Discretionary Aggregation

Michael Ebert
Mannheim Business School
University of Mannheim[*]

Dirk Simons
Mannheim Business School
University of Mannheim[†]

Jack Stecher
Tepper School of Business
Carnegie Mellon University[‡]

August 25, 2014

[*] ebertm@bwl.uni-mannheim.de (*corresponding author*)
[†] simons@bwl.uni-mannheim.de
[‡] jstecher@andrew.cmu.edu

Discretionary Aggregation

Abstract

Consider the following problem: a firm's manager has many sources of private information about the firm's value, and wishes to maximize the firm's share price on a market characterized by rational expectations. Everyone knows that the manager is informed, but the manager alone knows how many signals he has received. If the manager discloses anything, it must contain the truthful net impact of all his private sources of information. But he can reveal more, by partitioning the set of private signals and disclosing subtotals. What will the manager disclose? We show that he will disclose (a sufficient statistic for) all his information if and only if the news is sufficiently bad. Otherwise, he will disclose only the net result of his signals. That is, discretionary aggregation creates incentive to provide detailed information in bad states, and to provide only summary information in good states.

Keywords: Aggregation; Disclosure; Reporting Discretion; Sender-Receiver Games; Strategic Communication; Persuasion; Conservatism

I Introduction

The purpose of this paper is to study the consequences of discretionary aggregation in financial reporting. We address the following question: do firms' managers, whose sole goal is to maximize their company's share price, use discretion in aggregation to hide information from the market? The answer, as we demonstrate below, is only if the news is sufficiently good, i.e., only if the news makes the firm's market value sufficiently high. Otherwise, the manager optimally discloses everything.

We illustrate this point using an analytical model, in which we study a voluntary disclosure game between a risk neutral manager, who wants to maximize his firm's share price, and a risk neutral market, which prices the firm's share using rational expectations. The manager receives a collection of signals that are informative about the value of the firm's sole project. Only he knows how many signals he has received; ex ante, he can receive any positive finite number of signals.

He can choose to disclose, at zero cost, a report that aggregates his signals. Additionally, the manager can choose the granularity of the aggregation. He does so by partitioning the set of signals any way he likes, then sending a costless message giving the subtotal of the signals on each partition set. Our interest is in how much detail the manager optimally provides, and, if the manager chooses a nontrivial partition, how he chooses to group the signals.

By construction, the only way we enable a manager to hide information is through aggregation. The manager cannot lie or disclose half-truths. He cannot double-count, alter, or omit any signals. In this sense, if the manager provides any information, then he must tell the truth, the whole truth, and nothing but the truth. The whole truth is, however, less complete than it may sound. It must contain the net effect of his information, but still may not be fully disaggregate.

To illustrate, imagine the manager receives three signals from support $\{-1, 1\}$. Suppose two realized signals equal 1 and one equals -1. If the manager reports anything, he must reveal that his signals sum to 1; this precludes him from saying, “I received two signals, each equal to 1,” without mentioning the negative signal. If he aggregates all three signals into one report, he declares, “The signals sum to 1,” but the market cannot infer that the manager has received three signals. For all the market knows, the manager may have received a single signal, equal to 1. If instead the manager reports one subtotal equal to 2 and another equal to -1, the market still infers that the sum of the signals is 1, but it can also learn that at least three signals arrived. By providing subtotals, the manager raises the floor on the number of signals that could have arrived, without reducing the ceiling. This is similar in spirit to Shin (1994), though the driving force is different.

We find that the manager always discloses something, i.e., that silence is never optimal. The reasons are fairly standard, similar to those in Grossman and Hart (1980), Grossman (1981), Milgrom (1981), and Milgrom and Roberts (1986), although the fact that the market cannot bound the number of signals he receives makes the argument require some technical subtleties (see Dye and Finn, 2007). But the unraveling does not extend beyond the aggregate signal.

Whether the manager discloses more than the aggregate total hinges on the value of the firm, given the manager's private information. It is the manager's posterior distribution of the firm value that matters, and not the market's prior beliefs. To understand why, return to the previous example, where the manager has received three signals: $(1, 1, -1)$ from support $\{-1, 1\}$, but only the manager knows he has received three signals. Suppose the manager's posterior probability that the firm's sole project succeeds is above $1/2$. If, instead of reporting $(1, 1, -1)$, the manager reports (1), then he conceals two offsetting signals. This pushes the posterior probability of the firm's success away from $1/2$, thereby increasing firm's expected value. On the other hand, if the posterior probability of the firm's success is below $1/2$, then the manager prefers to reveal as many offsetting signals as possible. Doing so makes the posterior probability of success as close to $1/2$ as possible, thereby increasing the firm's expected value.

The issue our game addresses, aggregation, is of primary importance in accounting. A firm's manager has tremendous latitude on which line items to include in an annual report, and which to aggregate into a summary account. Investors reviewing a balance sheet are in no position to assess how many transactions are collapsed into a reported line item, but instead must form conjectures (see the discussion in Arya et al., 2000).

One way that firms can provide disaggregate information is through segment reporting. Under US GAAP (ASC 280) and under International Financial Reporting Standards (IFRS 8), firms providing segment-level reports must base their subtotals on information used in their internal management structure. A firm discloses the items that it does not allocate, and then assigns the remaining items to the reporting segments. Segment disclosures can be quite coarse. For example, as of 2013, Amazon operates in many jurisdictions, but reports two segments: North American and International operations.[1]

The reason that there is room for nondisclosure of the details is that the market, though aware of the net results, cannot directly observe how many signals the manager received. A reader of Amazon's 2013 annual report knows that Amazon uses at least two segments for internal reporting. Otherwise, the number of internal segments is completely opaque to the reader. As in Dye (1985) and Jung and Kwon (1988), the manager's ability to plead ignorance creates room for the manager not to disclose a sufficient statistic for the amount of information used to construct the aggregate report.[2]

The emphasis on a sufficient statistic is crucial. The case of Amazon is typical: few firms report more than three segments, reflecting only a modest increase in the granularity of reporting—on average, roughly half a segment—since the adoption of the current standards in 1997 in the US and in 2006 under IFRS (see Nichols and Street, 1999, Herrmann and Thomas, 2000, Street et al., 2000, Berger and Hann, 2003, Ettredge et al., 2006). What matters is not the number of segments, but the informativeness of the segments, and there is evidence that segment disclosures, under the current standards, reveal more information than under the previous standards of segmentation simply based on geography or industry;

see Behn et al. (2002) and Ettredge et al. (2005).

The driving force of our results is that aggregation hides offsetting signals, obscuring neutral news. A firm's manager benefits from revealing neutral news if and only if, on balance, the information about a firm is bad. Revealing good news in this case raises the market's assessed probability of good news. We spell out the details of this argument in the remaining sections, with proofs in the appendix.

II The Model

Set up

There are two players, a firm's risk neutral manager, whose objective is to maximize the firm's share price, and a risk neutral market, which prices the firm using rational expectations. Intuitively, the market can be thought of as minimizing the squared difference between the firm's value and the share price, so that investors price the shares at their mean value conditional on all public information. The market is purely a secondary market, not used to generate new financing for the firm. This assumption allows us to focus on the share price without delving into corporate finance issues.

The firm's value is determined by a single Bernoulli draw with unknown success probability $\widetilde{\pi}$. As we describe below, the manager privately receives an unknown and unbounded number of signals about $\widetilde{\pi}$, though it is common knowledge that he will receive at least one signal. Upon receiving the signals, the manager may issue his disclosure to the market, which values the firm at $E[\widetilde{\pi}|$ disclosure]. See Figure 1.

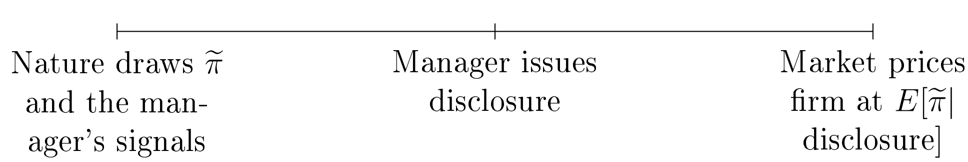

Figure 1: Sequence of events.

Information

The success probability $\widetilde{\pi}$ has a commonly known Beta(α, β) prior, which is the standard conjugate prior for the Bernoulli distribution. If $\alpha = \beta = 1$, the prior of $\widetilde{\pi}$ is uniform. Given α, β, the expectation of $\widetilde{\pi}$ is $E[\widetilde{\pi}] = \alpha/(\alpha + \beta)$. Upon observing σ successes and ϕ failures, the posterior given a Beta(α, β) prior is Beta$(\alpha + \sigma, \beta + \phi)$.

