simpson's Paradox is closely related to Sunder (1983), who provides an accounting application (cost allocation across activities) of the paradox.

For the remainder of the paper, Section II provides the intuition behind discretionary classification of a single item in isolation. We model an 'item' as a two-dimensional transaction or event with a generally accepted numerical assignment for the first (hard) dimension. The second dimension is soft, and a numerical threshold for that dimension cannot be established. Instead, the second dimension is used for classification purposes. For example, when disclosing non-GAAP earnings, a firm may choose to treat a numerical loss (first dimension) as transitory (second dimension) by excluding it from non-GAAP earnings. If the second dimension's underlying characteristic is continuous, a gray area will result where the item's classification makes the transition from transitory to permanent.⁴ Here the classifications become uninformative as the amount of discretion reaches its maximum level. Sections III and IV model the setting where the firm has two items to be classified. This extension permits us to contrast to the simple illustration, and to explain how certain classifications now are informative as the amount of discretion reaches its

³ For example, by Ijiri (1967) and Sorter (1969). Ijiri (p. 120) suggests that "any aggregation generally involves the loss of information in that the resulting total 'value' may be composed of many – possibly infinitely many – different components." Sorter (p. 13) adds that "less rather than more aggregation is appropriate" for the financial statements, and that it is up to the user to do the aggregation himself if the need so arises.

⁴ For example, Dichev et al.(2013, p. 13) indicate that a key issue with an interviewed standard setter was whether, and how, to treat non-permanent items.

maximum level. Section V introduces the notion of within-category aggregation and Section VI concludes.

II. A Single Transaction Illustration

The disciplining effect of consistency is possible only when a firm must classify more than one item. To provide a benchmark and simple introduction to our modeling of such classifications, consider a simple setting where there is a *single* item is to be classified. That is, the firm observes the realization of a random but quantifiable economic event or transaction, say a cash flow, \tilde{x} . This economic event or transaction must be reported to a risk-neutral investor where the true effect of both variables on the value of the firm is simply αx , where $\tilde{\alpha} \in \left[0, \frac{2}{\sqrt{\pi}}\right]$, and $\tilde{\alpha}$ is uniformly distributed and independent of \tilde{x} .⁵ We require that α represent a piece of evidence observable and understood by both the firm and a third party enforcer (e.g., an auditor) rather than a subjective estimate that exists only in the mind of management. But unlike x itself, α lacks a generally accepted reporting mechanism for its specific value. Instead, we assume that the firm classifies α as either L (low) or H (high). We further assume that the firm has *limited* discretion over this classification. In particular, whenever α is extremely low, the firm must classify the reported value of \tilde{x} as an L , and when α is extremely high, the firm must classify the reported value of \tilde{x} as an H . That is, as α becomes extreme, the third party will force the appropriate disclosure. However, for intermediate values of α , the firm has full discretion and may report α as either an L or an H . We represent the firm's degree of discretion with the interval depicted in Figure 1.

-- Place Figure 1 about here --

Any x with an $\alpha < \frac{1}{\sqrt{\pi}} - \varepsilon$ must be classified and reported as an L , and any x with an $\alpha > \frac{1}{\sqrt{\pi}} + \varepsilon$ must be classified and reported as an H . Accordingly, the exogenous parameter ε

⁵ The independence assumption is made for parsimony and provides us with an intuitive benchmark.

models the degree of discretion present in the classification scheme.⁶ We assume that because the value of the firm is αx , a firm that wishes to present itself in the most favorable light will choose that classification which maximizes the value of the firm as determined by the investor on the basis of the reported classification $C \in \{L, H\}$ and x . The investor, in turn, estimates the expected value of $\tilde{\alpha}$ conditional on x and its inference of the firm's classification strategy. The firm, in turn, makes its classification given the investor's expectations, resulting in a Bayesian Nash equilibrium. It is straightforward to verify that the following is such an equilibrium:

-- Place Table 1 about here --

Whenever $x < 0$, the firm will use all of its available discretion to report $C = L$ so as to mitigate the effect of the investor's perceived weight applied to the negative value of x . Thus, the firm will report classification L whenever $x < 0$ and $\alpha \in \left[0, \frac{1}{\sqrt{\pi}} + \varepsilon\right]$, which is the union of the lower non-discretionary region and the discretionary region. Therefore the expected value of $\tilde{\alpha}$ given $x < 0$ and $C = L$ equals $\frac{1}{2} \left(\frac{1}{\sqrt{\pi}} + \varepsilon \right)$. Similar reasoning establishes the rest of Table 1, which provides the equilibrium weights applied by the investor to the various realizations of \tilde{x} . An important feature of this simple model is that the classifications become uninformative as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$.

An interesting feature of this particular setting is that *conditional* on the classification being L or H , the investor's weight attached to a negative x is always greater (by an amount, ε) than the weight attached to a positive x . Consequently, if we only compare reported items within a classification, this will lead to the conclusion that for this simple setting that *bad news has a higher weight than good news*. Upon returning to the model, this result might be surprising given that on average, due to the independence of $\tilde{\alpha}$ and \tilde{x} , good and bad news should both have a weight of

⁶ To illustrate, firms with established business models, or members of a very competitive (and standardized) industry may have small ε values, while emerging technologies or first movers may have large ε values.

$\frac{1}{\sqrt{\pi}}$. While the result may appear to suggest conservatism, it is actually a manifestation of *Simpson's Paradox*, in which a trend that appears in different groups of data disappears when these groups are combined. It is interesting to note here that the Simpson's Paradox effect is a direct implication of focusing on classification since, by its definition, Simpson's Paradox requires that classifications be made before a paradox can exist.⁷

III. Model

Allowing for two dimensions permits us to not only examine the effect of extreme items, as we did above, but just as importantly, to consider the disciplining effect of natural orderings. That is, items to be classified into one of two categories are generally not identical and frequently can be ordered as to which category is more or less representative of that item. We model this setting with a circle of area one circumscribed within a square. See Figure 2.

-- Place Figure 2 about here --

For this setting, the firm reports the realizations of two random variables, \tilde{x}_1 and \tilde{x}_2 , and privately observes the realization random variables, $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$, where $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ is uniformly distributed over a *circle* with radius $\frac{1}{\sqrt{\pi}}$, which has an area of one, and where $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ is independent of $(\tilde{x}_1, \tilde{x}_2)$.

The true underlying value of the firm equals $\alpha_1 x_1 + \alpha_2 x_2$. Similar to the one-item illustration, the firm must classify each α_i realization as either an *L* or an *H*, and does so to maximize the outside risk-neutral investor's inferred expected value. The classification decision is designated as (C_1, C_2) where $C_i \in \{L, H\}$ is the classification for x_i . The investor, in turn, estimates the expected values of $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ conditional on the disclosure of x_1 and x_2 , the disclosed classification, and the inferred classification strategy of the firm. The firm, in turn makes its classification given the investor's expectation, resulting in a Bayesian Nash equilibrium.

⁷ See Sunder (1988) for another accounting illustration of Simpson's Paradox.

The inner circle with a radius of ε represents the realizations of $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ for which classification discretion is allowed. For reference, the large circle (radius $\frac{1}{\sqrt{\pi}}$) has been divided into four quadrants. If $\varepsilon = 0$, then the firm would have no discretion and must report classifications (H, H) , (L, H) , (L, L) and (H, L) when the (α_1, α_2) realization falls into quadrants I, II, III, and IV respectively. As ε increases, the regions of no-discretion shrink. The outer unshaded ring represents (α_1, α_2) realizations for which no classification discretion is never allowed. We assume the following property:

Consistency: The firm is prohibited from reporting (H, L) when $\alpha_1 < \alpha_2$, and prohibited from reporting (L, H) when $\alpha_1 > \alpha_2$.

Consistency reflects the idea that an observer such as an auditor is able to examine α_1 and α_2 , and recognize whether one is less than the other and that any reported classification should not violate that ordering. The ordering property essentially requires that the firm be consistent and keep its story straight. While the probability that $\alpha_1 = \alpha_2$ is zero implies that the realizations (α_1, α_2) will be strictly ordered with probability one, those realizations for which $\alpha_1 = \alpha_2$ create a useful boundary, as discussed below. To appreciate the implications of the ordering property, consider Figure 3, which depicts a cross-hatched semicircle:

-- Place Figure 3 about here --

The dashed diagonal represents all (α_1, α_2) pairs for which $\alpha_1 = \alpha_2$, and divides the discretionary region depicted in Figure 2 into two semicircles of equal area. For all (α_1, α_2) falling into the interior of cross-hatched semicircle, we have $\alpha_1 > \alpha_2$, and for all (α_1, α_2) falling into the interior of complementary (white) semicircle, we have $\alpha_1 < \alpha_2$. Thus, by the ordering property, a report of (L, H) is not allowed when (α_1, α_2) lies in the cross-hatched semicircle. Similarly, when (α_1, α_2) lies in the complementary (white) semicircle, a classification of (H, L) is not allowed.

The ordering property does not apply to an (H, H) or an (L, L) classification. For those cases, as long as (α_1, α_2) lands in *either* semicircle, either an (H, H) or (L, L) classification *is permitted*, because now there is no violation of the ordering property with either classification. Consequently, Figure 3 suggests that reports of (L, H) or (H, L) may be less manipulable than a report of (H, H) or (L, L) , and, as we will see, be more informative.

