

Consistent Good News and Inconsistent Bad News

Rick Harbaugh, John Maxwell, and Kelly Shue*

Draft Version, April 2016

Abstract

Good news is more persuasive when it is more consistent, and bad news is less damaging when it is less consistent. We show when Bayesian updating supports this intuition so that a biased sender prefers more or less variance in the news depending on whether the mean of the news exceeds expectations, and we apply the result to selective news distortion by a manager of multiple projects. If news from the different projects is generally good, boosting relatively bad projects increases consistency across projects and provides a stronger signal that the manager is skilled. But if the news is generally bad, instead boosting relatively good projects reduces consistency and provides some hope that the manager is unlucky rather than incompetent. We test for evidence of such distortion by examining the consistency of reported segment earnings across different units in firms. As predicted by the model, we find that firms report more consistent earnings when overall earnings are above rather than below expectations. Firms appear to shift the allocation of overhead and other costs to help relatively weak units in good times and relatively strong units in bad times. The mean-variance news preferences that we identify apply in a range of situations beyond our career concerns application, and differ from standard mean-variance preferences in that more variable news sometimes helps and better news sometimes hurts.

*Harbaugh, Maxwell: Indiana University. Shue: University of Chicago and NBER. We thank conference and seminar participants at ESSET Gerzensee, Queen's University Economics of Organization Conference, NBER Corporate Finance Meetings, Indiana University, the University of Chicago, the University of Toronto, and UNSW for helpful comments. We are also grateful to Mike Baye, Phil Berger, Bruce Carlin, Archishman Chakraborty, Alex Edmans, Alex Frankel, Emir Kamenica, and Anya Kleymenova for helpful comments. We thank Tarik Umar for excellent research assistance.

1 Introduction

Consistently good performance at multiple tasks is a strong signal of competence, while consistently bad performance is a strong signal of incompetence. In a career concerns environment where individuals want to appear skilled, this creates an incentive to selectively distort reported performance on different tasks to affect the perceived consistency of performance. For instance, a manager might shift accounting costs to make one unit appear to do better than it really did, or might devote more time to strengthening the reported performance of one unit than another. In such cases, should a manager help a weaker unit shore up its performance, or instead focus on further improving a stronger unit?

We investigate this question when a manager's performance on each of multiple projects is news about the manager's skill, and the precision of these signals is uncertain. In a Bayesian model, we confirm the intuition that more consistent news about performance across the projects makes the news more reliable and hence has a stronger effect on the posterior estimate of the manager's skill. As a result, when news from the different projects is generally favorable, the manager increases his perceived skill the most by shoring up the performance of worse performing projects.¹ This makes reported performance more consistent across projects and hence makes the generally good performance on all projects a stronger signal of the manager's competence. But when the news is generally unfavorable, the manager should focus resources on making the least bad projects look better. This makes reported performance less consistent and hence makes the generally bad performance a weaker signal of the manager's incompetence.

We test the model's prediction that reported good news is more consistent than reported bad news using the variance of corporate earnings reports for different units or segments within conglomerate firms. Since some costs are shared by different units, managers can shift reported earnings across units by adjusting the allocation of these costs. We find evidence that managers shift costs to inflate the reported earnings of worse performing units when the firm is doing well overall. This makes it appear that all the units are doing similarly well, which is a more persuasive signal of management's abilities than if some units do very well while others struggle. But when the firm is doing poorly, managers shift costs to inflate the reported earnings of the relatively better performing units. This makes it appear that at least some units are doing not too badly, so there is more uncertainty about management's abilities and the overall evidence of bad performance is weaker.

Our empirical tests account for the possibility that segment earnings may be relatively more consistent during good times due to other natural factors. For example, bad times may cause higher volatility across corporate segments. To isolate variation in the consistency of segment

¹We abstract away from other issues that have been analyzed in the literature such as performance on some tasks being more observable than others, some tasks having higher returns for particular managers, or some tasks being complementary with each other. See Holmstrom and Milgrom (1992) and related literature.

earnings news that is likely to be caused by strategic distortions of cost allocations, we compare the consistency of segment earnings to that of segment sales. Like earnings, the consistency of segment sales may vary with firm performance for natural reasons. However, sales are more difficult to distort because they are reported prior to the deduction of costs. We find that segment sales do not appear to be more consistent when firm news is good rather than bad. As a further test, we compare the consistency of segment earnings in real multi-segment firms to that of counterfactual firms constructed from matched single-segment firms. Again, we find that the consistency of matched segment earnings does not vary with whether the firm is releasing good or bad news.

Our analysis is based on the underlying problem that the accuracy of the news generating process is often uncertain and is estimated from the news itself. That this creates a temptation to manipulate the data was first noted in Babbage’s (1830) canonical typology of scientific fraud, which defines “trimming” of data as “in clipping off little bits here and there from those observations which differ most in excess from the mean and in sticking them on to those which are too small” so as to reduce the variance.² We show conditions under which such distortion helps a sender in a Bayesian environment, provide a game theoretic model of such distortion, and test the predictions of the model. Our results on detecting news distortion based on the relative variance of good and bad news fits into the literature on the statistical detection of data manipulation and selective reporting.³ However, in our analysis, distortion need not be actual fraud, but could be accounting adjustments, data cleaning and smoothing of outliers, or even the reallocation of time and other resources to help different projects. Such actions may be legally and contractually permissible, and may even be fully anticipated. Thus, our analysis applies to many situations such as marketing and advocacy, in which biased senders face some truth-telling constraints but have the ability to distort the consistency of news.

We contribute to the literature on “good news and bad news” in two ways. First, we consider a state q and a noisy scalar signal y , where $y = q + \varepsilon$ for some independent ε with symmetric quasiconcave density g . We assume that the prior f for the state q is symmetric and logconcave, and that for variability parameter $\rho' > \rho$, the noise density $g(y - q|\rho')$ is uniformly more variable than $g(y - q|\rho)$. When the news is good, $y > E[q]$, we show that $E[q|y, \rho] > E[q|y, \rho']$, so the posterior estimate is decreasing in the variability of the news, and when the news is bad, $y < E[q]$, we show that $E[q|y, \rho] < E[q|y, \rho']$ so more variability helps rather than hurts.⁴ Second, we apply this result to our case where there is a vector x of news generated by a process of uncertain

²As a result “the average given by the observations of the trimmer is the same, whether they are trimmed or untrimmed.” Trimming, which is our focus, is distinguished from creation of fake data (“forging”) and from selective choice of data (“cooking”).

³For instance, reporting or publication bias based on significance leads to an asymmetric “funnel effect” in sample means and sample sizes (Egger et al., 1997), and manipulation or publication bias leads the observed distribution of significant results to be inconsistent with test power.

⁴The relation of this result to second order stochastic dominance is analyzed in Section 3.4.

variability so that the variability of the news process is estimated from its consistency as captured by the standard deviation s of x . We show when more consistency of this vector of news makes the mean \bar{x} of the news more persuasive. That is, we show conditions under which the observed \bar{x} and s have a distribution $g(\bar{x} - q|s)$ that satisfies our conditions on $g(y - q|\rho)$ where $y = \bar{x}$ and $\rho = s$. Since an increase in any data point x_i has the same effect on \bar{x} , while the effect on s is strictly increasing in the size of x_i , this generates an incentive to selectively distort higher or lower news based on whether the overall news is good or bad.

Uncertainty over the data generating process induces sender preferences over the mean and the variance of the news that differ significantly from traditional mean-variance preferences by an investor over the underlying state (e.g., Meyer, 1987). First, whether the manager prefers more or less variance in the news depends on the overall favorability of the reports, whereas in a standard mean-variance model an investor always prefers less variance. Second, the manager does not always prefer higher news because of the “too good to be true” effect whereby one piece of very good news is treated as so unreliable that the posterior estimate reverts back toward the prior as the news gets better (Dawid, 1973; O’Hagan, 1979; Subramanyan, 1996). We contribute to this latter literature by considering multiple pieces of news from the same data generating process. We show how raising the best news makes not just that news but all the news appear less reliable, which strengthens the too good to be true effect. Conversely, we show that selectively distorting the data so as to focus on shoring up weaker news can allow for an increase in the mean of the news that avoids the too good to be true problem.

We consider both the naive case where the receiver treats distorted news as the true news, and the sophisticated case where the receiver rationally anticipates distortion in a perfect Bayesian equilibrium. Under our assumptions that the mean of the news is fixed and total distortion of the news is limited, we find that the sender’s incentive to selectively distort good or bad news is the same in either case. Hence the model’s predictions apply to both environments. We find that a sophisticated receiver can back out the true news when the news is generally bad, but that there is some loss in information due to partial pooling of reports when the news is generally good.⁵

We analyze ex-post distortion of the sender’s realized news, but our focus on the variability of the news is similar to that of the Bayesian persuasion literature which analyzes ex-ante commitment to an information structure (e.g., Kamenica and Gentzkow, 2011). In particular, the nonlinearity of the posterior mean due to interaction with the prior implies that the sender can benefit from ex-ante commitment to revealing information more or less exactly in different regions. For instance, a firm might adopt an accounting system that is designed to reveal more detailed good news than

⁵Note that if the receiver believes that the sender might be strategic or instead might be an “honest” type who reports the true news, then highly consistent good news (or highly inconsistent bad news) can be suspicious, which mitigates the incentive to distort the news. Stone (2015) considers a related problem in a cheap talk model of binary signals where too many favorable signals is suspicious.

bad news. Since we consider multidimensional news, the variability of the news is important not just ex ante as in the Bayesian persuasion literature, but also ex post once the news has been realized.

The closest approach to ours in the earnings management literature is by Kirschenheiter and Melumad (2002) who consider the incentive to smooth overall firm earnings across time so as to maximize perceived profitability. They focus on the case in which variability is inferred from a single piece of news (rather than from the consistency of multiple pieces of news as in our setting). Their analysis is also complicated by the firm’s need to anticipate uncertain future earnings when deciding whether to overreport or underreport current earnings. By considering distortion across earnings segments rather than time, we can focus on the underlying mechanism that is implicit in their approach – good results are more helpful when they are more consistent, and bad results are less damaging when they are inconsistent. We then show that this same idea applies in a more general statistical environment with multiple pieces of news, analyze the resulting mean-variance news preferences, and apply the idea to a range of other situations.