For readers unfamiliar with the Beta distribution, we review some of its properties. If we restrict α and β to positive integers (harmless for our purposes), the Beta pdf is

$$
\begin{aligned}
f_{\Pi}(\pi) &= \frac{1}{B(\alpha,\beta)}\pi^{\alpha-1}(1-\pi)^{\beta-1} = \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!}\pi^{\alpha-1}(1-\pi)^{\beta-1} \\
&= (\alpha+\beta-1)\binom{\alpha+\beta-2}{\alpha-1}\pi^{\alpha-1}(1-\pi)^{\beta-1}
\end{aligned}
$$

Here $B(\alpha, \beta)$ is Euler's Beta function, which for positive integers α, β is equal to $(\alpha-1)!(\beta-1)!)/((\alpha+\beta-1)!)$. Other than the first term, the Beta distribution is exactly the likelihood function of $\widetilde{\pi}$, given that we have observed $\alpha - 1$ successes and $\beta - 1$ failures. Because $\widetilde{\pi}$ is a Bernoulli success probability, the Beta prior assumption is equivalent to assuming that prior information came from observing the outcomes of $(\alpha + \beta - 2)$ prior Bernoulli trials. The normalizing coefficient, $(\alpha + \beta - 1)$, makes the density integrate to 1. This coefficient is the number of possible outcomes that could have been observed: given $(\alpha+\beta-2)$ prior Bernoulli trials, the number of possible successes could have been any number in $\{0, \ldots, \alpha + \beta - 2\}$, giving $\alpha + \beta - 1$ possibilities.

No one ever observes the realized success probability π. However, the manager receives $\widetilde{N}$ signals that are informative about $\widetilde{\pi}$. Both the signals and $\widetilde{N}$ are the manager's private information, though it is common knowledge that $\widetilde{N}$ comes from a Poisson(λ) prior (conditional on at least one signal arriving):

$$
(\forall N \in \mathbb{Z}_{++})\ Pr(\widetilde{N} = N) = \frac{\lambda^N}{N!(e^\lambda - 1)}
$$

The manager's signals, denoted $\widetilde{y} := (\widetilde{y}_1, \ldots, \widetilde{y}_N)$, are independent signed Bernoulli draws with success probability π (similar to the coding of information in Arya et al., 2000): after Nature chooses the realizations (π, N), she chooses N independent Bernoulli draws, coding failures as -1, and gives each draw to the manager:

$$\text{Given } (N, \pi), \quad (\forall i \in \{1, \ldots, N\})$$

$$\widetilde{y}_i \overset{\text{iid}}{\sim} 2 \cdot \text{Bernoulli}(\pi) - 1 = \begin{cases} 1, & \text{with probability } \pi \\ -1, & \text{with probability } 1 - \pi \end{cases}$$

Disclosure

After observing $(y_1, \ldots, y_N)$, the manger chooses whether to disclose, and if so, in how much detail. If the manager discloses, his report must convey the *net* number of positive signals. We denote a fully aggregate report by $m^1(y)$:

$$m^1(y) := \sum_{i=1}^{N} y_i$$

We have an initial observation on the fully aggregate report:

Proposition 1. *If the market receives report $m^1(y)$, along with enough information to infer N correctly, then the market effectively has full information. That is, $\langle N; m^1(y) \rangle$ jointly convey all of the manager's economically relevant information.*

If the manager chooses to disclose more than the fully aggregate report, he is required to do so as follows: first, after receiving N signals, the manager partitions $\{1, \ldots, N\}$ into v

nonempty, mutually disjoint, exhaustive subsets, which we denote $S_1, \ldots, S_v$:

$$(\forall j \in \{1, \ldots, v\}) \quad S_j \neq \varnothing$$

$$(\forall i, j \in \{1, \ldots, v\}) \; i \neq j \Rightarrow S_i \bigcap S_j = \varnothing$$

$$\bigcup_{j=1}^{v} S_j = \{1, \ldots, N\}$$

Next, given the manager's private partition, he discloses the net successes on each partition set. The manager's report, denoted $m^v(y)$, is then a v-tuple:

$$m^v(y) := \left(\sum_{j \in S_i} y_j \right)_{i=1}^{v}$$

For example, suppose that $N = 5$ and $y = (1, 1, 1, -1, -1)$, indicating that the manager has received five signals, with three successes and two failures. If the manager issues a fully aggregate report, then $m^1(y) = 1$, indicating that the firm has had one net positive signal. A fully disaggregate report, $m^5(y) = y$, reports each signal separately.

By disclosing $m^5(y)$, the manager shows the market that $N \geq 5$. The manager can convey the same information by splitting the signals as $(S_1 = \{1, 2, 3\}, S_2 = \{4, 5\})$, then reporting $m^2(y) = (3, -2)$. An alternative partitioning into two subtotals would convey less information: if the manager were to split the signals as $(S_1 = \{1\}, S_2 = \{2, 3, 4, 5\})$, then the report $m^2(y) = (1, 0)$ would indicate that $N \geq 3$, because of the requirement that $S_2 \neq \varnothing$. As a last example, if the manager were to choose $(S_1 = \{1\}, S_2 = \{2\}, S_3 = \{3, 4, 5\})$, then the report $m^3(y) = (1, 1, -1)$ would indicate that $N \geq 3$ and would not reveal that any subtotal contained more than one signal.

III Results

Background and Mathematical Results

For this subsection only, assume the manager always issues the aggregate report $m^1(y)$. Below, we investigate the market's inferences given that the manager is strategic in his aggregation and disclosure decisions, in contrast to the analysis of the market's inferences here, when the manager acts purely mechanically.

By Proposition 1, given N and $m^1(y)$, the number of positive signals is $(N + m^1(y))/2$, and the number of negative signals is $(N - m^1(y))/2$. Also, because the manager always receives at least one signal, and because offsetting signals come in pairs, $N \geq \max(|m^1(y|), 1)$ and, for some $h \in \mathbb{Z}_+$, $N = |m^1(y)| + 2h$. This also means that it is always the case that $N + |m^1(y)|$ is even: given the information system, an odd (even) net number of successes can only derive from an odd (even) number of signals.

We have the following result:

Lemma 1. *Let $\alpha = \beta = 1$, so that $\widetilde{\pi} \sim U[0,1]$. Then, given a sample of N independent signed Bernoulli trials and given $m \in \{-N, -N+2, \ldots, N\}$, the probability of observing m successes is 1/(N+1).*

More generally, if $\widetilde{\pi} \sim Beta(\alpha, \beta)$, then given a sample of N independent signed Bernoulli trials with success probability $\widetilde{\pi}$, the probability of m net successes is a scaling constant times the hypergeometric distribution:

$$Pr\left(m \text{ net successes } \middle| N \text{ trials}, N + |m| \text{ even}\right) = \left[\frac{\alpha + \beta - 1}{\alpha + \beta + N - 1}\right] \frac{\binom{\alpha+\beta-2}{\alpha-1}\binom{N}{\frac{N+m}{2}}}{\binom{\alpha+\beta+N-2}{\alpha+\frac{N+m}{2}-1}}$$

$$= \left[\frac{\text{Prior \#(total successes possibilities)}}{\text{Posterior \#(total successes possibilities)}}\right] \cdot \text{Hypergeometric probability mass}$$

We use Lemma 1 to derive the market's posterior distribution about N given the report $m^1(y)$. The next step in this direction is finding the joint probability of $m^1(\widetilde{y})$ and $\widetilde{N}$. We suppress the conditions that $N \geq \max(1, |m^1(y)|)$ and that $N + |m^1(y)|$ is even from here onward unless confusion can arise.