To illustrate the issues involved, suppose that $x_1 = -1$ and $x_2 = 1$, meaning that, *ceteris paribus*, the firm would like the investor to place the lowest weight possible on x_1 and the highest weight possible on x_2 . Conjecture for now that a reported classification of (L, H) results in the investor's highest inferred equilibrium value for the firm and a reported classification of (H, L) results in the lowest equilibrium value. If the firm reports (L, H) , the investor uses its knowledge of this preference to infer that (α_1, α_2) has fallen into either the white rim for (L, H) or the white semicircle (complementary to the hatched semicircle as depicted in Figure 3) and will calculate the conditional expectations for $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$, and use accordingly to weight x_1 and x_2 . Given that (L, H) was the preferred choice of the firm, suppose next that the firm reports a classification other than (L, H) . If the firm reports a classification different than (L, H) , the investor infers that (α_1, α_2) did not fall into the above-mentioned areas, because otherwise, the firm would have reported (L, H) . The investor can use this deviation from most preferred classification to make inferences about (α_1, α_2) . If, for example, the actual disclosure were (L, L) , the investor would then know that (α_1, α_2) has fallen into either the segment of the white rim for (L, L) or the cross-hatched semicircle and will calculate the weights for x_1 and x_2 accordingly. If instead, the firm reported (H, H) as the alternative to the preferred classification, the investor would then infer that (α_1, α_2) has fallen into either that segment of the white rim for (H, H) or the cross-hatched semicircle. Finally, suppose that (L, H) was the preferred choice of the firm, but that the firm reports classification (H, L) . If this were the least preferred choice of the firm, the investor knows that the firm has been unable to report its second choice of (H, H) or (L, L) , which would imply that (α_1, α_2) must have fallen into the outer rim for (H, L) . The investor accordingly creates a

hierarchy of firm choices and corresponding conditional inferences. A similar description applies to the case where (H, L) is the most preferred disclosure and (L, H) is the least preferred. Thus for the cases of (L, H) or (H, L) being most preferred, we conjecture a three-level hierarchy:

1. Most preferred classification
2. Second choice classification
3. Least preferred classification

This hierarchy is used by the investor to value x_1 and x_2 .

Next, consider the case where (H, H) or (L, L) is the firm's most preferred disclosure. Suppose that $x_1 = 1$ and $x_2 = 1$, meaning that the firm would like the investor to place the highest possible weights on both x_1 and x_2 . Conjecture that (H, H) is now the firm's preferred disclosure. If, indeed, the firm discloses (H, H) the investor infers that (α_1, α_2) has landed in either the white outer rim for (H, H) or the shaded area of Figure 2, and will calculate its conditional expectation for $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ accordingly. Next suppose that (H, H) was the preferred disclosure, but that another disclosure ((L, L) , (L, H) or (H, L)) has occurred. Now the investor may infer that (α_1, α_2) did not fall in white outer rim for (H, H) and did not fall the shaded area of Figure 2 either, and can pin down the location of $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ precisely to the white rim corresponding to the actual disclosure ((L, L) , (L, H) or (H, L)). A similar description applies to the case where (L, L) is the preferred disclosure. When (H, H) or (L, L) are most preferred, the investor accordingly creates a two-level hierarchy of firm choices and corresponding inferences.

1. Most preferred classification
2. Least preferred classification

Note that when (H, H) or (L, L) are most preferred, the hierarchy is reduced from three levels to two. This occurs because a rejection of (H, H) or (L, L) leaves no remaining discretion, forcing

the firm out to the white rim (least favored), whereas as rejection (L, H) or (H, L) leaves enough discretion to choose an (H, H) or (L, L) before the firm is forced out to the white rim.

To develop a further sense of all of the possible classifications, we next calculate the weights for the most and least preferred classifications for both hierarchies above. (These calculations are found in the appendix.) Suppose first that the firm reports its least preferred classification. Given the definitions made above, this means that all other options are not available and that (α_1, α_2) lies in the corresponding segment of the white outer rim. Due to the symmetry of the problem, we can let $\omega_i \equiv \omega_L$ represent the weight for least preferred classification $C_i = L$, and $\omega_i \equiv \omega_H$ represent the weight for a least preferred classification when $C_i = H$.

$$\omega_L = \frac{3\pi^{3/2}\varepsilon + \pi(3 - 4\varepsilon^2) - 4 - 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)}$$

$$\omega_H = \frac{3\pi^{3/2}\varepsilon + \pi(3 - 4\varepsilon^2) + 4 + 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)}$$

Next suppose that the firm makes its most preferred classification and that it is either (L, H) or (H, L) . In this case, symmetry again permits us to determine that an x_i classified as an L will have weight μ_L , and the other x_j classified as an H will have weight μ_H where

$$\mu_L = \frac{3\pi^2\varepsilon^2 - 4(\sqrt{2} - 1)\pi^{3/2}\varepsilon^3 + 3\pi - 4}{3\pi^{3/2}(1 + \pi^2\varepsilon)}$$

$$\mu_H = \frac{3\pi^2\varepsilon^2 + 4(\sqrt{2} - 1)\pi^{3/2}\varepsilon^3 + 3\pi + 4}{3\pi^{3/2}(1 + \pi^2\varepsilon)}$$

Finally, if a firm makes its most preferred choice (H, H) or (L, L) , then the investor knows that the (α_1, α_2) realization is contained in either in the shaded circle of Figure 2 or the corresponding

quarter of the outer white rim. Let λ_H indicate the weight applied to x_1 and x_2 each, when the classifications are (H,H) and λ_L indicate the weight applied to x_1 and x_2 each, when the classifications are (L,L) ,

$$\lambda_L = \frac{9\pi^2 \varepsilon^2 + 4\pi^{3/2} \varepsilon^3 + 3\pi - 4}{3\pi^{3/2}(1 + 3\pi\varepsilon^2)}$$

$$\lambda_H = \frac{3\pi^2 \varepsilon^2 - 4\pi^{3/2} \varepsilon^3 + 3\pi + 4}{3\pi^{3/2}(1 + \pi\varepsilon^2)}$$

The weights above are represented in Figure 4.

-- Place Figure 4 about here --

from the *ex ante* mean $\frac{1}{\sqrt{\pi}}$ provides an indication of the classification's informativeness. As discretion increases, reporting a least preferred classification becomes more informative as the ω weight deviates from the *ex ante* mean. Furthermore note that as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$, the μ weights do not converge to the uninformative weight, $\frac{1}{\sqrt{\pi}}$. This is due to the fact that even though $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$, a firm reporting (L,H) or (H,L) does not have complete discretion. Finally, the λ weights converge to the uninformative weight $\frac{1}{\sqrt{\pi}}$ as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$, reflecting the absence of discipline arising from of the lack of an ordering limitation on discretion for this type of disclosure.

Consequently, as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$, the (L,L) and (H,H) preferred choices become uninformative, while the (L,H) and (H,L) preferred choices continue to be informative. This contrasts the simple

model of Section II where the classifications become uninformative as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$ and thereby emphasizes the role of multiple items in disciplining the classifications made by an otherwise opportunistic firm. We will see next, however, that as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$, that the most preferred (L, H) and (H, L) classifications become rare.

To see this, consider Figure 5.

-- Place Figure 5 about here --

Figure 5 exhibits the firm's *most preferred classifications* for (x_1, x_2) pairs for $\varepsilon = .3$ and $\varepsilon = .5$. The boundary between (L, H) most preferred and (L, L) most preferred disclosure is found by comparing the investor's valuations assuming that a disclosure is most preferred. For example, the value $x_1\lambda_L + x_2\lambda_H$ is used for an (L, H) disclosure and $x_1\mu_L + x_2\mu_L$ is used for and (L, L) disclosure. The boundary between (L, H) preferred disclosure and (L, L) preferred disclosure is found by setting these two values equal and solving for x_2 as a function of x_1 . The same approach is issued for all of the other boundaries and is summarized in the appendix.

Note how the most preferred status areas for (L, H) and (H, L) shrink with ε . In the limit, as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$, it can be shown that most preferred status for (L, H) and (H, L) vanishes and the graph indicating the areas of most preferred status only for (L, L) and (H, H) which become two equal triangles separated by a diagonal line with a slope equal to one. When this occurs, the weights for x_1 and x_2 converge to the ex ante weight of $\frac{1}{\sqrt{\pi}}$ as noted in Figure 4, and (L, L) and (H, H) are the *only* most preferred disclosures. This implies that overall, any classification becomes uninformative due to the fact the firm is always free to choose its most preferred (L, L) or (H, H) and both of these classifications become uninformative. However, while rare, the classifications (L, H) and (H, L) remain informative as indicated in Figure 4.

While we have assumed that $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ is *independent* of $(\tilde{x}_1, \tilde{x}_2)$, to the extent that quantitative factors (i.e., (x_1, x_2)) influence the investor's interpretation of (C_1, C_2) , the equilibrium classifications and (α_1, α_2) are clearly *dependent*. This becomes even more apparent when we consider the firm's second choice classifications.