The remainder of the paper proceeds as follows. Section 2.1 provides a simple example with two projects that shows how the consistency of performance news affects updating. In Section 2.2 we develop statistical results on consistency and variability. In Section 2.3 we show how induced preferences over the mean and variance of the news affects distortion incentives, and in Section 2.4 we consider equilibrium distortion in a sender-receiver game with rational expectations. In Section 3 we consider a range of different applications with mean-variance news preferences, and extend the model to asymmetric news weights. In Section 4 we test for distortion using our main application of segment earnings reports. Section 5 concludes the paper.

2 The model

2.1 Example

A manager has n projects where performance x_i on each is an additive function of the manager’s ability q and a measurement error ε_i , so $x_i = q + \varepsilon_i$. The prior distribution of q is given by the symmetric logconcave density f and the ε_i are i.i.d. normal with zero mean and a s.d. σ_ε with non-degenerate independent distribution H . The manager, who may or may not know the realization of q , knows the realized values of x_i and can shift some resources to selectively boost reported performance \tilde{x}_i on one or more projects at the expense of lower reported performance on other projects. A receiver does not know q or the true x_i but knows the prior distributions and sees the performance reports \tilde{x}_i after the manager may have distorted them. What distortions make the manager look best?

Suppose that the manager wants to maximize the receiver’s expectation of the manager’s ability

q given the priors and the reports. For simplicity, in this example assume the receiver naively believes that the reported x_i are the true x_i . The receiver's posterior estimate $E[q|x]$ is a mixture of the prior and the performance news x with the weight dependent on how accurate the news is believed to be. Since the ε_i are i.i.d. normal, the news $x = (x_1, \dots, x_n)$ can be summarized by the news mean $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, and news variance $s^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n - 1)$. Letting ϕ be the density of the standard normal distribution,

$$\begin{aligned} E[q|\bar{x}, s] &= \frac{\int_{-\infty}^{\infty} qf(q) \int_0^{\infty} \prod_{i=1}^n \phi(x_i - q, \sigma_\varepsilon^2) dH(\sigma_\varepsilon) dq}{\int_{-\infty}^{\infty} f(q) \int_0^{\infty} \prod_{i=1}^n \phi(x_i - q, \sigma_\varepsilon^2) dH(\sigma_\varepsilon) dq} \\ &= \frac{\int_{-\infty}^{\infty} qf(q)g(\bar{x} - q|s) dq}{\int_{-\infty}^{\infty} f(q)g(\bar{x} - q|s) dq} \end{aligned} \quad (1)$$

where

$$g(\bar{x} - q|s) = \int_0^{\infty} \frac{1}{(\sigma_\varepsilon \sqrt{2\pi})^n} e^{-\frac{ns^2 + n(\bar{x} - q)^2}{2\sigma_\varepsilon^2}} dH(\sigma_\varepsilon) \quad (2)$$

captures the impact of the news. For instance, if H has density $h = 1/\sigma_\varepsilon^2$ then $g((\bar{x} - q)/(s/\sqrt{n}))$ is the density of a standard t -distribution with $n - 1$ degrees of freedom.⁶

When the news is more consistent as measured by a lower standard deviation s , the receiver infers that the x_i are less noisy in the sense that there is more weight on lower values of σ_ε in (2). This makes g less variable so the news mean \bar{x} is a more precise signal of q . Hence the posterior estimate of q in (1) puts more weight on the news relative to the prior. If the news is more favorable than the prior, then this greater weight on the news helps the manager.

To see this suppose there are two projects, the prior $f(q)$ for manager ability is normal with mean 0 and s.d. 2, and the prior for the variance of project performance is $h = 1/\sigma_\varepsilon^2$. Suppose that performance on both projects is generally good, $\bar{x} > 0$, but one project is doing better. The manager can shift resources to strengthen the better project at the expense of the worse project which lowers consistency, or instead can shore up the worse project at the expense of the better project which raises consistency. This is seen in Figure 1(a) where the different strategies lead to a more or less variable distribution of $g(\bar{x} - q|s)$. In particular, suppose the manager can distort the true performance $x = (3/2, 7/2)$ to either $x = (1, 4)$ or $x = (2, 3)$. As seen in the figure both reports have the same mean $\bar{x} = 5/2$ but favoring the worse project leads to more consistent performance with $s = \sqrt{1/2}$ rather than $s = \sqrt{9/2}$, and this consistency implies that the mean \bar{x} is a more precise estimate of q .

It would seem that more precise good news should lead to stronger updating of q , and this is seen in the upper right quadrant of Figure 1(b) which shows contour sets for $E[q|x]$. The report $x = (2, 3)$ makes the news seem more reliable than the report $x = (1, 4)$, so the audience puts more weight on the news relative to the prior distribution of q , leading to very different posterior

⁶This Jeffreys prior for H corresponds to the inverse gamma distribution with parameters $\alpha = 1, \beta = 0$.

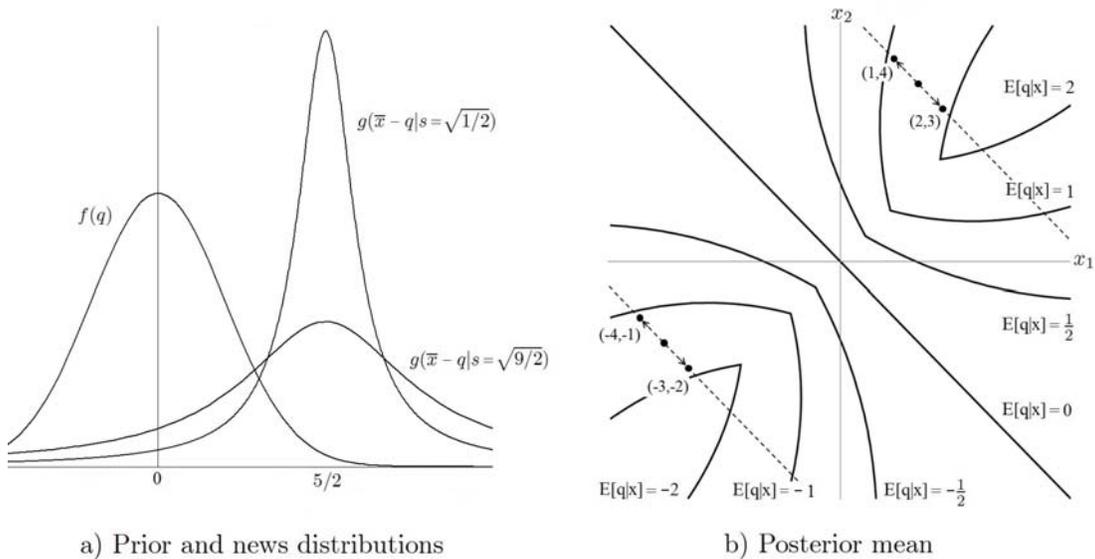


Figure 1: Effects of selective news distortion on consistency and posterior estimate

estimates of the manager’s ability. These effects are reversed if overall performance is worse than the prior. Looking at the lower left quadrant of the figure, when both projects are doing poorly the manager wants to shift resources to the better project and thereby increase the chance the overall bad outcome was due to the noisiness of the environment, so that the receiver relies less on the news and more on the prior distribution of q .

These differential incentives to distort the news imply that the variance of selectively distorted news will be lower (i.e., the news will be more consistent) when it is favorable rather than unfavorable. With enough instances of such situations, distortion can then be detected probabilistically from this predicted difference. In the following we allow for any number of data points, for different priors, for different sender preferences beyond just maximizing the posterior mean, for different pieces of news having different precision, and analyze a sender-receiver game where the receiver rationally anticipates distortion by the sender. We find that the same incentives to distort the consistency of the news remain and the same implications for distortion detection continue to hold.

2.2 Consistency and precision

Continuing the statistical framework introduced in the example, we are interested more generally in when greater consistency of the news as represented by a lower standard deviation s implies the mean \bar{x} is a more precise signal of q , and when a more precise signal of q is more persuasive in that it implies stronger updating in the direction of the signal. Looking back at Figure 1(a) notice that $g(\bar{x} - q | s = \sqrt{1/2})$ is more peaked than $g(\bar{x} - q | s = \sqrt{9/2})$. In fact, the former signal is more precise

in the stronger sense of uniform variability, i.e., for $s' > s$ the ratio $g(\bar{x} - q|s)/g(\bar{x} - q|s')$ is strictly increasing below the mode at $\bar{x} = q$ and strictly decreasing thereafter, so that g is increasingly more spread out for higher values of s on each side of its mode.⁷ The following lemma shows that this holds more generally. All proofs are in the Appendix.

Lemma 1 *Suppose for a given q that $x_i = q + \varepsilon_i$ for $i = 1, \dots, n$ where i.i.d. $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ and σ_ε^2 has independent non-degenerate distribution H . Then the distribution of \bar{x} satisfies $g(\bar{x} - q|s') \succ_{UV} g(\bar{x} - q|s)$ for $s' > s$.*

This result establishes that more consistent news makes the mean of the news a more precise signal of q in the strong sense of uniform variability. We now show generally when ordering of a signal y by uniform variability orders the effect on the posterior estimate for good and bad news.⁸

Lemma 2 *Suppose $g(y - q|\rho)$ is a symmetric quasiconcave density where $g(y - q|\rho') \succ_{UV} g(y - q|\rho)$ for $\rho' > \rho$, and $f(q)$ is independent, symmetric, and logconcave. Then $E[q|y, \rho] \geq E[q|y, \rho']$ if $y \geq E[q]$ and $E[q|y, \rho] \leq E[q|y, \rho']$ if $y \leq E[q]$.*

Regarding these conditions, just to ensure that seemingly good news really is good news, i.e., that news $y > E[q]$ implies $E[q|y] > E[q]$, we need that the prior $f(q)$ is symmetric and that g is symmetric and quasiconcave (Chambers and Healy, 2012).⁹ The additional logconcavity and uniform variability conditions, which are both likelihood ratio conditions under our continuity assumptions,¹⁰ ensure that more precise good news is better. The approach of the proof in the Appendix is to break f into its two parts above and below y . If the mode of f is below y , there is more mass in the lower part of f so an increase in ρ that spreads out g lowers the mean, and the opposite if the mode of f is above y . The likelihood ratio conditions ensure that the densities do not interact to produce counterintuitive results over some regions.

Applying Lemma 1 and letting \bar{x} and s take the roles of y and ρ in Lemma 2 establishes the following result that more consistent news is indeed more persuasive in the sense of moving the posterior estimate in the direction of the mean of the news.