Lemma 2. *Given that the number of independent signals* $\widetilde{N} \sim Poisson(\lambda)|\widetilde{N} \geq 1$ *and that the success probability* $\widetilde{\pi} \sim Beta(\alpha, \beta)$*, the joint probability of receiving* N *signals with* m *net successes is*

$$Pr(m \wedge N) = \left[\frac{(\alpha+\beta-1)\binom{\alpha+\beta-2}{\alpha-1}}{(e^{\lambda}-1)}\right] \cdot \left[\frac{\lambda^N}{(\alpha+\beta+N-1)\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!}\right] \Bigg/ \binom{\alpha+\beta+N-2}{\alpha+\frac{N+m}{2}-1}$$

In the special case of a uniform prior (i.e., $\alpha = \beta = 1$*), this becomes*

$$Pr(m \wedge N) = \left(\frac{1}{e^{\lambda}-1}\right) \cdot \left[\frac{\lambda^N}{(N+1)!}\right]$$

That is, the joint distribution of (N, m) *has a closed-form solution, which decomposes into three parts: a scaling factor that is independent of* (N, m)*,* λ^N*, coming from the Poisson prior, and a combinatorial expression in* m *and* N*, coming from the hypergeometric distribution in Lemma 1.*

The marginal distribution of the net number of successes is directly obtained by summing $Pr(m|N)Pr(N)$ over feasible values of N. This allows us to derive the posterior distribution of N given the net successes m:

Proposition 2. *Given that the number of net successes is* m*, the posterior distribution of the number of draws has the following closed-form solution:*

$$Pr(N|m) = \left[\frac{\lambda^N\left(\alpha+\frac{N+m}{2}-1\right)!\left(\beta+\frac{N-m}{2}-1\right)!}{\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!(\alpha+\beta+N-1)!}\right] \Bigg/ \sum_{\substack{k \geq \max(|m|,1) \\ k+|m| \text{ even}}} \left[\frac{\lambda^k\left(\alpha+\frac{k+m}{2}-1\right)!\left(\beta+\frac{k-m}{2}-1\right)!}{\left(\frac{k+m}{2}\right)!\left(\frac{k-m}{2}\right)!(\alpha+\beta+k-1)!}\right] \quad (1)$$

In the special case of a uniform prior, this becomes

$$Pr(N|m, \alpha = \beta = 1) = \left[\frac{\lambda^N}{(N+1)!}\right] / \left[\sum_{\substack{k \geq \max(|m|,1) \\ k+|m| \ \mathit{even}}} \frac{\lambda^k}{(k+1)!}\right] \tag{2}$$

The numerator consists of λ^N, coming from the Poisson prior, times a ratio of combinatorial expressions: the number of ways to order the posterior number of successes and the posterior number of failures, divided by the number of ways to order the success and failures from the sample and the number of ways to order the total number of draws (plus a degrees of freedom correction). The denominator is the analogous expression, summed over all possible numbers of draws given the net successes (requiring $N \geq 1$).

In the special case of a uniform prior, Equation (2) has an especially simple form and good asymptotic properties. If we multiply the numerator and denominator by λ, we can rewrite Equation (2) with the denominator as a power series:

$$Pr(N|m; \alpha = \beta = 1) = \left[\frac{\lambda^{N+1}}{(N+1)!}\right] / \left[\sum_{\substack{k \geq \max(|m|,1) \\ k+|m| \ \text{even}}} \frac{\lambda^{k+1}}{(k+1)!}\right] \tag{3}$$

The denominator of (3) is

$$\frac{\lambda^{|m|+1}}{(|m|+1)!} + \frac{\lambda^{|m|+3}}{(|m|+3)!} + \dots$$

If the number m of net successes is odd, this sum is the tail of the series

$$\sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!} = \cosh \lambda$$

If instead m is even, the sum in the denominator is the tail of the series

$$\sum_{k=0}^{\infty} \frac{\lambda^{2k+1}}{(2k+1)!} = \sinh \lambda$$

Note that $\cosh \lambda = (e^{\lambda} + e^{-\lambda})/2$ and $\sinh \lambda = (e^{\lambda} - e^{-\lambda})/2$. As λ increases, the value of $e^{-\lambda}$ becomes small compared with e^{λ}, and both $\cosh \lambda$ and $\sinh \lambda \approx e^{\lambda}/2$. For example, if $\lambda = 4$, then $e^{\lambda} \approx 54.60$ and $e^{-\lambda} < 0.02$, making the ratio $(\cosh 4)/(\sinh 4) \approx 1.0007$. We therefore have the following:

Proposition 3. *Let $\widetilde{N} \sim Poisson(\lambda)|N \geq 1$. Let m be the net number of successes on $\widetilde{N}$ signed Bernoulli draws, with an unknown success probability coming from a uniform prior, i.e., $\alpha = \beta = 1$. Then the posterior distribution $Pr(N|m)$ is approximated by a scaled Poisson. Specifically, for $N \geq \max(|m|, 1)$ satisfying $|m| + N$ even,*

1. *If $m = 0$, then*

$$Pr(N|m = 0, \alpha = \beta = 1) \approx 2\frac{\lambda^{N+1}e^{-\lambda}}{(N+1)!}$$

2. *If $m > 0$ and m is even, then*

$$Pr(N|m, \alpha = \beta = 1) \geq \left(\frac{2(|m|/2)!}{\lambda^{|m|/2}}\right)\left(\frac{\lambda^{N+1}e^{-\lambda}}{(N+1)!}\right)$$

3. *If m is odd, then*

$$Pr(N|m, \alpha = \beta = 1) \geq \left(\frac{2((|m|+1)/2)!}{\lambda^{(|m|+1)/2}}\right)\left(\frac{\lambda^{N+1}e^{-\lambda}}{(N+1)!}\right)$$

In case 1, the posterior converges in λ to the shown approximation. In cases 2 and 3, the posterior either converges in λ to the shown approximation, or it differs only in the scaling factor. In all cases, the posterior converges in λ to a scaled Poisson.

The approximations in Proposition 3 provide a convenient way to estimate $E[\widetilde{N}|m, \alpha =$

$\beta = 1]$. For example, if $m = 0$, then

$$E[\widetilde{N}|\alpha = \beta = 1, m] = \sum_{N=2}^{\infty} 2N \frac{\lambda^{N+1} e^{-\lambda}}{(N+1)!} = 2e^{-\lambda} (\lambda \cosh \lambda - \sinh \lambda)$$

Because $\cosh \lambda \approx \sinh \lambda \approx e^{\lambda}/2$ for large enough λ, this value is approximated by

$$E[\widetilde{N}|\alpha = \beta = 1, m] \approx 2e^{-\lambda}(\lambda - 1)e^{\lambda}/2 = \lambda - 1$$

Other approximations are similar.

Summing up, a nonstrategic, fully aggregate report produces a closed-form, analytically tractable posterior distributions of the number of signals received, which is approximately Poisson in the special case of a uniform prior.

Strategic Disclosure

We now focus on the manager's strategic aggregation decisions. In particular, we show that the manager always chooses to disclose at least the aggregate report. Whether the manager discloses more than this depends on whether his posterior $E[\widetilde{\pi}|y] > 1/2$.

To make the definitions of disclosure strategies and of equilibrium precise, we introduce some notation. For arbitrary sets A, B, write B^A for the collection of functions from $A \longrightarrow B$. For $n \in \mathbb{Z}_{++}$, let

$$Y := \bigcup_{n=1}^{\infty} \{-1, 1\}^n$$

The set Y represents the possible collections of signals the manager might receive. We define the set of possible disclosures, denoted Z, as

$$Z = \left(\bigcup_{n=1}^{\infty} \bigcup_{v=1}^{n} \left\{ z \in \{-n+v-1, \ldots, n-v+1\}^v \,\middle|\, \sum_{j=1}^{v} z_j \in \{-n, -n+2, \ldots, n\} \right\} \right) \bigcup \{\varnothing\}$$

That is, if we fix $n \in \mathbb{Z}_{++}$, then the disclosure, if made, can have anywhere from 1 to n

subtotals. If the disclosure has v subtotals, then each subtotal can be based on at most $n - v + 1$ signals. This means that each subtotal cannot be larger than $n - v + 1$ in absolute value. Moreover, the sum of the subtotals cannot be larger than n in absolute value, and can only differ from n in absolute value by an even number. This is because any such difference must be driven by the number of signals that cancel, i.e., pairs (y_i, y_j) with $y_i = -y_j$. Non-disclosure is also feasible.

For $x \in Y \cup Z$, let $\ell(x)$ be the length of x. We can now precisely define a strategy for the manager:

Definition 1. A *disclosure strategy* is a function $m \in Z^Y$ such that, for all $y \in Y$, letting $v(y) = (\ell \circ m)(y)$, there exists a partition of the coordinate positions, $S_1, \ldots, S_{v(y)} \subset \{1, \ldots, \ell(y)\}$, satisfying truthful aggregation. That is:

1. $(\forall j \in \{1, \ldots, v(y)\})\ S_j \neq \varnothing$ (nondegeneracy)

2. $(\forall j, j' \in \{1, \ldots, v(y)\})\ S_j \cap S_{j'} = \varnothing$ (no double counting)

3. $\bigcup_{j=1}^{v(y)} S_j = \{1, \ldots, \ell(y)\}$ (no omissions)

4. $(\forall j \in \{1, \ldots, v(y)\})\ m_j(y) = \sum_{i \in S_j} y_i$ (truthful aggregation)

Intuitively, m aggregates $y \in Y$ into a comprehensive collection of distinct subtotals, where the composition and number of subtotals is part of the strategy.