IV. Second Choice Classifications

To complete the analysis, consider the second choice disclosures. Consider first the case where the most preferred classification is (L, H) or (H, L) . Imagine a diagonal line drawn on Figure 5 representing all (x_1, x_2) pairs such that $-x_1 = x_2$, descending from the upper left-hand corner, $(-1, 1)$, to the lower right-hand corner, $(1, -1)$. Due to the linearity of the objective functions, it turns out that the firm's second most preferred choice will simply depend on which side of the that diagonal, $x_1 = -x_2$, that (x_1, x_2) lies. Return to the example where $x_1 = -1$ and $x_2 = 1$, and (L, H) the preferred classification. It would seem that for a quadrant II (x_1, x_2) realization falling below the $-x_1 = x_2$ diagonal, that because $|-x_1| > x_2$, the negative event dominates. Consequently, the firm might, as a second choice limited to (L, L) or (H, H) , prefer both weights to be low rather than for both weights to be high, with the reverse holding for a (x_1, x_2) realization falling above the $-x_1 = x_2$ diagonal in quadrant II. Accordingly, we conjecture a second choice of (L, L) for a quadrant II (x_1, x_2) realization falling below the $-x_1 = x_2$ diagonal, and a second choice of (H, H) for a quadrant II (x_1, x_2) realization falling above the $-x_1 = x_2$ diagonal. Conditional on either of these second choice disclosures, the investor may conclude that (α_1, α_2) has fallen into either the hatched semicircle of Figure 3 or the appropriate region in the white outer rim. Finally, if (L, H) is the most preferred disclosure, and a disclosure of (L, L) or (H, H) is not made, the investor may deduce that (α_1, α_2) has fallen into white rim for least favored (L, H) .

The following expressions summarize the equilibria and are displayed in Figure 6. (These are derived in the appendix.)

-- Place Figure 6 about here --

Suppose first that the firm's most preferred choice is (L, H) in quadrant II (x_1, x_2) lies below the diagonal, and its second choice is (L, L) , then a disclosure of (L, L) will have weights (γ, μ_L) where

$$\gamma = \frac{3\pi^2 \varepsilon^2 + 4(\sqrt{2} + 1)\pi^{3/2} \varepsilon^3 + 3\pi - 4}{3\pi^{3/2}(1 + \pi^2 \varepsilon)}$$

Similarly, if the firm's most preferred choice is (L, H) in quadrant II (x_1, x_2) lies above the diagonal, and its second choice is (H, H) , then a disclosure of (H, H) will have weights (μ_H, β) where

$$\beta = \frac{3\pi^2 \varepsilon^2 - 4(\sqrt{2} + 1)\pi^{3/2} \varepsilon^3 + 3\pi + 4}{3\pi^{3/2}(1 + \pi^2 \varepsilon)}$$

A symmetric argument can be made for the case where (x_1, x_2) lands in the region where (H, L) is most preferred. If (x_1, x_2) lies below the diagonal which means that its second choice is (L, L) and disclosure of (L, L) will have weights (μ_L, γ) , and if (x_1, x_2) lies above the diagonal which means that its second choice is (H, H) , a disclosure of (H, H) will have weights (β, μ_H) .

Consider the case where (L, H) is the most preferred choice, but the firm has made (L, L) its choice. When $\varepsilon \approx 0$, the weights (γ, μ_L) are approximately (μ_L, μ_L) reflecting the fact that there is little discretion available and that a disclosure of (L, L) reflects that (α_1, α_2) almost surely lies in the third quadrant of Figure 2 where the conditional expected values of $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ are both μ_L . A symmetric argument applies for a most preferred disclosure of (H, L) .

An implication of Figures 4 and 6 taken together is that while the ω weights representing the least preferred choice are more informative than the λ weight for the two-level hierarchy corresponding to (L, L) and (H, H) as most preferred, when we then consider the second choice

preference corresponding to a three level hierarchy and (L,H) and (H,L) as most preferred, we no longer have an unambiguous ordering of informativeness as we descend the hierarchy. In particular, the γ and β weights move toward, and then away from no-information benchmark of $\frac{1}{\sqrt{\pi}}$ as ε increases, while the companion μ weight is identical to the weight for most preferred (L,H) and (H,L) . Consequently we have shown that it is not always the case that a less preferred classification is more informative than more preferred one (Hirst et al., 1995).

V. Aggregation

We assumed in the previous sections that when an item was classified as (H,H) or (L,L) , the investor was able to observe both x_1 and x_2 . We saw that as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$ the classification overall became uninformative. In this section, we gather results from the previous ones to serves as a comparison to disclosures with within-category aggregation, and will see that by introducing within category aggregation, the classifications will remain informative as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$.

We modify the model by requiring the firm to aggregate x_1 and x_2 by reporting their sum whenever a (L,L) or (H,H) classification is made, the most preferred status will reverse itself with (L,H) and (H,L) becoming the most preferred disclosures as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$. Figure 4 shows that these disclosure remain informative as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$, thus establishing that within-category aggregation may enhance the informativeness of the disclosures relative to the (non-aggregation) model examined above.

To make this statement more precise, suppose that when (H,H) or (L,L) is reported that instead of indicating the values of x_1 and x_2 , the firm reports the sum $\sum x_i$. This assumption recognizes that without sub-categorization, classifying two items into the same category will result in their

aggregation. We retain the assumption, however, that the values of x_1 and x_2 are individually indicated when the classification is (L, H) or (H, L) . To simplify the analysis of within-category aggregation, see Figure 7.

-- Place Figure 7 about here --

Figure 7 assumes that $(\tilde{x}_1, \tilde{x}_2)$ is uniformly distributed over a square of length 2 and mean $(0, 0)$. Suppose for a moment, the investor *only* learned the value of $\sum x_i$. This would allow the investor to calculate expected value of an $\sum x_i$ disclosure as $\left(\frac{\sum x_i}{2}, \frac{\sum x_i}{2}\right)$. For example, suppose that the investor learned only that $\sum x_i = .5$. Then the conditional expected value of $(\tilde{x}_1, \tilde{x}_2)$ is the average of $(-1, .5)$ and $(.5, -1)$ or $(-.25, -.25)$.

Next consider the role of classification. The weights for the most preferred and least preferred disclosures will not change so the value associated with a most preferred (L, H) disclosure remains $\lambda_L x_1 + \lambda_H x_2$. If instead, given a particular realization of (α_1, α_2) , the firm wished to report an aggregate $\sum x_i$ as its most preferred disclosure, then it would be indifferent among all (x_1, x_2) giving rise to $\sum x_i$, because this is all the investor can see. Consequently, the value associated with a most preferred (L, L) disclosure of $\sum x_i$ will be $\lambda_L \frac{\sum x_i}{2} + \lambda_L \frac{\sum x_i}{2} = \lambda_L \sum x_i$. A similar argument can be made for a most preferred (H, H) resulting in an investor value of $\lambda_H \sum x_i$. Using these weights, we graph Figure 8, as the analogue of Figure 5.

-- Place Figure 8 about here --

Figure 8 presents a complementary analogue to Figure 5. Figure 8 exhibits the firm's most preferred classifications for (x_1, x_2) pairs for $\varepsilon = .3$ and $\varepsilon = .5$ given within-category aggregation.

Note how the preference for (L, L) and (H, H) shrinks with ε . In the limit, as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$, the preference for (L, L) and (H, H) vanishes and resulting figure will indicate that most preferred areas remaining only for (L, H) and (H, L) separated by a graph with slope equal to one. That is, with probability one, α_1 and α_2 can be strictly ranked, and with probability one half, the most preferred preference will be realized. On the other hand, with probability one-half, the most preferred classification will not be realized, but then the firm will classify (x_1, x_2) as either (L, L) or (H, H) depending on the sign of $\sum x_i$, but because $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ is assumed independent of $(\tilde{x}_1, \tilde{x}_2)$ the classification will be uninformative. Thus, as $\varepsilon \rightarrow \frac{1}{\sqrt{\pi}}$, while aggregated disclosures are uninformative, the *process of aggregation* is informative relative to the case of no aggregation described above. The net effect however, is for informativeness to survive in the face of wide discretion.

VI. Conclusion

Our model sheds light on the economics of the classification of disclosures opportunistically made by a firm. To the extent that all disclosures must be classified, our model examines a universal feature of modern financial reporting. By noting certain limitations on this discretion (discretion over classifying extreme items, and the inherent ordering of the items) we are able to show how the classifications are informative. Furthermore, due to the strategic nature of classification, the convention of aggregating within-category items (i.e., reporting a simple sum) may actually increase the overall informativeness of classification relative to classification made without the aggregation convention for the classifications that we study.

Appendix

This appendix contains a sequence of arguments deriving the results figures found in the text.

1. Deriving the weights when (L, H) and (H, L) are most preferred classifications.

Consider Figure 9.

--Insert Figure 9 about here--

Quadrant II exhibits a smaller gray quarter circle with radius ε , and a larger quarter circle with radius $\frac{1}{\sqrt{\pi}}$ containing the gray quarter circle. The centroid of smaller gray quarter circle is

$\left(\frac{1}{\sqrt{\pi}} - \frac{4\varepsilon}{3\pi}, \frac{1}{\sqrt{\pi}} + \frac{4\varepsilon}{3\pi}\right)$ and the centroid of the larger quarter circle is $\left(\frac{1}{\sqrt{\pi}} - \frac{4}{3\pi^{3/2}}, \frac{1}{\sqrt{\pi}} + \frac{4}{3\pi^{3/2}}\right)$. The probabilities of these two regions are $\frac{\pi\varepsilon^2}{4}$ and $\frac{1}{4}$ respectively.