⁷The uniformly more variable order, introduced by Whitt (1985) along with the stronger logconcave order, is satisfied by most standard symmetric distributions including the t-distribution used in Figure 1.

⁸The closest result we know of is by Hautsch, Hess, and Muller (2012), who consider a normal prior and normal news of either high or low precision, with a noisy binary signal of this precision. Most of the related literature considers expectations of convex or concave functions of the state, e.g., SOSD results for concave $u(q)$. As we show in Section 3.4, the effects of news precision on the posterior estimate of $u(q)$ can be ordered for all news only if u is linear.

⁹As Chambers and Healy show, Milgrom's results on when y' is more favorable than y do not rule out $y' > y > E[q]$ but $E[q] > E[q|y'] > E[q|y]$ without additional conditions on f , i.e., two pieces of seemingly good news can be ranked by MLR dominance and yet both be bad. See Finucan (1973) and O'Hagan (1979) for related results.

¹⁰Logconcavity of f is equivalent to $f(q - a) \succ_{MLR} f(q)$ for any $a > 0$. Uniform variability is, for $\rho' > \rho$, equivalent to $g(y - q|\rho) \succ_{MLR} g(y - q|\rho')$ for $y < q$ and $g(y - q|\rho') \succ_{MLR} g(y - q|\rho)$ for $y > q$.

Proposition 1 *Suppose for a given q that $x_i = q + \varepsilon_i$ for $i = 1, \dots, n$ where i.i.d. $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ and σ_ε^2 has independent distribution H , and $f(q)$ is independent, symmetric, and logconcave. Then $\frac{d}{ds}E[q|\bar{x}, s] \leq 0$ if $\bar{x} \geq E[q]$ and $\frac{d}{ds}E[q|\bar{x}, s] \geq 0$ if $\bar{x} \leq E[q]$.*

We use the case of a biased sender who wishes to maximize $E[q|\bar{x}, s]$ as our main application of the general mean-variance news preferences introduced below.

2.3 Mean-variance news preferences

To analyze the sender’s distortion incentives, it is helpful to think more generally of sender preferences over the news. Normality of ε_i implies \bar{x} and s are sufficient statistics for the information in the news x , so any information-based preferences must be a function of these statistics rather than the fine details of the news. We will consider general mean-variance news preferences $U : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ by a sender such that, denoting partial derivatives by subscripts,

$$\begin{aligned} U_s(\bar{x}, s) &> 0 \text{ for } \bar{x} < E[q] \\ U_s(\bar{x}, s) &= 0 \text{ for } \bar{x} = E[q] \\ U_s(\bar{x}, s) &< 0 \text{ for } \bar{x} > E[q] \end{aligned} \tag{3}$$

for all $(\bar{x}, s) \in \mathbb{R} \times \mathbb{R}_+$.¹¹ As established above in Proposition 1, U satisfies these conditions if $U = E[q|\bar{x}, s]$, and clearly the conditions are also satisfied if U is any increasing function of $E[q|\bar{x}, s]$. In Section 3 we provide other situations where U satisfies these conditions. As analyzed further in Section 3.4, these are sender preferences over the mean and variance (or standard deviation) of the news x , not investor or other receiver preferences over the mean and variance of the state q as in traditional mean-variance models (e.g., Meyer, 1987). We do not restrict the sign of $U_{\bar{x}}(\bar{x}, s)$ and in Section 3.2 we consider the issue of “too good to be true” news preferences where $U_{\bar{x}}(\bar{x}, s)$ is not monotonic.

To see the incentive to selectively distort the news, note that

$$\frac{d\bar{x}}{dx_j} = \frac{1}{n}, \quad \frac{ds}{dx_j} = \frac{x_j - \bar{x}}{(n-1)s} \tag{4}$$

so every data point has the same effect on \bar{x} , but the effect on the variance is increasing in the size of x_j relative to \bar{x} . Since a lower s helps when $\bar{x} > E[q]$ and hurts when $\bar{x} < E[q]$, the marginal gain is higher from increasing lower data points in the former case, and from increasing higher data points in the latter case. In particular if the best news is exaggerated this increases \bar{x} and also increases s , so the effects on the posterior estimate counteract each other if $\bar{x} > E[q]$ but reinforce each other if $\bar{x} < E[q]$. And if the worst news is exaggerated this increases \bar{x} but also decreases s ,

¹¹We focus on preferences over summary statistics of multiple signals, but the analysis also applies to preferences over one signal with known variability, $U(y, \rho)$, when the variability parameter ρ can be directly influenced.

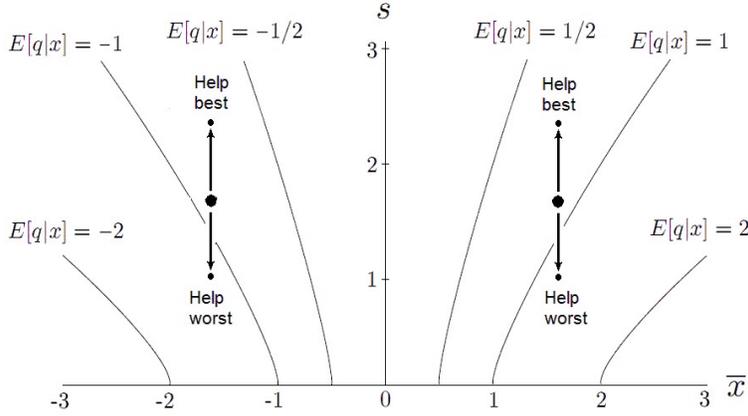


Figure 2: Segment earnings management, $U = E[q|\bar{x}, s]$

so the effects on the posterior estimate reinforce each other if $\bar{x} > E[q]$ but counteract each other if $\bar{x} < E[q]$.

Proposition 2 For U satisfying (3) and $x_i < x_j$, $\frac{d}{dx_i}U(\bar{x}, s) \geq \frac{d}{dx_j}U(\bar{x}, s)$ if $\bar{x} \geq E[q]$, and $\frac{d}{dx_i}U(\bar{x}, s) \leq \frac{d}{dx_j}U(\bar{x}, s)$ if $\bar{x} \leq E[q]$.

Figure 2 shows the posterior mean as a function of \bar{x} and s for $n = 4$ and a prior of zero. Applied to our main application of segment earnings management described further in Section 4, if management wants to increase the posterior estimate of their ability q based on the performance news of different units x_i , the figure represents indifference curves in (\bar{x}, s) space. First suppose that the true news is generally favorable, $x = (0, 1, 2, 3)$ as in the right side of the figure. Under the constant mean constraint suppose the best news is exaggerated from $x_4 = 3$ to $x_4 = 4$ and the worst news lowered from $x_1 = 0$ to $x_1 = -1$. The mean stays the same but s increases so the posterior estimate falls as seen from “Help best”. If instead, the worst news is raised from $x_1 = 0$ to $x_1 = 1$ and best news lowered from $x_4 = 3$ to $x_4 = 2$ then s falls and the posterior mean rises substantially as seen from “Help worst”. If the true news is generally unfavorable, $x = (-3, -2, -1, 0)$ as in the left side of the figure, then these effects are reversed.

2.4 Optimal and equilibrium distortion

So far we have analyzed the sender’s marginal incentive to distort different news, but the actual distortion choice depends on what constraints the sender faces, and may also depend on whether or not the receiver anticipates distortion. We consider both optimal distortion when the receiver is “naive” and does not anticipate distortion, and also equilibrium distortion when the receiver

is “sophisticated” and rationally anticipates distortion. We will show that the simple distortion strategy when the receiver is naive is also an equilibrium strategy when the receiver is sophisticated.

Let $\tilde{x}(x)$ be the sender’s pure strategy of reporting \tilde{x} based on true news x . The receiver estimates the posterior distribution of q given her priors f and H , the observed news \tilde{x} , and her beliefs that map \tilde{x} to the set of probability distributions over \mathbb{R}^n . In the naive receiver case the receiver does not anticipate distortion so receiver beliefs put all weight on $x = \tilde{x}$. In the sophisticated receiver case the receiver’s beliefs are consistent with the sender’s strategy along the equilibrium path. Therefore if $\tilde{x}(x)$ is one-to-one the receiver puts all weight on $x = \tilde{x}^{-1}(\tilde{x}(x))$. If not the receiver weights the distribution of x according to $\tilde{x}(x)$ and Bayes’ rule given f and H . If the sender makes a report that is off the equilibrium path the beliefs put all weight on whichever type is willing to deviate for the largest set of rationalizable payoffs, i.e., we impose the standard D1 refinement (Cho and Kreps, 1987).

We assume the sender’s preferences $U(\bar{x}, s)$ satisfy (3) so, from the sender’s perspective, the receiver’s beliefs about the distribution of the true x can be summarized by the receiver’s belief about the distribution of s which we denote by $\mu(s|\tilde{x})$. The sender maximizes her expected utility

$$\int_0^\infty U(\bar{x}, s) d\mu(s|\tilde{x}), \quad (5)$$

subject to a constant mean constraint and a maximum distortion constraint,

$$\sum_i \tilde{x}_i - x_i = 0 \text{ and } \sum_i |\tilde{x}_i - x_i| \leq d, \quad (6)$$

where $d > 0$ is the maximum total distortion across the news.

First consider the naive receiver case. When the news is generally unfavorable, $\bar{x} < E[q]$, the sender wants to increase s as much as possible. Supposing $n = 2$ and looking back at the bottom left quadrant of Figure 1(b) the sender will want to make the smaller x_i even smaller. So starting at any point on the dashed line that represents a constant mean \bar{x} , the sender will want to move further away from the center where $x_1 = x_2$ and toward either edge by taking $d/2$ from the smaller x_i and adding it to the larger x_i . If $n > 2$, this same logic applies. Relabeling the news such that $x_1 \leq \dots \leq x_n$, from (4) for any s the largest increase in s occurs when x_1 is decreased and x_n is increased, so the sender simply decreases x_1 by $d/2$ and increases x_n by $d/2$, which satisfies (6).

When the news is generally favorable, $\bar{x} > E[q]$, the sender wants to decrease s as much as possible. From the upper right quadrant of Figure 1(b) the sender wants to move inward along the dashed line toward the center where $x_1 = x_2$. Therefore if $x_2 - x_1 \geq d$ the sender increases x_1 by $d/2$ and decreases x_2 by $d/2$, and if not the sender just sets $x_1 = x_2$ without having to exhaust the total distortion budget. If $n > 2$, the sender will start by squeezing in the most extreme news x_1 and x_n . As they are moved inward they might bump into other data points, which then need to be moved in jointly from either side until the budget of $d/2$ distortion on either side, which maintains

the prior mean, is exhausted. If all the data starts out sufficiently close, the data is completely squeezed to the mean \bar{x} before the budget is exhausted. This strategy is specified in Proposition 3 below.