The market's strategy is simply to price the firm at its expected value, given all information and beliefs about manager's reporting strategy. To rule out degenerate equilibria, we impose rational expectations. The definition of equilibrium is as follows:

Definition 2. A *Perfect Bayesian Equilibrium* is a disclosure strategy $m^* \in Z^Y$, a pricing rule p^*, and a rule for updating beliefs about $\widetilde{\pi}$ conditional upon $m^*(y)$ such that

1. Given the report m^*, the price satisfies $p^* = E[\widetilde{\pi}|m^*]$.

2. Given the market's rule for updating beliefs,

$$m^* \in \arg\max_{m \in Z^Y} E[\widetilde{\pi}|m]$$

3. The market's beliefs are rational, with updating consistent with Bayes' rule, given m^*, and

4. If possible, the market's beliefs are consistent with Bayes' rule off the equilibrium path.

Our first result is that the manager always discloses at least the aggregate report:

Proposition 4. *In every Perfect Bayesian Equilibrium, a sufficient statistic for the number of net successes is disclosed. That is,*

$$(\forall y \in Y)\ (\ell \circ m^*)(y) > 0,$$

which, by conditions 2–4 of Definition 1, fully reveals $\sum_{i=1}^{\ell(y)} y_i$.

Although Proposition 4 shows that the manager always discloses the aggregate total, our main result is that the manager does not generically disclose a sufficient statistic for N. The aggregate is a sufficient statistic for whether $E[\widetilde{\pi}|y] > 1/2$. If it is, the manager prefers to keep the market's beliefs about N as low as possible. On the other hand, if $E[\widetilde{\pi}|y] < 1/2$, the manager discloses a sufficient statistic for N, in order to signal the largest possible number of offsetting pairs. The driving force of this result is that, given the number of net successes, the larger N is, the closer the posterior is to 1/2.

Theorem 1. *Let* $\widetilde{\pi} \sim Beta(\alpha, \beta)$. *Let* $s = \#\{i|y_i = 1\}$ *be the realized number of successes, and let* N *be the realized number of signals. In an essentially unique Perfect Bayesian Equilibrium, the manager fully reveals his information only if the posterior mean of* $\widetilde{\pi}$ *given his private information is at most* 1/2*, i.e., if and only if*

$$E[\widetilde{\pi}|y] = \frac{\alpha + s}{\alpha + \beta + N} \leq \frac{1}{2}$$

In this case, the market in equilibrium believes the disclosure is complete.

If the posterior mean of $\widetilde{\pi}$ is no less than 1/2, it is optimal for the manager to disclose the net number of successes $m^1(y)$ and no further information. In this case, the market calculates $E[\widetilde{\pi}|m^1(y)]$ using the conditional distribution in Proposition 2. The boundary case of $E[\widetilde{\pi}|y] = 1/2$ is consistent with both aggregate disclosure and full revelation.

Full disclosure, in the equilibrium of Theorem 1, can be achieved with as few as two segments, provided $E[\widetilde{\pi}|y] \leq 1/2$. Let

$$\begin{aligned} S_1 &= \{i \in \{1, \ldots, \ell(y)\} | y_i = 1\} \\ S_2 &= \{i \in \{1, \ldots, \ell(y)\} | y_i = -1\} \end{aligned}$$

In other words, put all the successes into one subtotal and all the failures into another. If the market believes that the manager has adopted this strategy, then its posterior belief conditional upon the disclosure is

$$E[\widetilde{\pi}|m^2(y)] = \frac{\alpha + \#(S_1)}{\alpha + \beta + \#(S_1) + \#(S_2)}$$

provided this posterior mean is at most 1/2. If the manager were to deviate and put some failures into subtotal $m_1^2(y)$ or some successes into subtotal $m_2^2(y)$, then the market would believe that there were fewer successes and failures, with each dropping by an equal number. This would drop from the posterior some signals that would have moved it closer to 1/2. Because the posterior is below 1/2, the effect would be to lower the posterior.

Conversely, if the posterior mean is above 1/2, then if the manager can successfully conceal two offsetting signals by putting them in the same subtotal, he will benefit by doing so. Each pair of offsetting signals that the market observes pushes the posterior mean closer to 1/2, which is undesirable when the posterior mean is higher. It therefore is optimal for the manager to conceal as many offsetting signals as possible, and this is achieved by putting all the signals into a single total.

In fact, even a *single* aggregate disclosure, $m^1(y)$, is consistent with full revelation in equilibrium, given $E[\widetilde{\pi}|y] \leq 1/2$. If $y_1 = \ldots = y_N$, then the manager sums all the signals (whether positive or negative) into one total. The market knows that it is not in the manager's interest to conceal offsetting signals when the posterior mean of $\widetilde{\pi}$ is below 1/2, and correctly infers that the reason the manager does not segment is that every signal had the same value. This means that the market attaches a different interpretation to an aggregate report, depending on the posterior mean. If the posterior mean is above 1/2, the market attaches positive probability to some offsetting signals being present. If the posterior mean is below 1/2, it does not.[3]

Figure 2 illustrates the intuition of Theorem 1 under the assumption of a uniform prior, i.e., $\alpha = \beta = 1$, and two different realizations of y which result in a net number of successes of 1 and -1, respectively. Panel (a) shows the market's posterior mean of $\widetilde{\pi}$, conditional upon the aggregate report $m^1(y)$ and its conjecture of N. Given the uniform prior, a net positive aggregate signal makes the posterior mean of $\widetilde{\pi} > 1/2$. However, the higher the market's conjecture of N, the closer the posterior moves downward toward 1/2. For a net negative aggregate signal, the opposite is true. Panel (b) shows the market's posterior mean of $\widetilde{\pi}$ conditional upon $m^1(y)$ and its average beliefs using (2). Observe that disaggregation drops *low* values of N from the set of feasible numbers of signals first. This causes disaggregation to lower the posterior mean of $\widetilde{\pi}$ when $m^1(y) = 1$ and to increase the posterior mean when $m^1(y) = -1$.

[FIGURE 2 AROUND HERE]

IV Discussion and Conclusion

The main insight of this paper is that firms with low expected values can credibly reveal information by making granular disclosures about their aggregate performance. Firms with high expected values cannot, and optimally choose to report coarser, more aggregate information than low expected value firms choose. The reasons are purely value maximization, and are unrelated to the usual suspects of risk attitudes, verification, or an asymmetric market response to good or bad news such as with debt finance. We study a manager who is risk neutral, does not lie, and is concerned solely with maximizing the value of the company's share price in a market with rational expectations.

The driving force of this result is that aggregation cancels offsetting signals. If a firm's value is low, then the more neutral news revealed, the better. If a firm's value is high, the more neutral news concealed, the better. The intuition is as follows: suppose you draw a random sample of size N from your favorite distribution with a finite mean, and sum the realizations. If the mean of the distribution is μ, then the mean of this sum is $N \cdot \mu$. Imagine that you want to convince us that μ is as high as possible, and that we can observe $N \cdot \mu$ but not N or μ directly. If $N \cdot \mu$ is negative, it is in your interest to show us the records of your individual draws. If $N \cdot \mu$ is positive, then the smaller you can convince us that N might be, the better off you are. With news that is good in aggregate, the firm's manager, like you in this toy example, would state the result and then be silent. The same manager, faced with bad news, has very strong incentive to put the bad news into a broader context, showing as much neutral news as possible.

This result is similar in spirit to earlier results by Penno (1996), Einhorn (2005), and Dziuda (2011). Penno studies a firm whose manager strategically chooses the precision of its disclosure. For firms with good news, low precision is desirable, while firms with bad news prefer higher precision. Einhorn studies an aggregation decision with two signals, differing in their precisions. She finds that the manager reports the disaggregate signal if and only if the more precise signal is good news for the firm. In our context, disaggregating reveals more precise information, and this is desirable if and only if revealing the number of offsetting signals is good news. Dziuda studies a possibly biased sales representative. Unlike our manager, her sales representative can withhold information, but, as in our setting, her representative cannot lie. Dziuda finds a benefit to revealing information about the number of received signals, even if doing so requires disclosing unfavorable information. Although our context and our notion of good news is very different from Dziuda's, Einhorn's, and Penno's, the results are clearly related.

There is a close connection between our main result and the literature on accounting conservatism. Gigler et al. (2009) demonstrate that, because conservative reporting makes it more difficult to report good news than bad news, a conservative regime leads to reported good news being more informative than reported bad news. Whether this property is desirable depends on the relative importance of Type I versus Type II errors. Our main result shows that discretionary aggregation has the opposite effect: reported good news is less informative than reported bad news, due to the manager's incentive to provide an explanation for bad news and not to provide any information that would lead the market to question good news. A potential issue for future work is the whether the optimal level of accounting conservatism depends on the degree of discretion in aggregation.