Consider next, the region corresponding to the larger quarter circle less the smaller gray quarter circle, or the white fraction of the larger quarter circle. This represents those realizations of α_1 and α_2 for which no discretion is allowed. The centroid of this region (ω_L, ω_H) obeys

$$\left(\frac{1}{\sqrt{\pi}} - \frac{4\varepsilon}{3\pi}\right) \left(\frac{\left(\frac{\pi\varepsilon^2}{4}\right)}{\frac{1}{4}}\right) + \omega_L \left(\frac{\left(\frac{\frac{1}{4} - \pi\varepsilon^2}{4}\right)}{\frac{1}{4}}\right) = \left(\frac{1}{\sqrt{\pi}} - \frac{4}{3\pi^{3/2}}\right), \text{ and}$$

$$\left(\frac{1}{\sqrt{\pi}} + \frac{4\varepsilon}{3\pi}\right) \left(\frac{\left(\frac{\pi\varepsilon^2}{4}\right)}{\frac{1}{4}}\right) + \omega_H \left(\frac{\left(\frac{\frac{1}{4} + \pi\varepsilon^2}{4}\right)}{\frac{1}{4}}\right) = \left(\frac{1}{\sqrt{\pi}} + \frac{4}{3\pi^{3/2}}\right), \text{ with the result that}$$

$$\omega_L = \frac{3\pi^{3/2}\varepsilon + \pi(3 - 4\varepsilon^2) - 4 - 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)} \text{ and } \omega_H = \frac{3\pi^{3/2}\varepsilon + \pi(3 - 4\varepsilon^2) + 4 + 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)}$$

Consider next Figure 10,

--Place Figure 10 about here--

For Figure 10, the darkened half circle represents the area of discretion for the report (L, H) . The darkened half-circle has centroid $\left(\frac{1}{\sqrt{\pi}} - \frac{2\sqrt{2}\varepsilon}{3\pi}, \frac{1}{\sqrt{\pi}} + \frac{2\sqrt{2}\varepsilon}{3\pi}\right)$, and the probability of being in this half circle equals $\frac{\pi\varepsilon^2}{2}$. Consequently the expected values of $\tilde{\alpha}_1$ and $\tilde{\alpha}_1$ given (L, H) equals

$$\begin{aligned}
 & E[\tilde{\alpha}_1 | (L, H)] \\
 &= \left(\frac{1}{\sqrt{\pi}} - \frac{2\sqrt{2}\varepsilon}{3\pi}\right) \left(\frac{\left(\frac{\pi\varepsilon^2}{2}\right)}{\frac{\pi\varepsilon^2}{2} + \frac{1}{4} - \frac{\pi\varepsilon^2}{4}}\right) + \left(\frac{3\pi^{3/2}\varepsilon + \pi(3 - 4\varepsilon^2) - 4 - 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)}\right) \left(\frac{\left(\frac{1}{4} - \frac{\pi\varepsilon^2}{4}\right)}{\frac{\pi\varepsilon^2}{2} + \frac{1}{4} - \frac{\pi\varepsilon^2}{4}}\right) \\
 &= \frac{3\pi^2\varepsilon^2 - 4(\sqrt{2} - 1)\pi^{3/2}\varepsilon^3 + 3\pi - 4}{3\pi^{3/2}(1 + \pi^2\varepsilon)} \\
 &= \mu_L
 \end{aligned}$$

$$\begin{aligned}
 & \text{and } E[\tilde{\alpha}_2 | (L, H)] \\
 &= \left(\frac{1}{\sqrt{\pi}} + \frac{2\sqrt{2}\varepsilon}{3\pi}\right) \left(\frac{\left(\frac{\pi\varepsilon^2}{2}\right)}{\frac{\pi\varepsilon^2}{2} + \frac{1}{4} - \frac{\pi\varepsilon^2}{4}}\right) + \left(\frac{3\pi^{3/2}\varepsilon + \pi(3 - 4\varepsilon^2) + 4 + 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)}\right) \left(\frac{\left(\frac{1}{4} - \frac{\pi\varepsilon^2}{4}\right)}{\frac{\pi\varepsilon^2}{2} + \frac{1}{4} - \frac{\pi\varepsilon^2}{4}}\right) \\
 &= \frac{3\pi^2\varepsilon^2 + 4(\sqrt{2} - 1)\pi^{3/2}\varepsilon^3 + 3\pi + 4}{3\pi^{3/2}(1 + \pi^2\varepsilon)} \\
 &= \mu_H
 \end{aligned}$$

The analysis for the (H, L) disclosure is symmetric.

2. Deriving the weights when (H, H) or (L, L) is most preferred

Consider Figure 11:

--Place Figure 11 about here--

The dotted fill circle in Figure 11, represents the area of discretion which has centroid $\left(\frac{1}{\sqrt{\pi}}, \frac{1}{\sqrt{\pi}}\right)$ and probability $\pi\varepsilon^2$. A an analysis similar to that for (L, H) establishes that

$$E[\tilde{\alpha}_1 | (H, H)] = E[\tilde{\alpha}_2 | (H, H)] = \frac{3\pi^2\varepsilon^2 - 4\pi^{3/2}\varepsilon^3 + 3\pi + 4}{3\pi^{3/2}(1 + \pi\varepsilon^2)} = \lambda_H.$$

A similar argument for (L, L) establishes that

$$E[\tilde{\alpha}_1 | (L, L)] = E[\tilde{\alpha}_2 | (L, L)] = \frac{9\pi^2\varepsilon^2 + 4\pi^{3/2}\varepsilon^3 + 3\pi - 4}{3\pi^{3/2}(1 + 3\pi\varepsilon^2)} = \lambda_L.$$

3. Deriving the boundaries for Figure 5

For each boundary indicated below we will solve for x_2 as a function of x_1 . Instead of directly solving for x_2 , we will solve first for \hat{x}_2 and the resulting boundary be

$$x_2 = \begin{cases} 0 & \text{if } \hat{x}_2 < 0 \\ \hat{x}_2 & \text{if } 0 \leq \hat{x}_2 \leq 1 \\ 1 & \text{if } \hat{x}_2 > 1 \end{cases}$$

Boundary for $(H, H)/(L, H)$: solve $x_1\mu_H + \hat{x}_2\mu_H = x_1\lambda_L + \hat{x}_2\lambda_H$ to obtain

$$\hat{x}_2 = \left(-\frac{(2 + 4\pi\varepsilon^2 + (\sqrt{2} - 2)\pi^{3/2}\varepsilon^3 + (3\sqrt{2} - 4)\pi^{5/2}\varepsilon^5)}{\sqrt{2}\pi^{3/2}\varepsilon^3(1 + 3\pi\varepsilon^2)} \right) x_1$$

Boundary for $(L, L)/(L, H)$: solve $x_1\mu_L + \hat{x}_2\mu_L = x_1\lambda_L + \hat{x}_2\lambda_H$

$$\hat{x}_2 = \left(-\frac{\pi(2\varepsilon^2 + \sqrt{2}\pi\varepsilon^3 - 2\pi^{3/2}\varepsilon^5 + 3\sqrt{2}\pi^{3/2}\varepsilon^5)}{(1 + 3\pi\varepsilon^2)(2 - 2\pi^{3/2}\varepsilon^3 + \sqrt{2}\pi^{3/2}\varepsilon^3)} \right) x_1$$

Boundary for $(L,L)/(H,L)$: solve $x_1\mu_L + \hat{x}_2\mu_L = x_1\lambda_H + \hat{x}_2\lambda_L$

$$\hat{x}_2 = \left(-\frac{(1+3\pi\varepsilon^2)(2-2\pi^{3/2}\varepsilon^3 + \sqrt{2}\pi^{3/2}\varepsilon^3)}{\pi\varepsilon^2(2+\sqrt{2}\pi\varepsilon - 2\pi^{3/2}\varepsilon^3 + 3\sqrt{2}\pi^{3/2}\varepsilon^3)} \right) x_1 \text{ and}$$

Boundary for $(H,H)/(H,L)$: solve $x_1\mu_H + \hat{x}_2\mu_H = x_1\lambda_H + \hat{x}_2\lambda_L$

$$\hat{x}_2 = \left(-\frac{\sqrt{2}\pi^{3/2}\varepsilon^3(1+3\pi\varepsilon^2)}{2+4\pi\varepsilon^2 - 2\pi^{3/2}\varepsilon^3 + \sqrt{2}\pi^{3/2}\varepsilon^3 - 4\pi^{5/2}\varepsilon^5 + 3\sqrt{2}\pi^{5/2}\varepsilon^5} \right) x_1$$