Now consider the sophisticated receiver case and suppose that the sender follows the same strategy as in the naive receiver case. Suppose $n = 2$, though the proof in the Appendix is more general. When $\bar{x} < E[q]$, reports on the equilibrium path can be inverted by a sophisticated receiver to back out the true x , but some reports are not on the equilibrium path. Looking again at Figure 1(b) if $d = 1$ then for any x such that $\bar{x} = -5/2$, a report on the dashed line between $(-3, -2)$ and $(-2, -3)$ should never be observed. As we show in the proof of Proposition 3, in such cases it is always the “worst type” $x_1 = x_2$ with the lowest s that is (weakly) willing to deviate to such an off the path report for the largest range of rationalizable payoffs. Therefore, by the D1 refinement, the receiver should assume that such a deviation was done by this type. Given such beliefs, even the worst type gains nothing from deviation. When $\bar{x} > E[q]$, if the reports for the projects differ, a sophisticated receiver can again invert the equilibrium strategy and back out the true x , but otherwise there is some pooling. In the example of Figure 1(b), if $d = 1$ then for all x between $(2, 3)$ and $(3, 2)$ the sender will report $(5/2, 5/2)$ so the receiver cannot invert the reports. In this case the receiver will form a belief over the true x that induces a distribution over s where s is always smaller than when the receiver is thought to be outside of the region between $(2, 3)$ and $(3, 2)$. Since the sender prefers a lower s and any other report will lead the receiver to infer the sender is outside this region with a higher s , the sender has no incentive to deviate.

Following this logic the optimal strategy when the receiver is naive is also an equilibrium strategy when the receiver is sophisticated, leading to the following proposition. Since the equilibrium is fully separating for $\bar{x} \leq E[q]$, the receiver correctly “backs out” the true values by discounting the reported values according to the equilibrium strategy. However the equilibrium is partially pooling for $\bar{x} \geq E[q]$, so some information is lost even though receivers correctly anticipate distortion.

Proposition 3 (i) Assume the receiver is naive. If $\bar{x} < E[q]$ then the sender’s optimal strategy is $\tilde{x}_1 = x_1 - d/2, \tilde{x}_n = x_n + d/2$, and $\tilde{x}_i = x_i$ for $i \neq 1, n$. If $\bar{x} > E[q]$ then (a) if $\sum_i |x_i - \bar{x}| \leq d$ then $\tilde{x}_i = \bar{x}$ for all i ; (b) if not then $\tilde{x}_i = \vec{x}$ for $i \leq a$, $\tilde{x}_i = \overleftarrow{x}$ for $i \geq b$, and $\tilde{x}_i = x_i$ for $a < i < b$ where \vec{x} solves $\sum_{i=1}^a (\vec{x} - x_i) = d/2$ subject to $x_a \leq \vec{x}$, and \overleftarrow{x} solves $\sum_{i=b}^n (x_i - \overleftarrow{x}) = d/2$ subject to $x_b \geq \overleftarrow{x}$. (ii) Assume the receiver is sophisticated. Then the sender’s strategy in (i) is a perfect Bayesian equilibrium.

This distortion strategy leads to a higher variance for \tilde{x} than for x when $\bar{x} < E[q]$, and lower variance for \tilde{x} than for x when $\bar{x} > E[q]$. By our symmetry assumptions on the prior density of q and on the news given q , the expected standard deviation of the true x is the same for any \bar{x} equidistant from the prior on either side. Therefore the reported standard deviation for \tilde{x} should

on average be higher below the prior than above the prior.¹²

Proposition 4 *The distortion strategy in Proposition 3 implies that, in expectation, $s(\bar{x})$ is higher when $\bar{x} < E[q]$ than when $\bar{x} > E[q]$.*

This result is our main testable implication of the model.

3 Applications and extensions

We now develop the mean-variance news preferences idea further by considering different applications where the sender's preferences $U(\bar{x}, s)$ have the general fan-shape of Figure 2 where $U_s > 0$ for \bar{x} below the prior and $U_s < 0$ for \bar{x} above the prior.¹³ An important counterexample to these preferences is where risk aversion may lead to $U_s < 0$ for all \bar{x} , as shown in Section 3.4. An extension to weighted means and weighted standard deviations is given in Section 3.5. This extension is used in our test of earnings management in Section 4. For each case we focus on the underlying distortion incentives, though the analysis can be extended in the same manner as above to equilibrium distortion.

3.1 Posterior probability

Rather than maximizing their estimated skill, a manager might want to maximize the estimated probability that they are competent so as to attain a promotion or avoid a demotion (Chevalier and Ellison, 1999). This can be modeled as maximizing the posterior probability that q exceeds some level. In the Appendix we establish Lemma 3 which is an equivalent to Lemma 2 for the posterior probability $F(q|y, \rho)$ rather than the posterior estimate $E[q|y, \rho]$. Letting $y = \bar{x}$ and $\rho = s$ then gives the following result that implies $U = \Pr[q > E[q]|\bar{x}, s]$ has the properties of (3). Hence the predictions for selective news distortion are the same as those for maximizing estimated skill.

Result 1 *For $\bar{x} \geq E[q]$, $\frac{d\Pr[q > a|\bar{x}, s]}{ds} \leq 0$ for all $a \leq \bar{x}$, and for $\bar{x} \leq E[q]$, $\frac{d\Pr[q > a|\bar{x}, s]}{ds} \geq 0$ for all $a \geq \bar{x}$.*

The introductory examples of Figures 1 and 2 showed how a manager would have an incentive to shift resources to help better or worse performing projects to maximize the posterior estimate of their ability. Figure 3(a) shows the same case as Figure 2 except the manager wants to maximize

¹²In practice there might be other constraints such as not all news can be distorted, or distortion might be costly, with some distortions more costly than others. This same prediction applies for the naive receiver case since the sender never benefits from a higher s when news is good or a lower s when news is bad. Whether the prediction extends generally to sophisticated receivers has not yet been verified.

¹³In some situations U might depend on details of the performance news rather than on \bar{x} and s , e.g., individual x_i might affect U directly because a manager's compensation depends on how individual units perform.

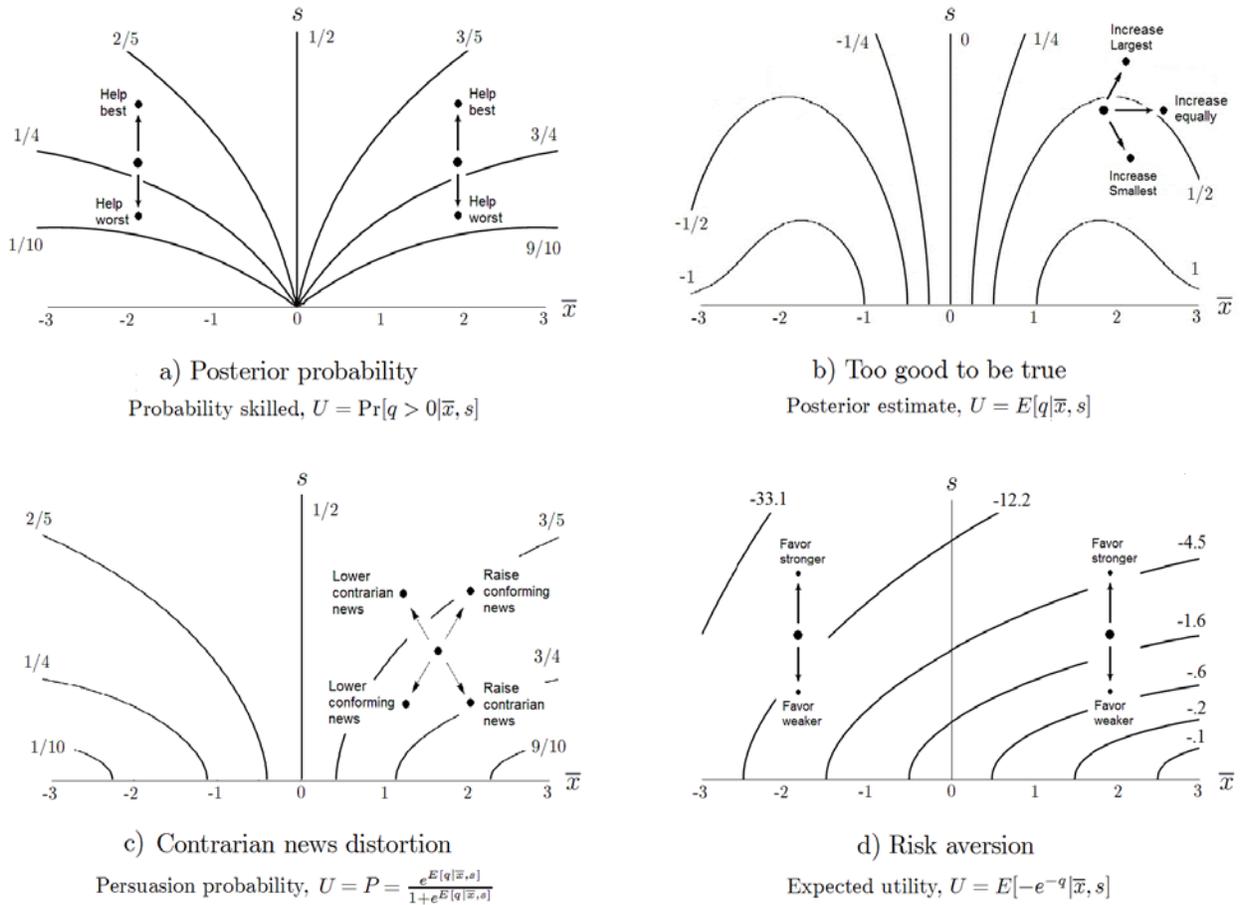


Figure 3: Applications with mean-variance news preferences $U(\bar{x}, s)$

the probability that his skill q is above the prior which is normalized to zero. As the above result indicates $\Pr[q > 0 | \bar{x}, s]$ is decreasing in s if $\bar{x} > 0$ and increasing in s if $\bar{x} < 0$, so the manager prefers to help the worst project in good times and the best project in bad times.¹⁴

3.2 Too good to be true

Can news be so good that it is no longer credible? Dawid (1973) and O'Hagan (1979) show that an increase in a single piece of news y can be “too good to be true” in that it might decrease $E[q|y]$, and in particular that $\lim_{y \rightarrow \infty} E[q|y] = E[q]$ if f has thinner tails than g , including the case where f is normal and g is the t -distribution. Subramanian (1996) shows that if f is normal

¹⁴We have assumed that the prior f is informative, but if f is uninformative and $h = 1/\sigma_\varepsilon^2$ then the posterior distribution of q is the t -distribution with $n - 1$ degrees of freedom, so the probability that $q < 0$ is given by $T_{n-1}(\bar{x}/s)$ and the indifference curves in the figure are linear. Hence this result generalizes the t -distribution case where an increase in s helps or hurts depending on the sign of \bar{x} .

and g is normal with uncertain variance, which includes the t -distribution case, that as $y > E[q]$ increases $E[q|y]$ is first increasing and then decreasing.