From an empirical viewpoint, the forces described here provide one reason that markets may react asymmetrically to bad news compared with good news. A bad news disclosure includes any information that would dampen the effect of the news, showing that the news is as close as possible to neutral. A good news disclosure may have omitted attenuating details. The market then rationally is muted in its reaction to good news, but has no corresponding reason to mute its response to bad news.

Our setting focuses on risk neutrality, but has natural extensions to different risk attitudes. A manager who is given incentive compensation designed to induce risk taking faces a trade-off in aggregation decisions. If the firm has a low expected value, the manager's benefits from revealing disaggregate information are partially offset by a reduction in volatility coming from more draws being revealed. A risk averse manager faces the opposite trade-off, and may reveal disaggregate information for a highly valued firm in order to reduce volatility. These forces change nothing essential in our argument, but may shift the manager's cutoff from 1/2 to a different critical value.

The framework of our model may be of independent interest from a purely technical perspective. Our Poisson-Signed Bernoulli-Beta (PSBB) setting has many convenient features. Some advantages of the individual components of our setting are as follow: first, for studying sender-receiver games where the sender of a message has received an unknown and unbounded number of private signals, it does not get any easier than a Poisson. Second, a central feature of aggregation, and how it can hide information, arises from having offsetting signals summed. Signed Bernoulli draws are the simplest environment in which this feature of aggregation arises. Third, the Beta distribution is the most natural one for an unknown probability of success on a (signed or unsigned) Bernoulli trial. It has the uniform as a special case, and in many important respects, it is simply the likelihood function of the binomial distribution in disguise. In addition to these important features, the PSBB setting has convenient properties arising from combining these distributions. Given the number of trials, the posterior distribution of the number of net successes is (other than a scaling factor) the standard hypergeometric distribution. Conversely, in the special case of a uniform prior, given the number of net successes, the posterior distribution of the number of trials is a scaled Poisson. These features make the PSBB setting extremely convenient: everything is closed-form, and almost everything can be expressed as a familiar distribution or something very close to it.

We end with an illustrative anecdote. In its 2009 annual report, Citigroup announced a change in its reporting segments, as part of the aftermath of the Great Recession. Their new reporting format included two segments, Citicorp and Citi Holdings. The annual report (iii) reads in part as follows: "Into Citicorp, we placed the businesses that are core to our strategy and that offer shareholders the greatest earnings potential within appropriate risk parameters.... In Citi Holdings, we assembled assets and businesses that are not central to our strategy.... Many are economically sensitive." In terms of our setting, Citicorp first observed their overall news, and for any financial services company, the news in 2009 was bad overall. In response, Citicorp chose to aggregate its positive news into one segment and its negative news into another. This is exactly what our model predicts, and was a credible way, given the economic conditions of the time, of informing the market of the company's overall financial condition.

Proofs

Proof of Proposition 1. Given $\langle N; m^1(y)\rangle$, the market can infer

$$\#\{i|y_i = 1\} = \frac{N + m^1(y)}{2} \qquad \text{and} \qquad \#\{i|y_i = -1\} = \frac{N - m^1(y)}{2}$$

Given y, the posterior distribution on $\widetilde{\pi}$ is

$$\widetilde{\pi}|y \sim Beta(\alpha + \#\{i|y_i = 1\}, \beta + \#\{i|y_i = -1\})$$

That is, $\langle N; m^1(y)\rangle$ conveys all the manager's information on $\widetilde{\pi}$.

The market has rational expectations, and the firm's value is distributed Bernoulli($\widetilde{\pi}$). Therefore, the only economically relevant information for the market is news that affects its posterior beliefs about $\widetilde{\pi}$. □

Proof of Lemma 1. The uniform case is a special case of the Beta(α, β) case, obtained immediately from setting $\alpha = \beta = 1$. So it suffices to consider the general Beta(α, β) setting.

The probability of observing m net successes is

$$\begin{aligned}
Pr\left(m\,|N, N + |m| \text{ even}\right) &= \int_0^1 \binom{N}{\frac{N+m}{2}} \pi^{\frac{N+m}{2}} (1-\pi)^{\frac{N-m}{2}} f_{\widetilde{\pi}}(\pi) d\pi \\
&= (\alpha + \beta - 1)\binom{\alpha+\beta-2}{\alpha-1}\binom{N}{\frac{N+m}{2}} \int_0^1 \pi^{\alpha+\frac{N+m}{2}-1}(1-\pi)^{\beta+\frac{N-m}{2}-1} d\pi \\
&= \left[\frac{\alpha+\beta-1}{\alpha+\beta+N-1}\right] \frac{\binom{\alpha+\beta-2}{\alpha-1}\binom{N}{\frac{N+m}{2}}}{\binom{\alpha+\beta+N-2}{\alpha+\frac{N+m}{2}-1}} \\
&\quad \cdot \int_0^1 \frac{1}{B\left(\alpha+\frac{N+m}{2}, \beta+\frac{N-m}{2}\right)} \pi^{\alpha+\frac{N+m}{2}-1}(1-\pi)^{\beta+\frac{N-m}{2}-1} d\pi \\
&= \left[\frac{\alpha+\beta-1}{\alpha+\beta+N-1}\right] \frac{\binom{\alpha+\beta-2}{\alpha-1}\binom{N}{\frac{N+m}{2}}}{\binom{\alpha+\beta+N-2}{\alpha+\frac{N+m}{2}-1}} \\
&= \left[\frac{\text{Prior }\#(\text{total successes possibilities})}{\text{Posterior }\#(\text{total successes possibilities})}\right] \cdot \begin{array}{l}\text{Hypergeometric} \\ \text{probability mass}\end{array}
\end{aligned}$$

□

Proof of Lemma 2. Using Lemma 1, we expand $Pr(m \wedge N) = Pr(m|N)Pr(N)$ as

$$
\begin{aligned}
Pr(m \wedge N) &= \left[\frac{\lambda^N}{N!\,(e^\lambda - 1)}\right]\left[\frac{\alpha+\beta-1}{\alpha+\beta+N-1}\right]\frac{\binom{\alpha+\beta-2}{\alpha-1}\binom{N}{\frac{N+m}{2}}}{\binom{\alpha+\beta+N-2}{\alpha+\frac{N+m}{2}-1}} \\
&= \left[\frac{\alpha+\beta-1}{(e^\lambda-1)}\right]\cdot\left[\frac{\lambda^N}{(\alpha+\beta+N-1)\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!}\right]\frac{\binom{\alpha+\beta-2}{\alpha-1}}{\binom{\alpha+\beta+N-2}{\alpha+\frac{N+m}{2}-1}} \\
&= \left[\frac{(\alpha+\beta-1)\binom{\alpha+\beta-2}{\alpha-1}}{(e^\lambda-1)}\right]\cdot\left[\frac{\lambda^N}{(\alpha+\beta+N-1)\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!}\right] \\
&\quad /\binom{\alpha+\beta+N-2}{\alpha+\frac{N+m}{2}-1}
\end{aligned}
$$

That is, the $N!$ terms cancel, and we can factor the terms that do not depend on N. The special case of a uniform follows from direct substitution. □

Proof of Proposition 2. We begin by solving for the marginal distribution of the number of net successes. From Lemma 2,

$$
\begin{aligned}
Pr(m) &= \sum_{\substack{k\geq\max(|m|,1)\\ k+|m| \text{ even}}} Pr\left(m \wedge \widetilde{N} = k\right) \\
&= \left[\frac{(\alpha+\beta-1)\binom{\alpha+\beta-2}{\alpha-1}}{(e^\lambda-1)}\right]\sum_{\substack{k\geq\max(|m|,1)\\ k+|m| \text{ even}}}\left[\frac{\lambda^k\left(\alpha+\frac{k+m}{2}-1\right)!\left(\beta+\frac{k-m}{2}-1\right)!}{\left(\frac{k+m}{2}\right)!\left(\frac{k-m}{2}\right)!(\alpha+\beta+k-1)!}\right]
\end{aligned}
$$