4. Deriving the weights when (L,H) is preferred, but another report is issued.

If (L,H) is preferred but not reported, we know that (α_1, α_2) is not located in the shaded upper semicircle of Figure 3, nor in the section of the white outer ring corresponding to (L,H) . From Figure 5, we may hypothesize that when (x_1, x_2) is near the borders of the (L,H) preference zone, that the second most preferred disclosure will be either (H,H) or (L,L) . If the firm issues the second most preferred disclosure, say, (L,L) , then we know that (α_1, α_2) is either located in the white outer rim for (L,L) or the lower unshaded semicircle of Figure 3. In this case, the centroid for the white segment is $(\omega_L, \omega_L) =$

$$\left(\frac{3\pi^{3/2}\varepsilon + \pi(3-4\varepsilon^2) - 4 - 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)}, \frac{3\pi^{3/2}\varepsilon + \pi(3-4\varepsilon^2) - 4 - 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)} \right)$$

and the centroid for the lower unshaded semicircle is $\left(\frac{1}{\sqrt{\pi}} + \frac{2\sqrt{2}\varepsilon}{3\pi}, \frac{1}{\sqrt{\pi}} - \frac{2\sqrt{2}\varepsilon}{3\pi} \right)$, so the

expected centroid over both regions equals $(\bar{\alpha}_1, \bar{\alpha}_2)$ where

$$\begin{aligned} \bar{\alpha}_1 &= \left(\frac{1}{\sqrt{\pi}} - \frac{2\sqrt{2}\varepsilon}{3\pi} \right) \left(\frac{\left(\frac{\pi\varepsilon^2}{2} \right)}{\frac{\pi\varepsilon^2}{2} + \frac{1}{4} - \frac{\pi\varepsilon^2}{4}} \right) + \left(\frac{3\pi^{3/2}\varepsilon + \pi(3-4\varepsilon^2) - 4 - 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)} \right) \left(\frac{\left(\frac{1}{4} - \frac{\pi\varepsilon^2}{4} \right)}{\frac{\pi\varepsilon^2}{2} + \frac{1}{4} - \frac{\pi\varepsilon^2}{4}} \right) \\ &= \frac{3\pi^2\varepsilon^2 + 4(\sqrt{2}+1)\pi^{3/2}\varepsilon^3 + 3\pi - 4}{3\pi^{3/2}(1+\pi^2\varepsilon)} = \gamma. \end{aligned}$$

$$\begin{aligned}\bar{\alpha}_2 &= \left(\frac{1}{\sqrt{\pi}} + \frac{2\sqrt{2}\varepsilon}{3\pi} \right) \left(\frac{\left(\frac{\pi\varepsilon^2}{2} \right)}{\frac{\pi\varepsilon^2}{2} + \frac{1}{4} - \frac{\pi\varepsilon^2}{4}} \right) + \left(\frac{3\pi^{3/2}\varepsilon + \pi(3-4\varepsilon^2) - 4 - 4\sqrt{\pi}\varepsilon}{3(\pi^{3/2} + \pi^2\varepsilon)} \right) \left(\frac{\left(\frac{1}{4} - \frac{\pi\varepsilon^2}{4} \right)}{\frac{\pi\varepsilon^2}{2} + \frac{1}{4} - \frac{\pi\varepsilon^2}{4}} \right) \\ &= \frac{3\pi^2\varepsilon^2 - 4(\sqrt{2}-1)\pi^{3/2}\varepsilon^3 + 3\pi - 4}{3\pi^{3/2}(1 + \pi^2\varepsilon)} = \mu_L\end{aligned}$$

As similar analysis establishes that a firm with an (L, H) preference who discloses as second choice (H, H) will have weights $(\bar{\alpha}_1, \bar{\alpha}_2)$

$$\bar{\alpha}_1 = \frac{3\pi^2\varepsilon^2 + 4(\sqrt{2}-1)\pi^{3/2}\varepsilon^3 + 3\pi + 4}{3\pi^{3/2}(1 + \pi^2\varepsilon)} = \mu_H$$

$$\bar{\alpha}_2 = \frac{3\pi^2\varepsilon^2 - 4(\sqrt{2}+1)\pi^{3/2}\varepsilon^3 + 3\pi + 4}{3\pi^{3/2}(1 + \pi^2\varepsilon)} = \beta$$

To show that the diagonal $-x_1 = x_2$ forms a boundary between the second choice disclosures for preferred choice (L, H) we compare the value $\mu_L x_1 + \mu_L x_2$ for second choice disclosure of (L, L) to value $\mu_H x_1 + \beta x_2$ of second choice disclosure of (H, H) evaluated at $-x_1 = x_2$ and find that they are equal.

Symmetric analysis apply to the case of second choice disclosures for most preferred (L, H) .

5. Establishing that (H, L) is least preferred when (L, H) most preferred and vice versa.

For example, If (L, H) most preferred then the valuation will be $\mu_L x_1 + \mu_H x_2$. By establishing boundaries, Figure 5 implies that the (L, L) region cannot be most preferred. From Figure 5 we see that $x_1 < 0 < x_2$ and from Figure 4 we see that $\mu_L < \mu_H$ which in implies that

$$\mu_L x_1 + \mu_H x_2 > \mu_H x_1 + \mu_L x_2.$$

Which implies that when (x_1, x_2) is contained in the (L, H) most preferred region of Figure 5 (H, L) cannot be most preferred.

Suppose next, contrary to assumption, that (H, L) is second preferred. If (H, L) were second preferred then the investor would know that (α_1, α_2) lies either in the white outer rim for (H, L) or hatched semicircle of Figure 3. Recall that if (L, L) we second preferred, then the investor would know that (α_1, α_2) lies either in the white outer rim for (L, L) or the hatched semicircle of Figure 3. Thus a second chose of (H, L) or (L, L) share the hatched semi-circle, and we need to evaluate only the effect of the outer white rim. If (H, L) is second preferred then

$$\omega_H x_1 + \omega_L x_2 > \omega_L x_1 + \omega_L x_2, \text{ or}$$

$$(\omega_H - \omega_L)x_1 > 0 .$$

This is contradicted however, by the fact that $x_1 < 0 < x_2$, demonstrating that (H, L) is least preferred when (L, H) . A symmetric argument establishes that (L, H) is least preferred when (H, L) is most preferred.

6. Boundaries for figure 8:

First, argument similar to argument 5 used above establishes that (H, L) is least preferred when (L, H) most preferred and vice versa.

Next, for each boundary indicated below we will solve for x_2 as a function of x_1 . Instead of directly solving for x_2 , we will solve first for \hat{x}_2 and the resulting boundary be

$$x_2 = \begin{cases} 0 & \text{if } \hat{x}_2 < 0 \\ \hat{x}_2 & \text{if } 0 \leq \hat{x}_2 \leq 1 \\ 1 & \text{if } \hat{x}_2 > 1 \end{cases}$$

Boundary for $(H, H)/(L, H)$: $\lambda_H(x_1 + \hat{x}_2) = \lambda_L x_1 + \lambda_H \hat{x}_2$

$$\hat{x}_2 = \left(\frac{2 - 2\pi^{3/2}\varepsilon^3 + \sqrt{2}\pi^{3/2}\varepsilon^3}{\sqrt{2}\pi^{3/2}\varepsilon^3} \right) x_1$$

Boundary for $(L, L)/(L, H)$: $\lambda_L(x_1 + \hat{x}_2) = \lambda_L x_1 + \lambda_H \hat{x}_2$

$$\hat{x}_2 = \left(\frac{\pi(2\varepsilon^2 + \sqrt{2}\pi\varepsilon^3 - 2\pi^{3/2}\varepsilon^5 + 3\sqrt{2}\pi^{3/2}\varepsilon^5)}{2 + 4\pi\varepsilon^2 - 2\pi^{3/2}\varepsilon^3 + \sqrt{2}\pi^{3/2}\varepsilon^3 - 4\pi^{5/2}\varepsilon^5 + 3\sqrt{2}\pi^{5/2}\varepsilon^5} \right) x_1$$

Boundary for $(L,L)/(H,L)$: $\lambda_L(x_1 + \hat{x}_2) = \lambda_H x_1 + \lambda_L \hat{x}_2$

$$\hat{x}_2 = \left(\frac{2 + 4\pi\varepsilon^2 - 2\pi^{3/2}\varepsilon^3 + \sqrt{2}\pi^{3/2}\varepsilon^3 - 4\pi^{5/2}\varepsilon^5 + 3\sqrt{2}\pi^{5/2}\varepsilon^5}{\pi\varepsilon^2(2 + \sqrt{2}\pi\varepsilon - 2\pi^{3/2}\varepsilon^3 + 3\sqrt{2}\pi^{3/2}\varepsilon^3)} \right) x_1 \text{ and}$$

Boundary for $(H,H)/(H,L)$: $\lambda_H(x_1 + \hat{x}_2) = \lambda_H x_1 + \lambda_L \hat{x}_2$

$$\hat{x}_2 = \left(\frac{\sqrt{2}\pi^{3/2}\varepsilon^3}{2 - 2\pi^{3/2}\varepsilon^3 + \sqrt{2}\pi^{3/2}\varepsilon^3} \right) x_1$$

References

Barberis, Nicholas and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics*, Vol. 68(2), pp. 161-199.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A Model of Investor Sentiment, *Journal of Financial Economics*, Vol. 49 pp. 307-343.

Dichev, Ilija, John Graham, Campbell Harvey, and Shiva Rajgopal, 2013, Earnings Quality: Evidence from the Field, *Journal of Accounting and Economics*, pp. 1-33.

Dye, Ronald, and Sri Sridhar, 2004, Reliability-Relevance Tradeoffs and the Efficiency of Aggregation, *Journal of Accounting Research*, Vol. 42, pp. 51-88.

Fan, Yun, Abhijit Barua, William Cready, and Wayne Thomas, 2010, Managing Earnings Using Classification Shifting: Evidence from Quarterly Special items, *The Accounting Review*, Vol. 85, pp. 1303-1323.