Applied to our environment with $y = \bar{x}$, these standard results imply that as \bar{x} increases with a fixed s the news can eventually become too good to be true.¹⁵ This effect is aggravated or mitigated when an individual x_i changes, depending on its position relative to the mean. An increase in $x_i > \bar{x}$ not only raises \bar{x} but also increases s , so in the region where $\frac{d}{d\bar{x}}E[q|\bar{x}, s] < 0$ an increase in x_i that raises s has the double effect of decreasing $E[q|\bar{x}, s]$ via both \bar{x} and s . For $x_i < \bar{x}$ the two effects counteract each other so the too good to be true effect is mitigated and potentially avoided.

Based on these differential effects, it is possible to increase \bar{x} and avoid the too good to be true problem entirely through selective distortion. Suppose the total distortion constraint is $\sum_i |\tilde{x}_i - x_i| \leq d$ for some given $d > 0$. If the sender reports $\tilde{x}_1 = x_1 + (d/2 + \varepsilon)$ and $\tilde{x}_n = x_n - (d/2 - \varepsilon)$ for $0 \leq \varepsilon \leq d/2$ then s falls discontinuously for any such ε while \bar{x} increases continuously as ε increases from zero. Therefore by the continuity of $E[q|\bar{x}, s]$ in s and \bar{x} , if $\frac{d}{ds}E[q|\bar{x}, s] < 0$ it is always possible to choose an ε that raises $E[q|\bar{x}, s]$ even in the range where $\frac{d}{d\bar{x}}E[q|\bar{x}, s] < 0$. These two results, and the equivalents for unfavorable news, are stated in the following proposition.

Result 2 (i) If $\frac{d}{d\bar{x}}E[q|\bar{x}, s] \geq 0$ then $\frac{d}{dx_i}E[q|\bar{x}, s] > 0$ for all $x_i < \bar{x}$, and if $\frac{d}{d\bar{x}}E[q|\bar{x}, s] \leq 0$ then $\frac{d}{dx_i}E[q|\bar{x}, s] < 0$ for all $x_i > \bar{x}$. (ii) For any $d > 0$ and x , there exists a distortion \tilde{x} such that $\tilde{x} > \bar{x}$ and $E[q|\tilde{x}, \tilde{s}] > E[q|\bar{x}, s]$, and an alternative distortion \tilde{x}' such that $\tilde{x}' < \bar{x}$ and $E[q|\tilde{x}', \tilde{s}'] < E[q|\bar{x}, s]$.

This result is shown in Figure 3(b) where the environment is the same as Figure 2 except the prior has lower variance so that as \bar{x} increases and becomes less reliable the posterior $E[q|\bar{x}, s]$ converges more quickly to the prior in the pictured range. As seen on the right side of the figure, increasing all the x_i keeps s the same and $E[q|\bar{x}, s]$ falls as the data becomes less believable relative to the prior, but if x is selectively distorted with increases in the smaller data points this is avoided. By the same logic even in the range of “too bad to be true” on the left side of the figure, it is possible to reduce $E[q|\bar{x}, s]$ further by selective reduction of x that focuses on reducing s by reducing the largest data points.

The same analysis extends to posterior probabilities. Dawid (1973) shows that not only does the mean revert to the prior when f has thinner tails than g , but the entire posterior distribution reverts to the prior distribution, so $\lim_{\bar{x} \rightarrow \infty} \Pr[q > a|\bar{x}, s] = 1 - F(a)$ for any a .¹⁶ The same selective

¹⁵If H is such that g is the t -distribution then news must eventually be too good to be true as O’Hagan shows. But if g is instead logconcave, which is also possible for mixtures of normals, then higher news must always be better by an application of Milgrom’s good news result.

¹⁶The basic model of Student (1908) already incorporates a version of the too good to be true idea due to its use of the same data to estimate both the mean and the standard deviation. In particular, letting $t(x)$ be the t -value, direct calculations show that $\lim_{x_i \rightarrow \infty} t(x) = 1$, so if $\bar{x} > 0$ and s is small enough that $t(x) > 1$, raising any x_i eventually undermines the reliability of all the data so much that significance decreases. Note that if $h = 1/\sigma_\varepsilon^2$ and, counter to

distortion strategy used for the posterior mean above can then also be used to avoid the too good to be true problem for the posterior probability.

3.3 Contrarian news distortion

The literature on news bias has focused on distortions that push a scalar news variable in the source’s favored direction at some reputational or other cost (e.g., Gentzkow and Shapiro, 2006). If there are multiple pieces of news, then the consistency of the news also becomes a factor that the source can manipulate.¹⁷ For instance, opponents of action on climate change are claimed to exaggerate evidence against the scientific consensus as part of a strategy of “seeding doubt” (e.g., Oreskes and Conway, 2010),¹⁸ while proponents are claimed to make the consensus appear stronger by downplaying opposing evidence. These are not the only distortion strategies available – opponents could instead focus on downplaying evidence for the consensus, while proponents could instead focus on exaggerating “scare stories” that are in the direction of the consensus.

Applying our model to such situations, define news as contrarian relative to other news if it is on the prior’s side of the mean of the news and conforming otherwise. That is, for $\bar{x} > E[q]$ we say news x_i is *contrarian* if $x_i < \bar{x}$ and *conforming* if $x_i > \bar{x}$, and for $\bar{x} < E[q]$ we say news x_i is *contrarian* if $x_i > \bar{x}$ and *conforming* if $x_i < \bar{x}$.¹⁹ For $\bar{x} > E[q]$ distorting contrarian news $x_i < \bar{x}$ downward increases s and also lowers \bar{x} , while distorting contrarian news upward decreases s and also raises \bar{x} . So both sides – those who want a higher estimate and who want a lower estimate – get a double effect from focusing on distorting contrarian news in their favored direction. And for $\bar{x} < E[q]$ distorting contrarian news $x_i > \bar{x}$ upward or downward also generates a double effect for either side. In contrast, distorting conforming news always creates a trade-off of either making the mean of the news more favorable but the consistency less favorable, or making the mean of the news less favorable but the consistency more favorable. If $P(\bar{x}, s)$ is the probability that the audience is persuaded to one side, which could be a function of $E[q|\bar{x}, s]$ or, as in Section 3.1, of $\Pr[q > E[q]|\bar{x}, s]$, we have the following result by application of Proposition 2.

Result 3 *Suppose the persuasion probability $P(\bar{x}, s)$ satisfies (3). For either side of a debate, $U = P$ or $U = 1 - P$, distorting contrarian news is more effective than distorting conforming news.*

our assumption, f is uninformative then $f(q|x)$ follows the t-distribution. So an informative f strengthens the too good to be true effect by making it hold for all $\bar{x} > 0$.

¹⁷Chakraborty and Harbaugh (2010) consider multidimensional news but without uncertainty over the news generating process. Their focus is on the implicit opportunity cost of pushing one dimension versus another.

¹⁸Recently released internal memos from Exxon indicate an explicit strategy to “emphasize the uncertainty in scientific conclusions” regarding climate change. NYT 11/7/2015.

¹⁹Recall that “news” in our model comes from the same data generating process so that the credibility of all the news rises and falls with its consistency. Data from different processes is modeled as contributing to the prior. This makes the question of whether a given analysis really follows standard methods, and hence has the spillover effects we analyze, of particular importance and hence a likely area of controversy.

Returning to our career concerns environment, suppose that some investors oppose current management and others support management, and the evidence is the performance of different units. Then, if performance is generally good, opponents gain most from “spinning” the performance of the weakest units as particularly bad, while proponents gain most from spinning the performance of such units as less bad than they appear, and the opposite if performance is generally bad. For instance, following a standard random utility model, suppose the probability that the board is persuaded to retain current management is $P = e^{E[q|\bar{x},s]} / (1 + e^{E[q|\bar{x},s]})$, as shown in Figure 3(c). Since the news mean is above the prior in the figure, opponents want to make contrarian evidence more damaging and opponents want to make it less damaging, and neither side benefits much from distorting conforming news. Given the definition of contrarian news, the same holds if the news mean is below the prior. In general the model implies that debates are likely to focus on the exact meaning of the most contrarian evidence, and such evidence is a good place to look for signs of distortion.

3.4 Risk aversion

As shown above, the mean-variance news model satisfies (3) when sender utility is a monotonic function of the posterior estimate of q or of the posterior probability that q exceeds the prior. If the sender cares more generally about the distribution of q then U might not satisfy (3). For instance, in an asset pricing context, if q is the true profitability of the firm then undiversified investors may be particularly concerned about the chance of low realizations of q . Given such risk aversion by investors, the firm will be concerned with how the variance of news about q affects the investor’s expected utility rather than just the expected value of q .

To see this, suppose the sender is a firm, q is the firm’s true value, and the receiver is a risk averse investor with utility $u(q)$. The investor’s valuation of the asset, and the payoff to the firm, are increasing in the investor’s expected utility $E[u(q)|\bar{x}, s]$. Following the general argument in Meyer (1987), since \bar{x} and s are sufficient statistics for x , and since we are taking the prior $f(q)$ as given, the investor must have a “mean-variance” utility function in our environment in the sense that no other information matters. However, due to the interaction between the news and the prior, risk aversion by the investor need not imply that the sender prefers lower news variance. For good news, lower variance is doubly helpful for the sender since it raises the posterior mean and reduces the chance of very bad outcomes. But for bad news, lower variance lowers the posterior mean even if it reduces the chance of very bad outcomes, so higher variance may be preferred. Hence this situation is distinct from standard mean-variance utility models that focus on the investor’s preference regarding variance in q rather than on the firm’s preference regarding variance in news about q .