Using the fact that $Pr(N|m) = Pr(m \wedge N)/Pr(m)$, we obtain the posterior distribution of N:

$$
\begin{aligned}
Pr(N|m) &= \frac{Pr(m \wedge N)}{Pr(m)} \\
&= \left[\frac{\lambda^N\left(\alpha+\frac{N+m}{2}-1\right)!\left(\beta+\frac{N-m}{2}-1\right)!}{\left(\frac{N+m}{2}\right)!\left(\frac{N-m}{2}\right)!(\alpha+\beta+N-1)!}\right] \\
&\quad /\sum_{\substack{k\geq\max(|m|,1)\\ k+|m| \text{ even}}}\left[\frac{\lambda^k\left(\alpha+\frac{k+m}{2}-1\right)!\left(\beta+\frac{k-m}{2}-1\right)!}{\left(\frac{k+m}{2}\right)!\left(\frac{k-m}{2}\right)!(\alpha+\beta+k-1)!}\right]
\end{aligned}
$$

The uniform special case comes from directly substituting $\alpha = \beta = 1$. □

Proof of Proposition 3. We have from (2)

$$Pr(N|m, \alpha = \beta = 1) = \left[\frac{\lambda^{N+1}}{(N+1)!}\right] / \left[\sum_{\substack{k \geq \max(|m|,1) \\ k+|m| \text{ even}}} \frac{\lambda^{k+1}}{(k+1)!}\right]$$

$$= \begin{cases} \left[\frac{\lambda^{N+1}}{(N+1)!}\right] / \left[\sum_{k=\frac{|m|}{2}}^{\infty} \frac{\lambda^{2k+1}}{(2k+1)!}\right], & \text{if } |m| > 0, \text{ even} \\ \left[\frac{\lambda^{N+1}}{(N+1)!}\right] / \left[\sum_{k=1}^{\infty} \frac{\lambda^{2k+1}}{(2k+1)!}\right], & \text{if } m = 0 \\ \left[\frac{\lambda^{N+1}}{(N+1)!}\right] / \left[\sum_{k=\frac{|m|+1}{2}}^{\infty} \frac{\lambda^{2k}}{(2k)!}\right], & \text{if } |m| \text{ odd} \end{cases}$$

In the special case where $m = 0$, the sum in the denominator includes all except the first term in the Taylor series expansion of $\sinh \lambda$ around 0. That means the denominator becomes

$$\sinh \lambda - \lambda$$

and hence that

$$\begin{aligned} Pr(N|m = 0, \alpha = \beta = 1) &= \frac{\lambda^{N+1}}{(N+1)!(\sinh \lambda - \lambda)} \\ &= \frac{\lambda^{N+1}}{(N+1)!\left(\frac{e^{\lambda}-e^{-\lambda}}{2} - \lambda\right)} \\ &\searrow 2\frac{\lambda^{N+1}e^{-\lambda}}{(N+1)!} \end{aligned}$$

The other cases make use of the Taylor Approximation theorem (which is in standard calculus texts):

Theorem (Taylor Approximation Theorem)**.** *Assume $f \in C^{N+1}[a, b]$. Fix $x_0, x \in [a, b]$. Let $P_N(x)$ be the N^{th}-order Taylor approximation of $f(x)$ around x_0:*

$$P_N(x) := \sum_{k=0}^{N} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

The remainder term, $R_N(x) := f(x) - P_N(x)$, has the form

$$R_N(x) = \frac{f^{(N+1)}(c)}{(N+1)!}(x - x_0)^{N+1}$$

where $c \in [x_0, x]$.

Using the Taylor Approximation Theorem, the fact that sinh and cosh are monotone on $[0, \infty)$, and the fact that $d(\cosh \lambda)/d\lambda = \sinh \lambda$ and $d(\sinh \lambda)/d\lambda = \cosh \lambda$, we obtain

$$Pr(N|m,\alpha=\beta=1) = \begin{cases} \left[\frac{\lambda^{N+1}}{(N+1)!}\right] / \left[\sum_{k=\frac{|m|}{2}}^{\infty} \frac{\lambda^{2k+1}}{(2k+1)!}\right], & \text{if } |m|>0, \text{ even} \\ \left[\frac{\lambda^{N+1}}{(N+1)!}\right] / \left[\sum_{k=\frac{|m|+1}{2}}^{\infty} \frac{\lambda^{2k}}{(2k)!}\right], & \text{if } |m| \text{ odd} \end{cases}$$

$$\geq \begin{cases} \frac{\lambda^{N+1-|m|/2}(|m|/2)!}{(N+1)!\sinh\lambda}, & |m|>0, \text{ even}, |m|/2 \text{ even} \\ \frac{\lambda^{N+1-|m|/2}(|m|/2)!}{(N+1)!\cosh\lambda}, & |m|>0, \text{ even}, |m|/2 \text{ odd} \\ \frac{\lambda^{N+1-(|m|+1)/2}((|m|+1)/2)!}{(N+1)!\cosh\lambda}, & |m| \text{ odd}, (|m|+1)/2 \text{ even} \\ \frac{\lambda^{N+1-(|m|+1)/2}((|m|+1)/2)!}{(N+1)!\sinh\lambda}, & |m| \text{ odd}, (|m|+1)/2 \text{ odd} \end{cases}$$

$$\to \begin{cases} \left[\frac{2(|m|/2)!}{\lambda^{|m|/2}}\right]\left[\frac{\lambda^{N+1}e^{-\lambda}}{(N+1)!}\right], & |m|>0, \text{ even}, |m|/2 \text{ even} \\ \left[\frac{2((|m|+1)/2)!}{\lambda^{(|m|+1)/2}}\right]\left[\frac{\lambda^{N+1}e^{-\lambda}}{(N+1)!}\right], & |m| \text{ odd} \end{cases}$$

The inequality arises from using the upper bound on the remainder from the Taylor approximation in the denominator. By the Taylor Approximation theorem, the exact value of the remainder may have a different coefficient on the sinh or cosh terms, but otherwise would be unchanged. Hence, other than the first term, the result would be unchanged, and the posterior in all cases converges to a scaled Poisson. □

Proof of Proposition 4. Let $\widehat{p}$ be the manager's conjecture of the market's belief $E[\widetilde{\pi}|$ nondisclosure$]$. Recall that $\alpha, \beta \in \mathbb{Z}_{++}$, and that $\ell(y) < \infty$. This means that, for any signal $y \in Y$, rational expectations require that the market's updated belief satisfy $0 < E[\widetilde{\pi}|$ nondisclosure$] < 1$. By the Archimedean property, there is some $n \in \mathbb{Z}_{++}$ such that, for all $N \geq n$,

$$\frac{\alpha+N}{\alpha+\beta+N} > \widehat{p} \tag{4}$$

Let $m > n$, so that (4) holds. Suppose the manager reports $m^1(y) = m$. By Proposition 1,

$$\begin{aligned} \frac{Pr(\widetilde{N}=m|m^1(\widetilde{y})=m)}{Pr(\widetilde{N}=m+2|m^1(\widetilde{y})=m)} &= \frac{\lambda^m(\alpha+m-1)!(\beta-1)!}{m!(\alpha+\beta+m-1)!}\cdot\frac{(m+2)!(\alpha+\beta+m+1)!}{\lambda^{m+2}(\alpha+m+1)!(\beta+1)!} \\ &= \frac{(m+2)(m+1)(\alpha+\beta+m+1)(\alpha+\beta+m)}{\lambda^2(\alpha+m+1)(\alpha+m)(\beta+1)\beta} \end{aligned} \tag{5}$$

The numerator of (5) is on the order of m^4, while the denominator is on the order of m^2. Thus, the ratio (5) grows arbitrarily large as m increases. Since $Pr(\widetilde{N}=m|m^1(\widetilde{y})=m)$ is bounded above, it must be the case that the denominator asymptotically approaches 0 for large enough m. A similar argument shows that $Pr(\widetilde{N}=m+2|m^1(\widetilde{y})=m)/Pr(\widetilde{N}=m+4|m^1(\widetilde{y})=m) \to \infty$ as m increases. By a straightforward inductive argument, it follows that, for large enough m, $Pr(\widetilde{N}=m|m^1(\widetilde{y})=m)$ gets arbitrarily close to 1. In particular, if m is sufficiently large, then rational expectations imply

$$E[\widetilde{\pi}|m] \approx \frac{\alpha+m}{\alpha+\beta+m} > \widehat{p}$$

This means that, for some value m^*, if $m^1(y) \geq m^*$, the manager prefers disclosing $m^1(y)$ to non-disclosure.