Hirst, D. Eric., Lisa Koonce, and Paul Simko, 1995, Investor reactions to financial analysts' research reports. *Journal of Accounting Research*, Vol. 33, pp. 335-51.

Ijiri, Yuji, 1967, *The Foundations of Accounting Measurement*, Englewood Cliffs, Prentice Hall.

McVay, Sarah, 2006, Earnings Management Using Classification Shifting: An Examination of Core Earnings and Special Items, *The Accounting Review*, Vol. 81, pp. 501-531.

Milgrom, Paul and John Roberts, 1986, Relying on the Information of Interested Parties, *RAND Journal of Economics*, Vol. 17, pp.18-32.

Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer, 2008, Coarse Thinking and Persuasion, *The Quarterly Journal of Economics*, pp. 577- 619.

Mullainathan, Sendhil and Andrei Shleifer, The Market for News, *The American Economic Review*, Vol. 95, pp. 1031-1053.

Schipper, Katherine, 2003, Principles-based Accounting Standards, *Accounting Horizons*, Vol. 17, pp. 61-72.

Sorter, George, 1969, An "Events" Approach to Basic Accounting Theory, *The Accounting Review*, Vol. 44, pp. 12-19.

Sunder, Shyam, 1983, Simpson's Reversal Paradox and Cost Allocation, *Journal of Accounting Research*, Vol. 21, pp. 222-233.

Conditional Expected Value of $\tilde{\alpha}$		
	L	H
$x < 0$	$\frac{1}{2} \left(\frac{1}{\sqrt{\pi}} + \varepsilon \right)$	$\frac{1}{2} \left(\frac{3}{\sqrt{\pi}} + \varepsilon \right)$
$x > 0$	$\frac{1}{2} \left(\frac{1}{\sqrt{\pi}} - \varepsilon \right)$	$\frac{1}{2} \left(\frac{3}{\sqrt{\pi}} - \varepsilon \right)$

Table 1: The investor's valuation weights

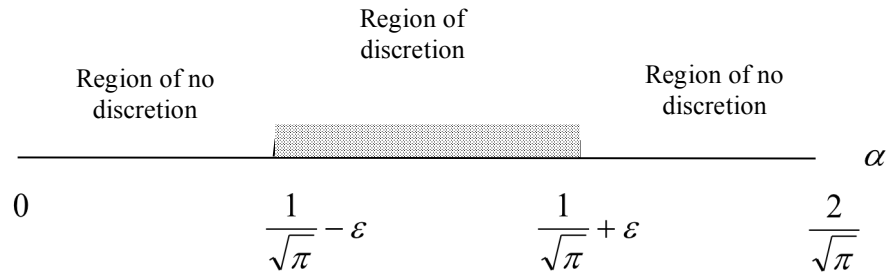


Figure 1: Discretion with a Single Item

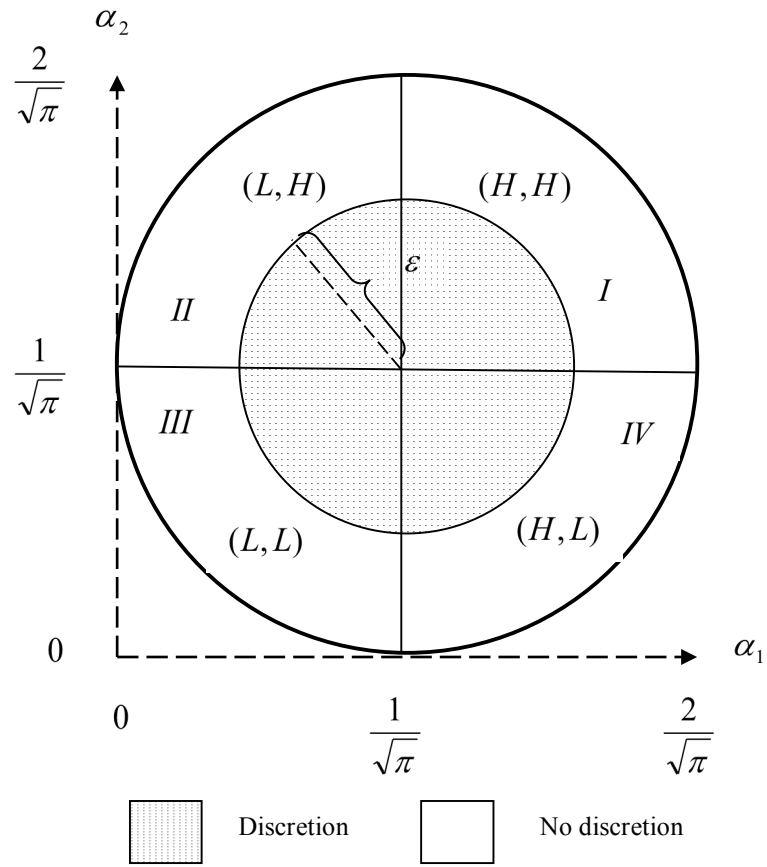


Figure 2: Discretion with Two Items. Roman numerals indicate quadrants.

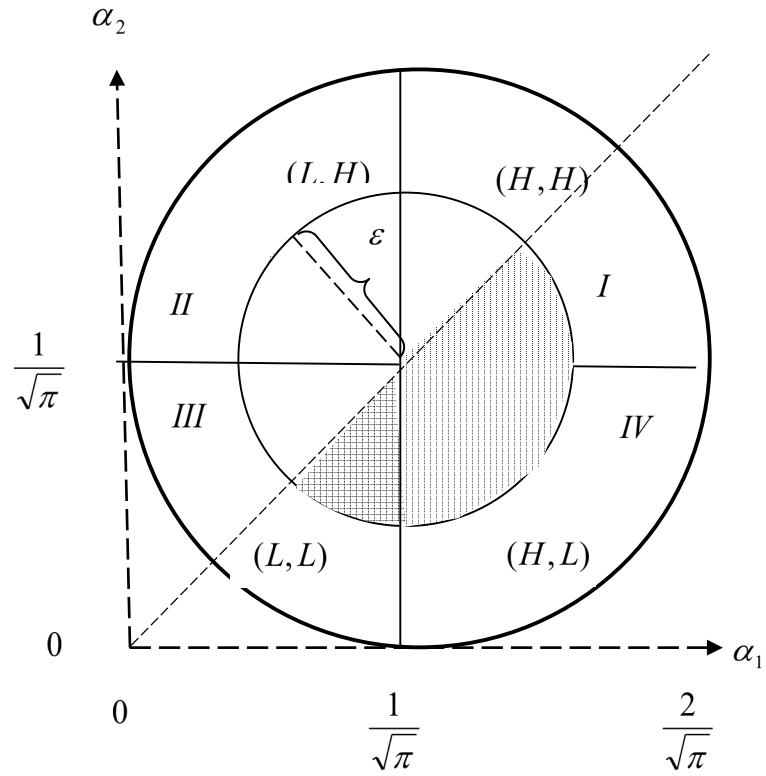


Figure 3: The cross-hatched region represents those realizations of $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ for which the firm may not classify an (x_1, x_2) as (L, H) . The Roman numerals indicate quadrants.

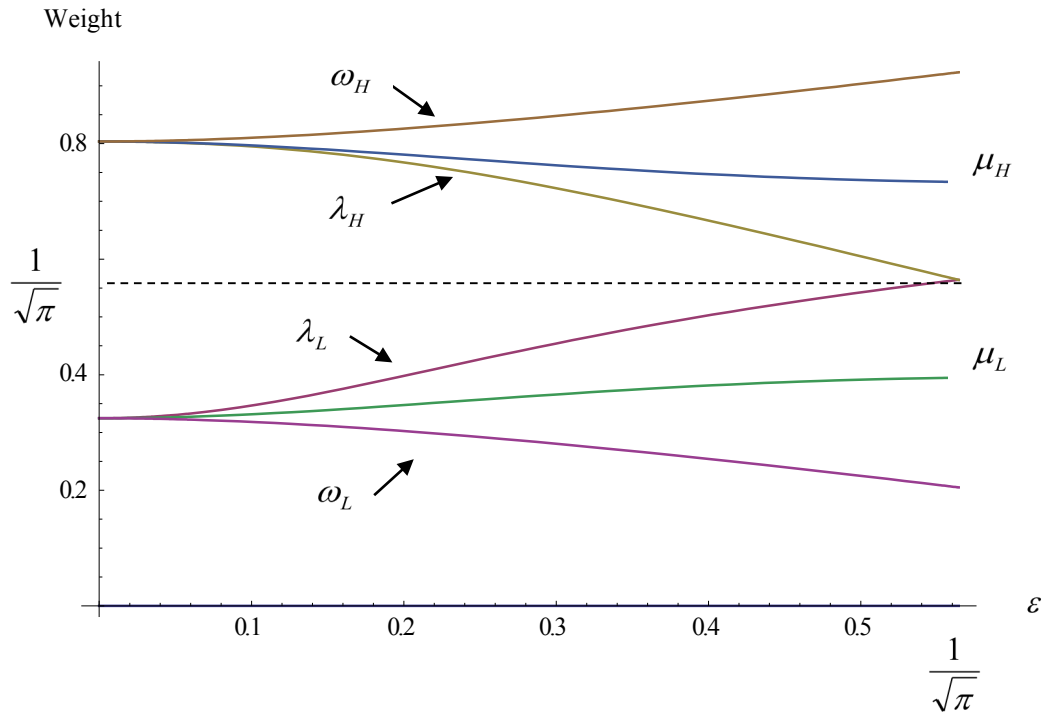


Figure 4: The weights for most preferred and least preferred disclosures

Legend for Figure 4:

ω_i is the weight for any least preferred C_i disclosure.