In particular, Lemma 4 in the Appendix establishes that for $\rho' > \rho$, $\int_{-\infty}^a F(q|\rho', y) dq >$

$\int_{-\infty}^a F(q|\rho, y)dq$ for all a if $y > E[q]$, and $\int_{-\infty}^a F(q|\rho', y)dq < \int_{-\infty}^a F(q|\rho, y)dq$ for all a if $y < E[q]$. The former result establishes that $F(q|\rho, y) \succ_{SO SD} F(q|\rho', y)$ if $y > E[q]$ which, together with Lemma 2 and Proposition 1, implies part (i) of the following for a risk averse sender. The latter result implies part (ii) for a risk-loving sender. Together parts (i) and (ii) imply part (iii), as already established directly in Proposition 1.

Result 4 *Suppose U is an increasing function of $E[u(q)|\bar{x}, s]$ where u is increasing. (i) For u concave $U_s \leq 0$ if $\bar{x} \geq E[q]$; (ii) for u convex $U_s \geq 0$ if $\bar{x} \leq E[q]$; and (iii) for u linear $U_s \leq 0$ if $\bar{x} \geq E[q]$ and $U_s \geq 0$ if $\bar{x} \leq E[q]$.*

Figure 3(d) shows the concave u case for constant absolute risk aversion, $u = -e^{-q}$. In the realm of good news, smaller s both increases $E[q|\bar{x}, s]$ and lowers uncertainty by making $f(q|\bar{x}, s)$ less spread out so the gains from reducing s are accentuated. In the realm of bad news, a smaller s decreases $E[q|\bar{x}, s]$ but reduces uncertainty so much that the sender is better off. So, in this example, risk aversion not only adds to the negative information effect of a higher s for good news, it is strong enough to reverse the positive information effect for bad news so that $U_s < 0$ for both good and bad news. Similarly, risk lovingness not only adds to the positive information effect of a higher s for bad news, it can be strong enough to reverse the negative information effect for good news so that $U_s > 0$ for both good and bad news.

3.5 Asymmetric news weights

So far we have assumed a symmetric environment where each project contributes equally to estimating the posterior value of q . If projects vary predictably in size, the noise terms for each project might have different variances rather than be identically distributed. In this extension we show that a weighted mean-variance model is the same statistically as the symmetric model with appropriate substitution of weighted parameters. Moreover, under natural assumptions that fit environments including our segment earnings application, the strategic implications are also the same.

Suppose that $\varepsilon_i \sim N(0, \sigma_\varepsilon^2/w_i)$ for each project where the weights $\sum_i w_i = W$ are known and σ_ε^2 is distributed according to H as before. Defining the weighted mean and standard deviation as $\bar{x}_w = \sum_{i=1}^n \frac{w_i}{W} x_i$ and $s_w = \left(\sum_{i=1}^n \frac{nw_i}{W} (x_i - \bar{x}_w)^2 / (n-1) \right)^{1/2}$, it is straightforward to verify that $g(\bar{x}_w - q, s_w)$ is the same as (2) with \bar{x}_w and s_w^2 in place of \bar{x} and s^2 , so the same result from Lemma 2 for uniform variability holds with these weighted sufficient statistics. Therefore in the environments above where we show that $U(\bar{x}, s)$ satisfies (3) then $U(\bar{x}_w, s_w)$ also satisfies the equivalent of (3).

Regarding distortion incentives, note that $\frac{d}{dx_i} \bar{x}_w = w_i/W$ so distortion of different news need not have the same effect on the weighted mean, and $\frac{d}{dx_i} s_w = n \frac{w_i}{W} (x_i - \frac{w_i}{W} \bar{x}_w) / ((n-1)s)$ so the ordering effects on the standard deviation are also sensitive to the weights. As we now show the

distortion incentives of the main model are restored if the effects of distortion on each piece of news are also weighted by the inverse of the variances. We show this weighting occurs naturally in our segment earnings environment when segment performance is segment earnings e_i divided by segment assets a_i , so Return on Assets (ROA) is $x_i = e_i/a_i$ and the firm distorts e_i .

Suppose there are n different earnings segments and different size segments are captured by each segment i being composed of m_i different subsegments of equal size. Let subsegment earnings of the k th subsegment of the i th segment be $e_{ik} = q + \varepsilon_k$ for $i = 1, \dots, n$ and $k = 1, \dots, m_i$ where ε_{ik} is i.i.d. normal with mean 0 and s.d. σ_ε . By normality, $Var[e_i] = Var[\sum_{k=1}^{m_i} e_{ik}] = m_i \sigma_\varepsilon^2$. Suppose that each subsegment has identical assets normalized to 1, so that total assets for each segment are $a_i = m_i$ and total firm assets are $A = \sum_{i=1}^n a_i$. Then the variance of segment ROA performance is $Var[x_i] = Var[e_i/a_i] = \frac{1}{a_i^2} Var[\sum_{k=1}^{m_i} e_{ik}] = \frac{1}{a_i^2} a_i \sigma_\varepsilon^2 = \frac{1}{a_i} \sigma_\varepsilon^2$, which is inversely proportional to segment size.

Using weights $w_i = \frac{a_i}{A}$ the weighted mean and standard deviation of the firm's performance are then

$$\bar{x}_w = \sum_{i=1}^n \frac{a_i}{A} \frac{e_i}{a_i} \text{ and } s_w = \left(\sum_{i=1}^n n \frac{a_i}{A} \left(\frac{e_i}{a_i} - \sum_{i=1}^n \frac{a_i}{A} \frac{e_i}{a_i} \right)^2 / (n-1) \right)^{1/2} \quad (7)$$

so

$$\frac{d}{de_i} \bar{x}_w = \frac{1}{A} \text{ and } \frac{d}{de_i} s_w = \frac{n}{A} \frac{(x_i - \bar{x}_w)}{(n-1)s}. \quad (8)$$

It is easier to manipulate the ROA of a smaller segment due to the smaller denominator, but smaller segments have less weight, so shifting earnings from one segment to another does not affect the weighted mean $\bar{x}_w = \sum_{i=1}^n \frac{a_i}{A} \frac{e_i}{a_i}$, which equals the firm's overall ROA $\sum_{i=1}^n \frac{e_i}{A}$. Since the ROA of larger segments has lower variance, any distortion of their ROA has a bigger impact on the posterior estimation of q , but any change in earnings has a relatively smaller impact on ROA for a bigger segment, so again the effects cancel out and the relative gains from distorting segment performance continue to hold. These effects are seen in (8), which reduces to the symmetric case of (4) for $a_i = 1$ and $A = n$.

Therefore the above case where $w_i = a_i/A$ and distortion is of e_i where $x_i = e_i/a_i$ is an example of the following more general result.

Result 5 *If $\varepsilon_i \sim N(0, \sigma_\varepsilon^2/w_i)$ where w_i is known then the above asymmetric model generates the same restrictions on $U(\bar{x}_w, s_w)$ as the symmetric model does for $U(\bar{x}, s)$, and also generates the same relative distortion incentives for the sender if distortion ability is inversely proportional to w_i .*

Since earnings segments often vary substantially in size, we use this weighted model for our empirical analysis of segment earnings distortion, and in particular we consider how changes in cost allocations across segments affect earnings e_i and hence affect segment ROA $x_i = e_i/a_i$.

4 Empirical Test – Earnings Management Across Segments

We now turn to an empirical test of the theory. In earnings reports, managers of US public firms have discretion in how to attribute total firm earnings to business segments operating in different industries. The reporting of earnings across segments is therefore one aspect of “earnings management,” whereby a managers tries to influence the short-run appearance of the firm’s profitability, or of her own managerial ability, by adjusting reported earnings. The shifting of total firm earnings across time is a well-studied topic in the theoretical and empirical literature (e.g., Stein, 1989; Kirschenheiter and Melamud, 1992), but the shifting of earnings across segments has not received as much attention. In particular, the strategy of influencing the consistency of earnings across segments has not, to our knowledge, been analyzed theoretically or empirically.²⁰

4.1 Overview of Empirical Setting

Segment earnings (also known as segment profits or EBIT) are a key piece of information used by boards and investors when evaluating firm performance and managerial quality. In a survey of 140 star analysts, Epstein and Palepu (1999) find that a plurality of financial analysts consider segment performance to be the most useful disclosure item for investment decisions, ahead of the three main firm-level financial statements (statement of cash flows, income statement, and balance sheet).

Under regulation SFAS No. 14 (1976—1997) and SFAS No. 131 (1997—present), managers exercise substantial discretion over the reporting of segment earnings.²¹ Firms are allowed to report earnings based upon how management internally evaluated the operating performance of its business units. In particular, segment earnings are approximately equal to sales minus costs, where costs consist of costs of goods sold; selling, general and administrative expenses; and depreciation, depletion, and amortization. As shown in Givoly et al. (1999), the ability to distort segment earnings is primarily due to the manager’s discretion over the allocation of shared costs to different segments.²² This discretion over cost allocations approximately matches our model of strategic dis-

²⁰The literature has considered issues such as withholding segment earnings information for proprietary reasons (Berger and Hann, 2007), the effects of transfer pricing across geographic segments on taxes (Jacob, 1996), and the channeling of earnings to segments with better growth prospects (You, 2014). Our analysis of the distortion of allocations across segments is also related to the literature on the “dark side of internal capital markets,” e.g., Scharfstein and Stein (2000).

²¹Prior to SFAS No. 131, many firms did not report segment-level performance because the segments were considered to be in related lines of business. SFAS No. 131 increased the prevalence of segment reporting by requiring that disaggregated information be provided based on how management internally evaluated the operating performance of its business units.