Accordingly, a necessary condition for non-disclosure is

$$E[\widetilde{\pi}|m^1(y)] \leq E[\widetilde{\pi}| \text{ no disclosure}] = \widehat{p}$$

Let m^* satisfy

$$E[\widetilde{\pi}|m^1(y) = m^* - 1] \leq \widehat{p} < E[\widetilde{\pi}|m^1(y) = m^*]$$

The manager would then not disclose if and only if $m^1(y) \leq m^* - 1$. Since $E[\widetilde{\pi}|m^1(y)]$ is monotone in the aggregate disclosure, and aggregate disclosures below $m^* - 1$ occur with positive probability, rational expectations require $E[\widetilde{\pi}|$ no disclosure$] < \widehat{p}$. This is a contradiction, implying there can be no cutoff $\widehat{p} > 0$ below which the firm does not release the aggregate disclosure. □

Proof of Theorem 1. First, we consider the case where $E[\widetilde{\pi}|y] < 1/2$. If the manager discloses $m^1(y)$, the market's beliefs are that this is a sufficient statistic for full revelation, i.e., all y_i have the same sign, and $m^1(y) = N$, provided $m^1(y) \neq 0$. Similarly, if the manager reports $m^v(y)$ for $v > 1$, the market can assume full revelation, with all the successes divided among the subtotals that are positive, and all the failures divided among the subtotals that are negative, provided none of the subtotals is equal to 0. To show that full disclosure is a Perfect Bayesian Equilibrium, it suffices for us to consider a case where the market treats any subtotal of 0 as containing exactly two signals, one success and one failure. The idea behind this belief is that the market treats the manager as playing the equilibrium if possible, and if not, the market treats the manager as playing the closest strategy to the equilibrium that is consistent with the disclosure.

On the equilibrium path, it suffices to consider the case where the manager discloses at most two subtotals, $m^2(y)$, putting all the successes into one subtotal and all the failures into the other. Let

$$S_1 = \{i \in \{1, \ldots, n\} | y_i = 1\} \qquad S_2 = \{i \in \{1, \ldots, n\} | y_i = -1\}$$

$$s = \sum_{i \in S_1} y_i = \#(S_1) \qquad f = N - s = \sum_{i \in S_2} |y_i| = \#(S_2)$$

If $s = 0$, the market's updated belief is

$$E[\widetilde{\pi}|m^1(y) = -f] = \frac{\alpha}{\alpha + \beta + f}$$

and any other disclosure, given the market's belief, would lead to the same posterior mean. Similarly, if $f = 0$, any disclosure is fully revealing, giving the market a (correct) posterior of

$$E[\widetilde{\pi}|m^1(y) = s] = \frac{\alpha + s}{\alpha + \beta + s}$$

If $\min(s, f) > 0$, the manager's strategy

$$E[\widetilde{\pi}|m^2(y)] = \frac{\alpha + s}{\alpha + \beta + s + f} < \frac{1}{2}$$

where the inequality is by hypothesis.

The argument is by induction. Define a *1-deviation* as a report $\langle s - 1, -(f - 1)\rangle$, generated as follows: for some $i \in S_1$ or some $j \in S_2$, let

$$S_1' = S_1 \backslash \{i\} \text{ and } S_2' = S_2 \cup \{i\} \quad \text{ or let } \quad S_1' = S_1 \cup \{j\} \text{ and } S_2' = S_2 \backslash \{j\}$$

Either way, let

$$s' = \sum_{i \in A_1'} y_i \quad \text{ and } \quad f' = -\sum_{i \in A_2'} y_i$$

If the investors believe that the report was truthful, then the aggregate signal would be interpreted as $s - 1 + f - 1 = s - f$, so there is no change in the aggregate report. However, the number of informative signals would appear to be $s - 1 + f - 1 = s + f - 2$. Thus, a 1-deviation is a report that hides a pair of opposed informative signals $\{-1, 1\}$.

The market's posterior belief, upon obtaining a message with a 1-deviation, is

$$E[\widetilde{\pi} | m^2(y) = (s', -f')] = \frac{\alpha + s - 1}{\alpha + \beta + s + f - 2}$$

The 1-deviation is therefore unprofitable if and only if

$$\begin{aligned} \frac{\alpha + s - 1}{\alpha + \beta + s + f - 2} &\leq \frac{\alpha + s}{\alpha + \beta + s + f} \\ \Leftrightarrow \quad \frac{\alpha + s}{\alpha + \beta + s + f} &\leq \frac{1}{2} \end{aligned} \tag{6}$$

By hypothesis, the posterior mean given y is below 1/2. Therefore, the manager is worse off by making a 1-deviation than by reporting truthfully.

Define an *h-deviation* as a 1-deviation from an *$(h-1)$-deviation*. Assume the posterior upon receiving an $(h-1)$-deviation is at most 1/2. Then, by an analogous argument, the posterior upon receiving an h-deviation and believing it to be fully revealing is weakly below the posterior upon receiving an $(h-1)$-deviation. Hence, if the posterior upon receiving a fully revealing disclosure is at most 1/2, then for any h, no h-deviation is profitable.

To finish the case where $E[\widetilde{\pi}|y] < 1/2$, suppose that the manager discloses a subtotal that equals 0. That is, for some $v \in \mathbb{Z}_{++}$, the manager's disclosure $m^v(y)$ contains at least one entry (which, without loss of generality, we can take to be $m_1^v(y)$) that equals 0. The market can infer that there must be at least two signals in subtotal $m_1^v(y)$, and that these are offsetting. If the manager includes exactly two offsetting signals into this subtotal, then his report is equivalent to including the positive signal in a different subtotal with a positive sum, and including the negative signal in a different subtotal with a negative sum. This means that a zero subtotal containing exactly two offsetting signals cannot add any benefit over separating the positive and negative signals. On the other hand, if $m_1^v(y)$ contains more than two signals, then it is equivalent, for some h, to an $h-$deviation, making the manager strictly worse off.

Suppose now instead that the posterior mean of $\widetilde{\pi}$ upon receiving a fully revealing disclosure is above 1/2. This case is symmetric: by (6), the manager is always better off making a 1-deviation in a report, if believed at face value, than in reporting truthfully. By an analogous inductive argument, it is clear that, for every $h \in \mathbb{Z}_{++}$, an h-deviation is profitable over an $(h-1)$-deviation. This establishes that fully-revealing disclosure cannot be optimal when $E[\widetilde{\pi}] > 1/2$, and in fact that the manager's best response to a market that takes a report at face value would be to report a fully aggregate disclosure. Equivalently, the manager could disclose multiple subtotals, provided they all have the same sign.

If instead the manager always reports $m^1(y)$ when the posterior mean is above 1/2, then the market estimates

$$E[\widetilde{\pi}|m^1(y)] = E\left[\left.\frac{\alpha + \frac{m^1(y)+\widetilde{N}}{2}}{\alpha+\beta+\widetilde{N}}\right| m^1(y)\right]$$
$$= \sum_{\substack{N \geq \max(|m^1(y)|,1) \\ N+|m^1(y)| \text{ even}}} Pr(N|m^1(y)) \cdot \frac{\alpha + \frac{m^1(y)+N}{2}}{\alpha+\beta+N} \tag{7}$$

by Proposition 1, calculating the posterior probability in (7) using (1). Off-equilibrium, if the manager's disclosure reveals a sufficient statistic implying $N > |m^1(y)|$, suppose the market believes that there were a large number of offsetting signals. For example, the market may believe that the manager is more likely to mistakenly reveal an offsetting pair of signals when the number of such pairs is sufficiently large. For a large enough number of presumed offsetting signals, (7) gets arbitrarily close to 1/2, and therefore eventually lower than $E[\widetilde{\pi}|m^1(y)]$.

Finally, note that investors learn from the aggregated report, whether the posterior is above or below 1/2:

$$E[\widetilde{\pi}|y] = \frac{\alpha+s}{\alpha+s+\beta+f} \geq \frac{1}{2}$$
$$\Leftrightarrow \quad s - f \geq \beta - \alpha \tag{8}$$

$s-f$ is the aggregate report and $\beta - \alpha$ is common knowledge.

If condition (8) holds, investors interpret a two-dimensional message as $\langle s, -f \rangle$. This means that the market can always determine which side of 1/2 the posterior is on. Therefore, full disclosure when $E[\widetilde{\pi}|y] < 1/2$ and full aggregation when $E[\widetilde{\pi}|y] > 1/2$ is a Perfect Bayesian Equilibrium.

Essential uniqueness also follows from the fact that the market can always determine from the aggregate report whether the posterior is above or below 1/2. This implies that full aggregation cannot be optimal given that $E[\widetilde{\pi}|y] < 1/2$ unless the market believes $N = \infty$, an event with prior probability 0. Therefore, rational expectations and Bayes' rule imply the suboptimality of full aggregation in this case, and the same inductive argument as above implies that the manager's optimal strategy must fully reveal N. The case where $E[\widetilde{\pi}|y] > 1/2$ is analogous.