μ_L and μ_H are the weights for most preferred disclosures (L, H) and (H, L) .

λ_L and λ_H are the weights for most preferred disclosures (L, L) and (H, H) .

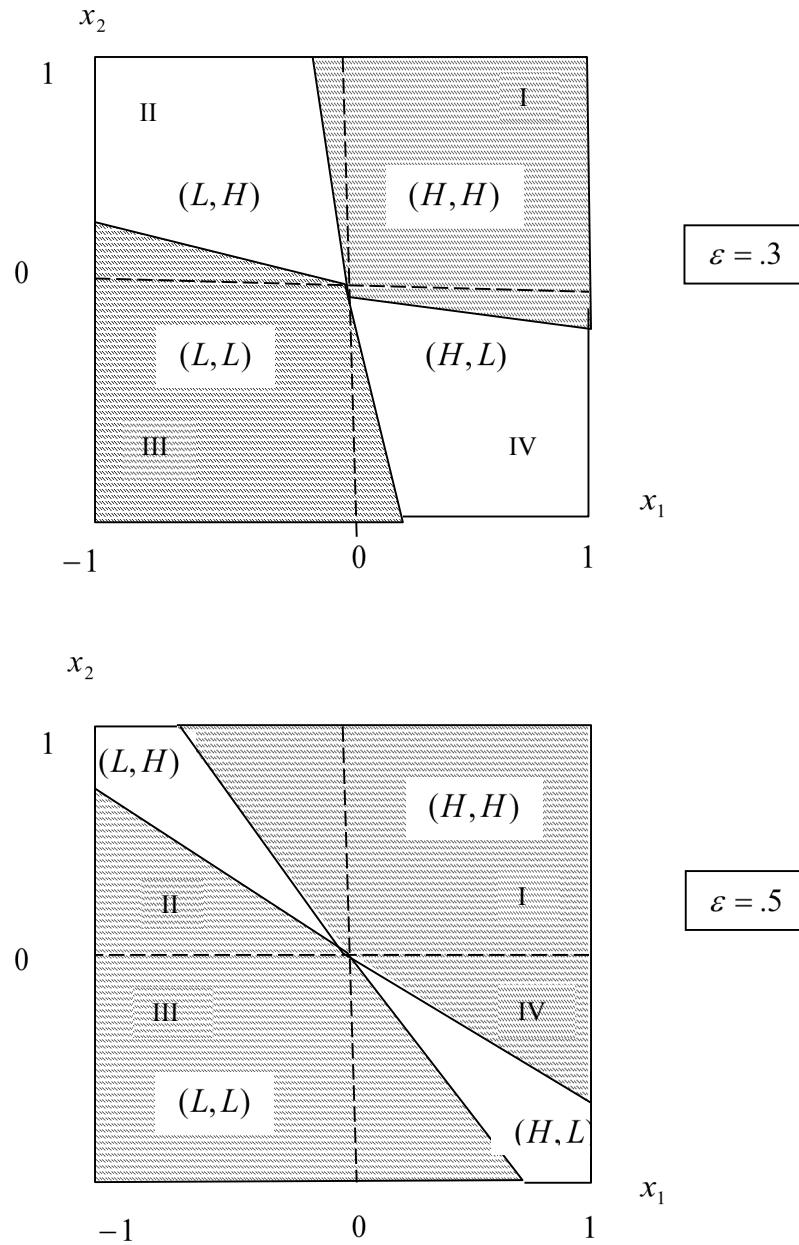


Figure 5: The firm's most preferred classifications for (x_1, x_2) pairs given $\varepsilon = .3$ and $\varepsilon = .5$. Roman numerals indicate quadrants. See the appendix for the boundary expressions.

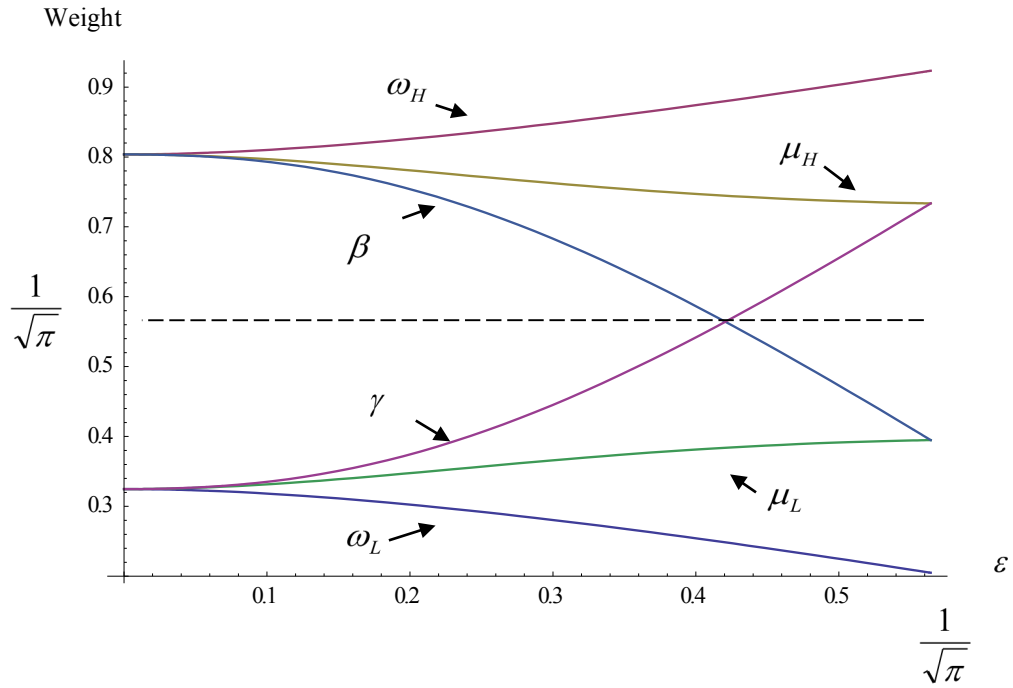


Figure 6: β and γ weights for second choice disclosures

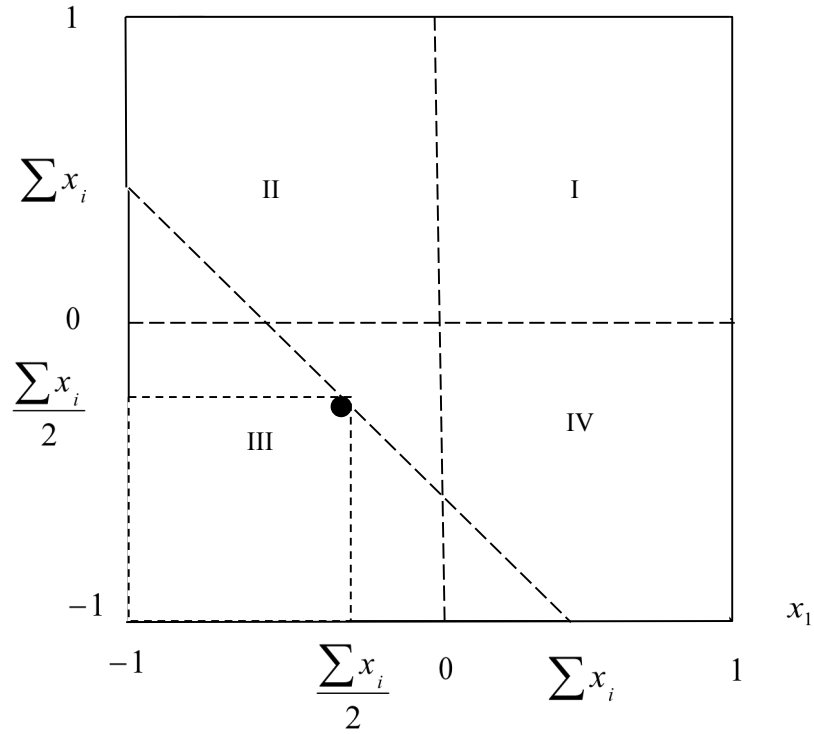


Figure 7: Implication of an aggregate disclosure assuming that $(\tilde{x}_1, \tilde{x}_2)$ is uniformly distributed over a square of length 2 and mean $(0,0)$, The conditional expected value of an $\sum x_i$ disclosure is $\left(\frac{\sum x_i}{2}, \frac{\sum x_i}{2}\right)$.