²²GE’s 2015 10Q statement offers an example of managerial discretion over segment earnings: “Segment profit is determined based on internal performance measures used by the CEO ... the CEO may exclude matters such as charges for restructuring; rationalization and other similar expenses; acquisition costs and other related charges; technology and product development costs; certain gains and losses from acquisitions or dispositions; and litigation settlements or other charges ... Segment profit excludes or includes interest and other financial charges and income taxes according to how a particular segment’s management is measured ... corporate costs, such as shared services, employee benefits and information technology are allocated to our segments based on usage.”

tortion of news under a fixed mean and total distortion constraint. We assume that total segment earnings are approximately fixed in a period and managers have a limited amount of discretionary costs that can be flexibly allocated across segments to alter the consistency of segment earnings. Our theory predicts that managers will distort segment earnings to appear more consistent when overall firm news is good relative to expectations. When firm news is bad, managers will distort segment earnings to appear less consistent.²³

By focusing on segment earnings management within a time period rather than firm-level earnings management over time, we are able to bypass an important dynamic consideration for the management of earnings over time. The manager can only increase total firm-level earnings in the current period by borrowing from the future, which limits the manager’s ability to report high earnings again next period. In contrast, distortion of the consistency of earnings across segments in the current period does not directly constrain the manager’s ability to distort segment earnings again next period. Nevertheless, dynamic considerations may still apply to how total reported earnings this period are divided across segments. For example, investors may form expectations of segment-level growth using reported earnings for a particular segment. In this first test of the theory, we abstract away from these dynamic concerns and consider a manager who distorts the consistency of segment earnings to improve short-run perceptions of her managerial ability, e.g., to improve the manager’s probability of receiving an outside job offer.

We empirically test whether segment earnings display abnormally high (low) consistency when overall firm performance is better (worse) than expected. Our analysis allows for the possibility that the consistency of segment earnings varies with firm performance for other natural reasons. For example, bad times may cause higher volatility across segments. In addition, performance across segments may be less variable during good times because good firm-level news is caused by complementarities arising from the good performance of related segments. There may also be scale effects, in that the standard deviation of news may naturally increase in the absolute values of the news. Therefore, we don’t use zero correlation between the consistency of segment earnings and overall firm performance as our null hypothesis.

Instead, we compare the consistency of reported segment earnings to a benchmark consistency of earnings implied by segment-level sales data. Like earnings, the consistency of segment sales may vary with firm performance for natural reasons. However, sales are more difficult to distort because they are reported prior to the deduction of costs. This benchmark consistency implied by

²³Our analysis is also motivated by anecdotal evidence that managers emphasize consistent or inconsistent segment news depending on whether overall firm performance is good or bad. For example, Walmart’s 2015 Q2 10Q highlights balanced growth following strong performance, “Each of our segments contributes to the Company’s operating results differently, but each has generally maintained a consistent contribution rate to the Company’s net sales and operating income in recent years.” In contrast, Hewlett-Packard CEO Meg Whitman highlights contrarian segment performance after sharply negative growth in five out of six segments in 2015, “HP delivered results in the third quarter that reflect very strong performance in our Enterprise Group and substantial progress in turning around Enterprise Services.”

segment sales leads to a conservative null hypothesis. Prior to strategic cost allocations, managers may have already distorted the consistency of segment sales through transfer pricing or the targeted allocation of effort and resources across segments. As a further test, we compare the consistency of segment earnings in real multi-segment firms to that of counterfactual firms constructed from matched single-segment firms. This benchmark captures natural changes in consistency that may be driven by industry trends among connected segments during good and bad times.

4.2 Data and Empirical Framework

We use Compustat segment data merged with I/B/E/S and CRSP for multi-segment firms in the years 1976-2014. We restrict the sample to business and operating segments (some firms report geographic segments in addition to business segments). We exclude observations if they are associated with a firm that, at any point during our sample period, contained a segment in the financial services or regulated utilities sectors, as these firms face additional oversight over their operations and accounting disclosure. In our baseline analysis, we also exclude very small segments (segments with assets in the previous year less than one-tenth that of the largest segment), although we explore how our results vary with size ratios in supplementary analysis.

We follow the framework in Section 3.5, which extends the model to a setting with asymmetric project weights. We measure segment earnings as EBIT (raw earnings) scaled by the segment assets (assets are measured as the average over the current and previous year). We measure firm earnings as the sum of segment EBIT divided by the sum of segment assets. This scaled measure of earnings is also known as return on assets (ROA). We focus on this scaled measure of earnings because it is commonly used by financial analysts, investors, and corporate boards to assess performance and is easily comparable across firms and segments of different sizes. Due to the scaling, firm earnings are equal to the weighted mean of segment earnings, with the weight for each segment equal to segment assets divided by total firm assets. As shown earlier in Section 3.5, firm earnings remain constant even if costs are strategically shifted across segments, which fits with our model in which managers can distort the consistency of news, holding the weighted mean of news constant. We measure the consistency of news as the weighted standard deviation of segment earnings, with higher standard deviation implying lower consistency.²⁴

We use segment sales data to construct a benchmark for how the consistency of segment earnings would vary with overall firm news in the absence of strategic cost allocations. Consider segment i in firm j in year t . Total firm earnings (unscaled) equal total sales minus total costs ($E_{jt} = Sales_{jt} - Costs_{jt}$) and segment earnings (unscaled) equal segment sales minus costs associated with the segment ($e_{ijt} = sales_{ijt} - costs_{ijt}$). For our first benchmark, we use a “proportional costs”

²⁴In supplementary results, omitted for brevity, we find similar results if we instead equal-weight each segment within a firm-year. Using equal weights, segment news is measured by EBIT scaled by assets within the segment, and firm news is measured as the equal-weighted mean of segment news.

assumption. We assume that, absent distortions, total costs are associated with segments according to the relative levels of sales for each segment. Predicted segment earnings (scaled by segment assets a_{ij}) can be estimated as:

$$\frac{\widehat{e}_{ijt}}{a_{ijt}} = \frac{1}{a_{ijt}} \left(sales_{ijt} - \frac{sales_{ijt}}{Sales_{jt}} \cdot Costs_{jt} \right). \quad (9)$$

We estimate the predicted consistency as the log of the weighted standard deviation of the predicted segment earnings:

$$\widehat{s}_{jt} \equiv \log \left(SD \left(\frac{\widehat{e}_{ijt}}{a_{ijt}} \right) \right). \quad (10)$$

Our baseline regression specification tests whether the difference between the actual standard deviation and predicted standard deviation of segment earnings depends on whether firm news exceeds expectations:

$$s_{jt} - \widehat{s}_{jt} = \beta_o + \beta_1 I_{jt}^{goodnews} + controls + \epsilon_{jt}. \quad (11)$$

$I_{jt}^{goodnews}$ is a dummy variable for whether overall firm news exceeds expectations. *Controls* include year fixed effects and the weighted mean of the absolute values of segment sales and earnings, to account for scale effects in the average relationship between standard deviations and means in the data. Standard errors are allowed to be clustered by firm.

We refer to $s_{jt} - \widehat{s}_{jt}$ as the abnormal standard deviation of segment earnings. Our null hypothesis is $\beta_1 = 0$, i.e., that differences between the actual and predicted standard deviations of segment earnings are unrelated to whether the firm is releasing good or bad news overall. This null hypothesis allows for the possibility that we predict the consistency of segment earnings with error, but requires that the prediction error is uncorrelated with whether firm news exceeds expectations. Our model of strategic distortion of consistency predicts that $\beta_1 < 0$, i.e., that the abnormal standard deviation of segment earnings is lower when firm news is good than when firm news is bad.

We can also use industry data to improve the predictions of earnings consistency absent cost allocation distortions. Instead of assuming that total costs would be associated with segments according to the relative levels of segment sales, we can further adjust using industry averages calculated from single-segment firms in the same industry. This helps to account for the possibility that some segments are in industries that tend to have very low or high costs relative to sales. Let γ_{it} equal the average ratio of costs to sales among single segment firms in the SIC2 industry corresponding to segment i in each year. Let $Z_{jt} \equiv \sum_i (\gamma_{it} \cdot sales_{ijt})$. Under an ‘‘industry-adjusted’’ assumption, total costs are associated with segments according to the relative, industry-adjusted,

level of sales of each segment:

$$\frac{\widehat{e}_{ijt}}{a_{ijt}} = \frac{1}{a_{ijt}} \left(sales_{ijt} - \frac{\gamma_{it} \cdot sales_{ijt}}{Z_{jt}} Costs_{jt} \right) \quad (12)$$

We can then substitute the above definition for Equation (9) and reestimate our baseline regression specification.

Our baseline specification assumes that the receiver focuses on earnings news in terms of the level of earnings scaled by assets, otherwise known as ROA. The receiver of news may alternatively focus on performance relative to other similar firms. We can extend our analysis to the case in which receivers of earnings news focus on earnings relative to the industry mean. We measure relative segment earnings as $\frac{e_{ijt}}{a_{ijt}} - m_{it}$, where m_{it} is the value-weighted mean earnings (also scaled by assets) for the segment's associated SIC2 industry in year t . We measure firm relative earnings as $\frac{E_{ijt}}{A_{ijt}} - M_{it}$, where $M_{ijt} \equiv \sum_i \left(\frac{a_{ijt}}{A_{jt}} \right) m_{it}$. Using these measures, firm-level relative earnings is equal to the weighted mean of segment relative earnings, with the weight for each segment again equal to segment assets divided by total assets. The predicted relative earnings for each segment is simply $\frac{\widehat{e}_{ijt}}{a_{ijt}} - m_{it}$, where $\frac{\widehat{e}_{ijt}}{a_{ijt}}$ is as defined in Equations (9) or (12). Using these measures, we can let $s_{jt} - \widehat{s}_{jt}$ equal the difference between the real and predicted log weighted standard deviations of relative segment earnings and reestimate our baseline regression specification in Equation (11).

In Equation (11), $I_{jt}^{goodnews}$ is a dummy variable for whether overall firm news exceeds expectations. In our baseline specifications, $I_{jt}^{goodnews}$ indicates whether total firm earnings exceeds the same measure in the previous year. In tests focusing on relative segment earnings, we can instead let $I_{jt}^{goodnews}$ be an indicator for whether total firm earnings exceeds the industry mean (M_{ijt}). In supplementary tests, we find similar results if $I_{jt}^{goodnews}$ is an indicator for whether total firm earnings exceeds zero, the “break even” point.

Finally, we can measure firm news continuously as (1) the difference between total firm earnings and the same measure in the previous year, or (2) the difference between total firm earnings and the industry mean (M_{ijt}). Our theory does not predict that the consistency of segment earnings should increase *continuously* with firm performance. Rather, the theory predicts a jump in abnormal consistency when firm performance exceeds expectations. For example, the theory predicts that managers will increase consistency when firm news exceeds expectations, but not more so when firm news greatly exceeds expectations. However, the empirically-measured relationship between the consistency of segment earnings and firm performance may be smooth because we use noisy proxies for the expectations of those viewing the segment news disclosures.