$\square$

Notes

[1]Crawford et al. (2012, 12) study IFRS 8 after 2009, the year in which it came into effect. In interviews with users and preparers of financial statements, they find that one of the main concerns raised was the choice of aggregation in constructing reporting segments. Additionally, they find an increase in the number of reporting segments in 2009, even for geographical information, despite the fact that geographical disclosure requirements became less restrictive under IFRS 8 than under the previous standard, IAS 14R. Given the drop in the market at the time—the Financial Times 100 and the Financial Times 250 indices fell 16.2% and 12.1%, respectively, from the start of 2008 to the end of 2009—the increase in the reporting granularity supports our results.

[2]For more on conditions when nondisclosure is feasible in equilibrium, see Verrecchia (1983), Dye (2001), Verrecchia (2001), and Beyer et al. (2010).

[3]As an extension, one might also consider a setting where the manager precommits to a disclosure policy, that is, restricts his choice of $m \in Z^Y$ such that, for all $y, y' \in Y$, if $\ell(y) = \ell(y')$ then (1) $v(y) = v(y')$, and (2) the partition $S_1, \ldots, S_{v(y)}$, defined in Definition 1, is the same for y and y'. It turns out that the results are entirely analogous to Theorem 1: the manager gives a fully aggregate disclosure whenever the *prior* mean, $E[\widetilde{\pi}] < 1/2$ and a fully disaggregate disclosure whenever the prior mean is below $1/2$. One difference is that the number of subtotals required for a fully revealing disclosure is larger under precommitment. To guarantee that a disclosure reveals everything, the manager would need to have at most two signals in each subtotal, and hence would need at least the ceiling of $N/2$ subtotals. This change is in the spirit of Sunder (1997, 89). Aside from the minimal message space size and the need to base the decision on ex ante information, the result is entirely as in Theorem 1. For related work on ex ante optimal aggregation rules, see Venezia (1978).

References

Arya, A., J. Fellingham, and D. Schroeder. 2000. Accounting information, aggregation, and discriminant analysis. *Management Science* 46: 790–806.

Arya, A., J. C. Fellingham, J. C. Glover, D. A. Schroeder, and G. Strang. 2000. Inferring transactions from financial statements. *Contemporary Accounting Research* 17: 365–385.

Behn, B. K., N. B. Nichols, and D. L. Street. 2002. The predicitive ability of geographic segment disclosures by U.S. companies: SFAS No. 131 vs. SFAS No. 14. *Journal of International Accounting Research* 1: 31–44.

Berger, P. G. and R. Hann. 2003. The impact of SFAS No. 131 on information and monitoring. *Journal of Accounting Research* 41: 163–223.

Beyer, A., D. A. Cohen, T. Z. Lys, and B. R. Walther. 2010. The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics* 50: 296–343.

Crawford, L., H. Extance, C. Helliar, and D. Power. 2012. Operating segments: The usefulness of IFRS 8. Technical report, Institute of Chartered Accountants of Scotland.

Dye, R. A.1985. Disclosure of nonproprietary information. *Journal of Accounting Research* 23: 123–145.

Dye, R. A.2001. An evaluation of "Essays on disclosure" and the disclosure literature in accounting. *Journal of Accounting and Economics* 32: 181–235.

Dye, R. A. and M. Finn. 2007. Equilibrium voluntary disclosures when firms possess random multi-dimensional private information. In R. Antle, F. Gjesdal, and P. J. Liang (Eds.), *Essays in Accounting Theory in Honour of Joel S. Demski*, Chapter 3, pp. 53–72. Springer-Verlag.

Dziuda, W.2011. Strategic argumentation. *Journal of Economic Theory* 146: 1362–1397.

Einhorn, E.2005. The nature of the interaction between mandatory and voluntary disclosures. *Journal of Accounting Research* 43: 593–621.

Ettredge, M. L., S. Y. Kwon, D. B. Smith, and M. S. Stone. 2006. The effect of SFAS No. 131 on the cross-segment variability of profits reported by multiple segment firms. *Review of Accounting Studies* 17: 91–117.

Ettredge, M. L., S. Y. Kwon, D. B. Smith, and P. A. Zarowin. 2005. The impact of SFAS No. 131 business segment data on the market's ability to anticipate future earnings. *The Accounting Review* 80: 773–804.

Gigler, F., C. Kanodia, H. Sapra, and R. Venugopalan. 2009. Accounting conservatism and the efficiency of debt contracts. *Journal of Accounting Research* 47: 767–797.

Grossman, S. J.1981. The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics* 24: 461–483.

Grossman, S. J. and O. D. Hart. 1980. Takeover bids, the free-rider problem, and the theory of the corporation. *The Bell Journal of Economics* 11: 42–64.

Herrmann, D. and W. B. Thomas. 2000. An analysis of segment disclosures under SFAS No. 131 and SFAS No. 14. *Accounting Horizons* 14: 287–302.

Jung, W.-O. and Y. K. Kwon. 1988. Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting Research* 26: 146–153.

Milgrom, P.1981. Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics* 12: 380–391.

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

Nichols, N. B. and D. L. Street. 1999. Segment information: What early adopters reported. *The Journal of Accountancy*: 37–41.

Penno, M.1996. Unobservable precision choices in financial reporting. *Journal of Accounting Research* 34: 141–149.

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

Street, D. L., N. B. Nichols, and S. J. Gray. 2000. Segment disclosures under SFAS No. 131: Has business segment reporting improved? *Accounting Horizons* 14: 259–285.

Sunder, S.1997. *Theory of Accounting and Control*. Thompson.

Venezia, I.1978. A decision-theoretic approach to the aggregation problem at the pre-data-collection stage. *Journal of the American Statistical Association* 73: 552–558.

Verrecchia, R. E.1983. Discretionary disclosure. *Journal of Accounting and Economics* 5: 179–194.

Verrecchia, R. E.2001. Essays on disclosure. *Journal of Accounting and Economics* 32: 97–180.

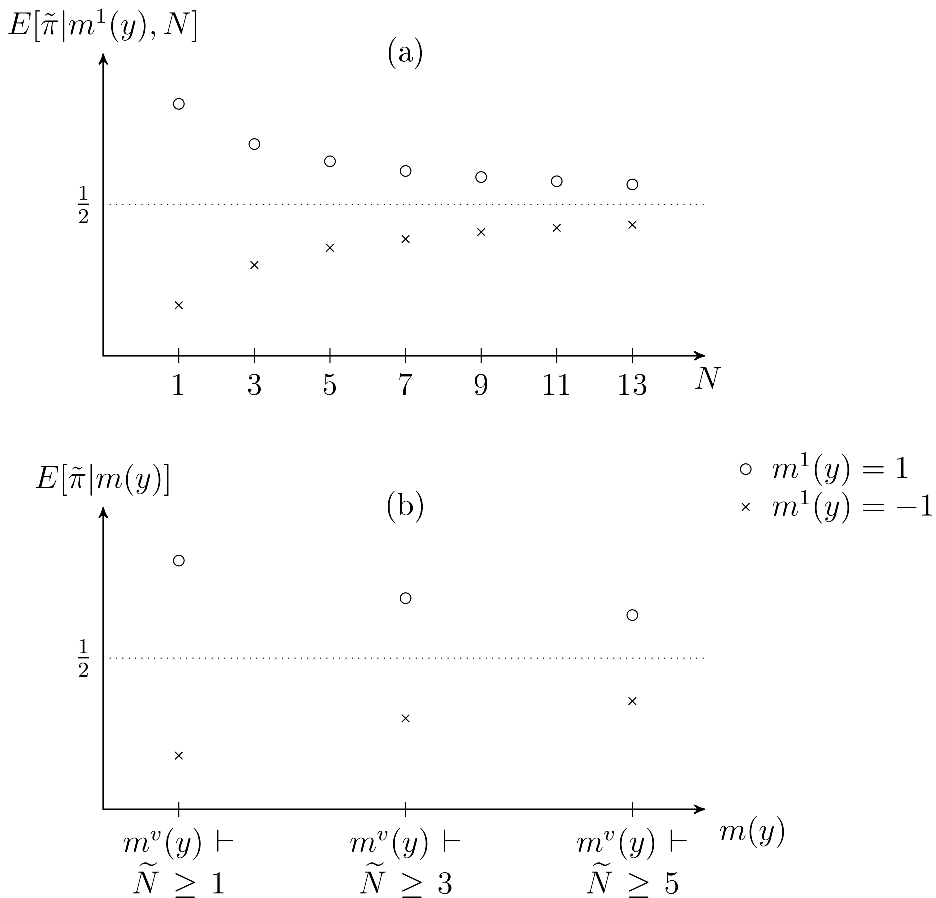

Figure 2: Illustration of rational beliefs for $E[\tilde{\pi}|m(y)]$ above and below 1/2