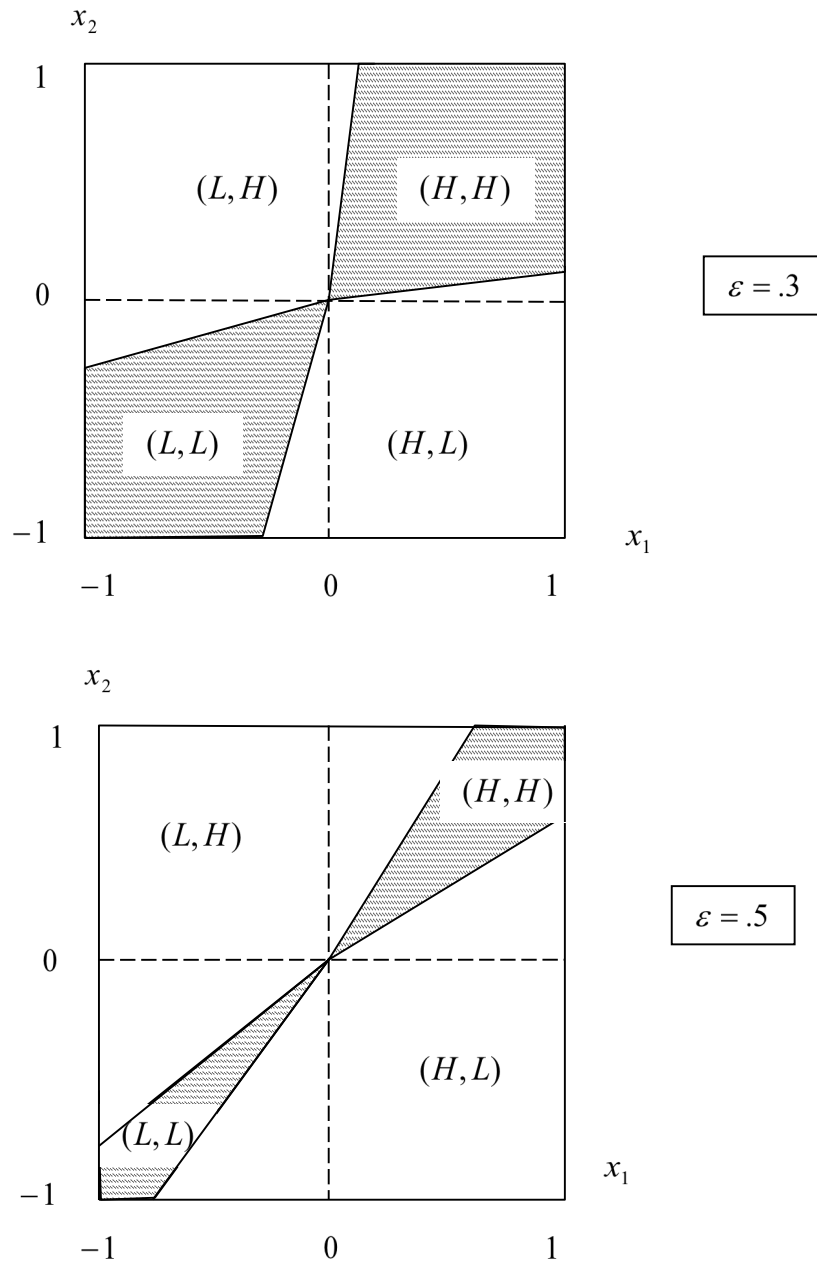


Figure 8: The firm's most preferred classifications for (x_1, x_2) pairs given $\varepsilon = .3$ and $\varepsilon = .5$ with aggregation of (H, H) and (L, L) items. See the appendix for the boundary expressions.

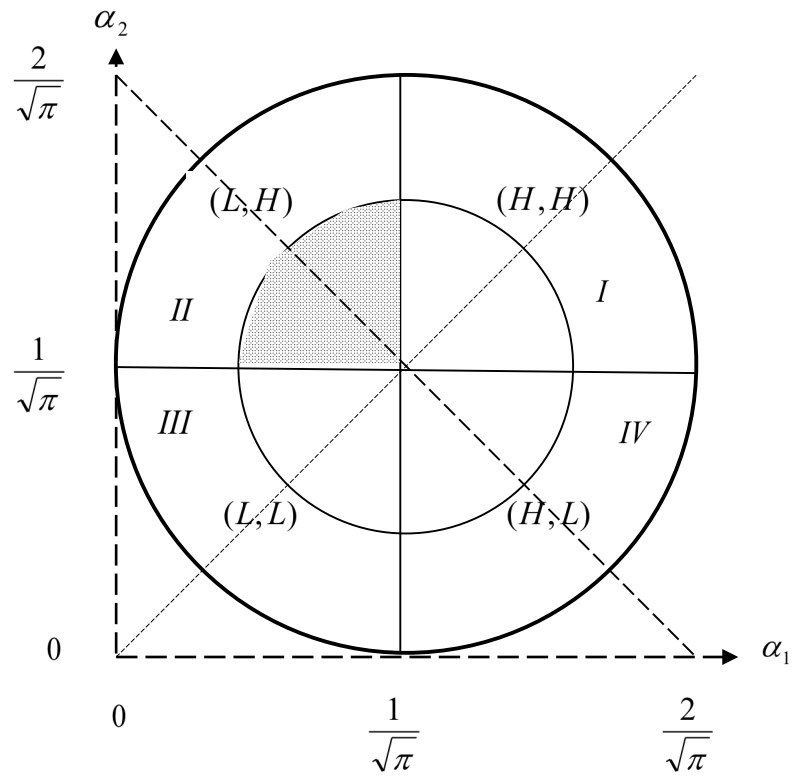


Figure 9: Appendix.

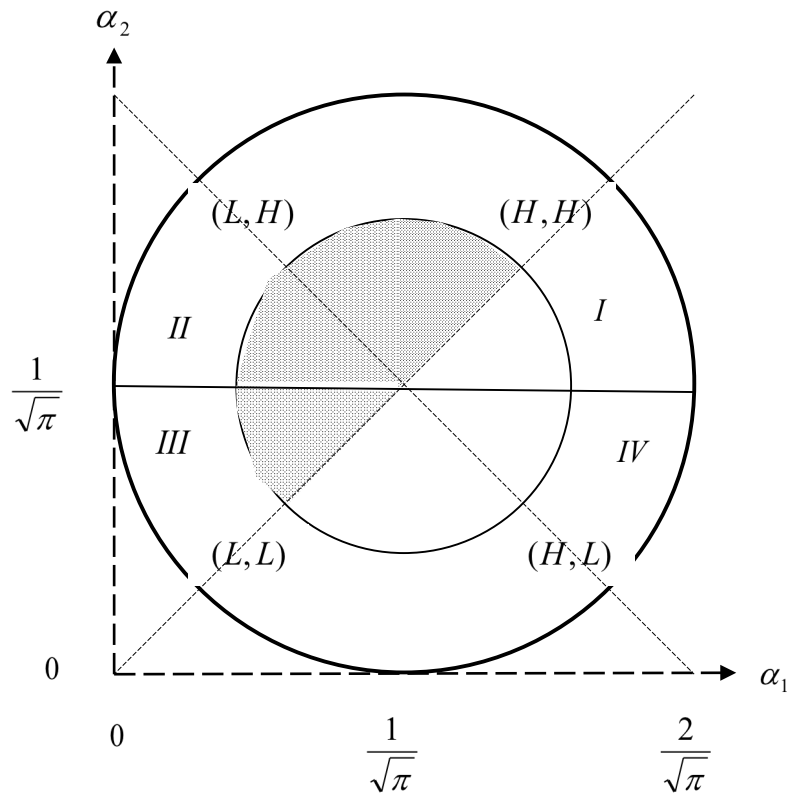


Figure 10: Appendix.

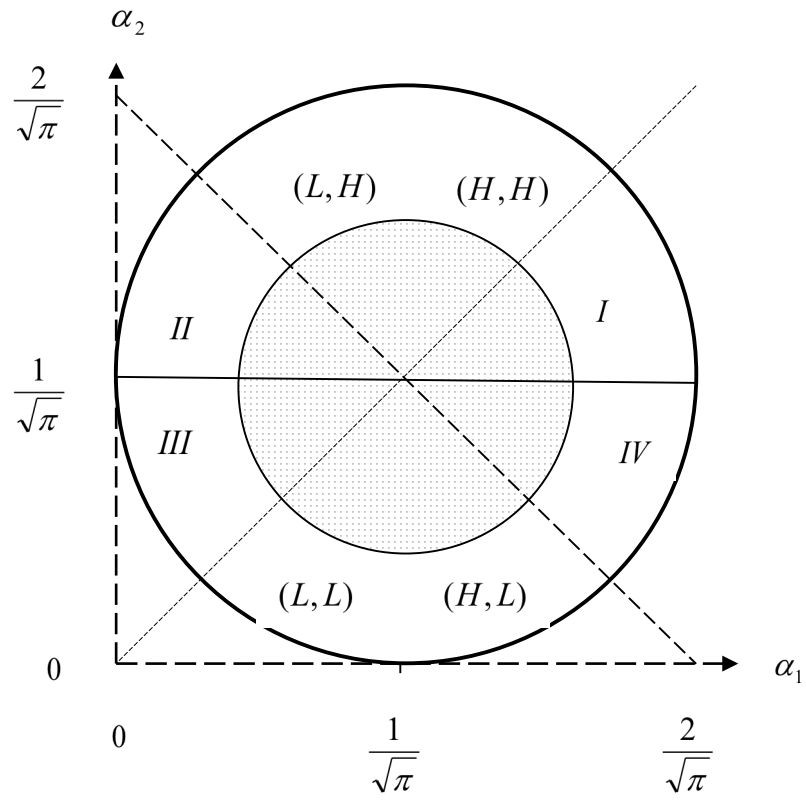


Figure 11: Appendix.