Table 1 summarizes the data. Our baseline regression sample consists of 4,297 firms, corresponding to 23,276 firm-years observations. This final sample is derived from an intermediate sample of 60,085 segment-firm-year observations. For a firm-year observation to be included in the sample,

Table 1
Summary Statistics

This table summarizes the data used in our baseline regression sample. Each observation represents a firm-year. Segment earnings equal segment EBIT divided by segment assets (the average of segment assets in the current and previous years). Segment sales are also scaled by assets. Firm earnings and sales are equal to the weighted means of segment earnings and sales, respectively, where the weights are equal to segment assets divided by total assets. All means and standard deviations are weighted and calculated using the segment data within each firm-year. *Good firm news* is an indicator for whether firm earnings in the current year exceeds the level in the previous year. *Good relative firm news* is an indicator for whether firm earnings exceeds the industry mean (calculated as in Section 4.2) in the same year. *Firm earnings* > 0 is an indicator for whether firm earnings is positive. Δ *Firm earnings* measures the continuous difference between firm earnings in the current and previous years. *Firm relative earnings* measures the continuous difference between firm earnings and industry mean earnings in the current year.

	Mean	Std. dev.	p25	p50	p75
Number of segments	2.575	0.936	2	2	3
Firm earnings (= mean earnings)	0.134	0.146	0.067	0.129	0.199
Std. dev. earnings	0.115	0.133	0.037	0.076	0.141
Log std. dev. earnings	-2.705	1.145	-3.309	-2.582	-1.962
Firm sales (= mean sales)	1.657	0.951	1.054	1.511	2.020
Std. dev. sales	0.545	0.573	0.184	0.371	0.701
Log std. dev. sales	-1.117	1.134	-1.694	-0.991	-0.356
Good firm news (dummy)	0.496				
Good relative firm news (dummy)	0.558				
Firm earnings > 0 (dummy)	0.895				
Δ Firm earnings (continuous)	0.013	0.097	-0.021	0.014	0.047
Firm relative earnings (continuous)	0.012	0.249	-0.052	0.008	0.073

we require that the firm reports the same set of segments in the previous year, which allows us to measure segment assets in the previous year as well as annual changes. The first year of a firm \times segment reporting format is excluded from the sample. We present summary statistics of the weighted means and standard deviations of segment earnings and sales. All measures of earnings and sales in this and future tables are scaled by assets unless otherwise noted.

Finally, we emphasize that throughout the empirical tests, we do not take a stand on whether investors, boards, or other receivers of firm earnings news are sophisticated or naïve about the distortion of consistency. The main prediction from the model does not require that receivers rationally expect distortion, just that they use the consistency of earnings as a measure of the precision of the overall earnings signal and that managers react by manipulating consistency. If receivers do anticipate distortion, then as shown earlier, the same predictions apply.

4.3 Empirical Results

Table 2
Consistency of Segment Earnings

The dependent variables in Columns 1 and 2 are the log standard deviation of segment earnings and sales, respectively. The dependent variables in Columns 3 and 4 are the abnormal log standard deviations of segment earnings, relative to predictions calculated using reported segment sales under a proportional costs assumption and industry-adjusted assumption, respectively. Control variables include the *good firm news* indicator, year fixed effects, and controls for the absolute means of segment earnings and sales. All means and standard deviations are weighted by segment assets divided by total assets. All earnings and sales measures are scaled by assets. Standard errors are clustered by firm. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	SD Earnings	SD Sales	Abnormal SD Earnings	
	(1)	(2)	(3)	(4)
Good firm news	-0.0975*** (0.0150)	0.0146 (0.0138)	-0.111*** (0.0179)	-0.0884*** (0.0198)
Cost assumption			Prop	Ind adj
Control for mean	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R ²	0.0643	0.129	0.179	0.0231
Obs	23276	23276	23276	23276

We begin by using regression analysis to test our model prediction that segment earnings are more consistent when firm news is good than when firm news is bad. Table 2 presents our main results. Column 1 regresses the log of the weighted standard deviation of segment earnings on an indicator for good firm news (whether firm earnings beat the same measure last year) as well as controls for year fixed effects and the absolute value of the mean of segment earnings and sales. We find support for the main model prediction that $\beta_1 < 0$. After controlling for general scale effects, the standard deviation of segment earnings is lower when firm news is good than when it is bad. In contrast, Column 2 shows that the standard deviation of segment sales does not vary significantly with the indicator for good firm news.

To more formally test whether the consistency of earnings appears to have been manipulated, Column 3 estimates our baseline specification as described in Equation (11), which compares the consistency of reported segment earnings to a benchmark consistency of segment earnings implied by segment-level sales data. This specification helps to account for other natural factors that may impact the variability of news across segments during good or bad times, assuming that these factors similarly impact segment sales. We find that good firm news corresponds to an abnormal 11% decline in the standard deviation of segment earnings. In Column 4, we find similar results after using an industry-adjusted assumption to create a benchmark consistency of segment earnings from the reported sales data. Under an industry-adjusted assumption, good firm news corresponds to an

abnormal 9 percent decline in the standard deviation of segment earnings. These results support a model in which management distort cost allocations so that good earnings news is consistent and bad earnings news is inconsistent.

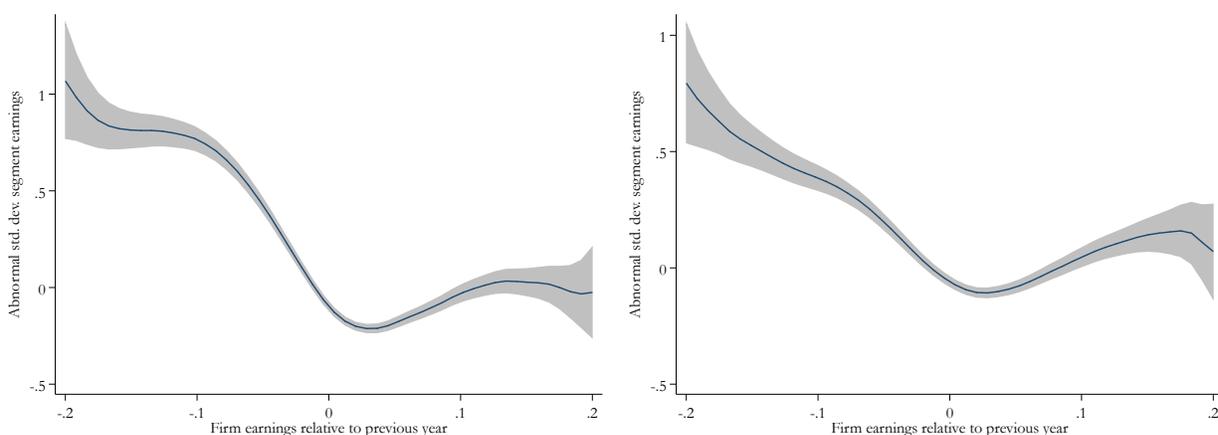
Figure 4

Real vs. predicted consistency of segment earnings

These graphs show how the abnormal standard deviation of segment earnings varies with overall firm news. The x-axis represents firm earnings in the current year minus firm earnings in the previous year. We measure the abnormal standard deviation of segment earnings as the difference between the log weighted standard deviation of segment earnings in the real data and the log weighted standard deviation of predicted segment earnings calculated from reported segment sales data. Predicted segment earnings are formed using a proportional costs assumption in Panel A and an industry-adjusted assumption in Panel B, as described in Section 4.2. We plot how the abnormal standard deviation of segment earnings varies with firm earnings, after controlling for fiscal year fixed effects. The curves represent local linear plots estimated using the standard rule-of-thumb bandwidth. Gray areas indicate 90 percent confidence intervals.

Panel A: Proportional costs assumption

Panel B: Industry-adjusted assumption



We can also visually explore the relationship between the consistency of segment earnings and firm-level news. Figure 4 shows that the abnormal standard deviation of segment earnings is lower when firm news is good than when firm news is bad. We measure firm-level news relative to expectations as the difference between firm earnings in the current year and the same measure in the previous year. We measure the abnormal standard deviation of segment earnings as the difference between the log standard deviation of segment earnings in the real data and the log standard deviation of predicted segment earnings calculated from reported segment sales data. Focusing on the difference between the actual and predicted measures helps to account for potential unobserved factors that may cause the standard deviation of segment earnings differ between good and bad times. These factors should similar affect reported segment sales. However, segment sales are more difficult to manipulate because, unlike earnings, they are reported prior to the discretionary

allocation of shared costs.

In Panels A and B, we use a proportional costs and industry-adjusted assumption, respectively, to predict how earnings would look using reported sales data. Our null hypothesis is that prediction error, i.e., the difference between the standard deviations of real and predicted earnings, should be uncorrelated with whether firm-level news exceeds expectations. Instead, we find that the abnormal standard deviation of earnings sharply declines around zero, when firm-level earnings exceeds the same measure in the previous year. This supports our model prediction that managers strategically allocate costs so that good news is consistent and bad news is inconsistent.

As noted previously, the model does not predict that the abnormal standard deviation of segment earnings should decline continuously with firm performance. Instead, the theory predicts a discrete drop in the abnormal standard deviation of earnings when firm performance exceeds expectations. However, because we proxy for expectations with error, the observed relationship may look more smooth. This is consistent with what we observe in both panels of Figure 4. The abnormal standard deviation of segment earnings declines approximately continuously with firm earnings. In addition, the slope is steepest near zero, when firm earnings in the current year is equal to earnings in the previous year, suggesting that the previous year's earnings represents a reasonable proxy for expectations.

Table 3
Consistency of Relative Segment Earnings

This table reestimates the results in Table 2 using performance measures relative to industry means. Relative earnings, sales, and abnormal relative earnings are as defined in Section 4.2. All other variables are as defined in Table 2. Standard errors are clustered by firm. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	SD Relative Earnings	SD Relative Sales	Abnormal SD Relative Earnings	
	(1)	(2)	(3)	(4)
Good relative firm news	-0.282*** (0.0260)	-0.126*** (0.0236)	-0.180*** (0.0300)	-0.149*** (0.0324)
Cost assumption			Prop	Ind adj
Control for mean	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R ²	0.0669	0.119	0.0728	0.0292
Obs	23276	23276	23275	23276

In our baseline analysis, we assume the receiver focuses on segment and firm earnings news in terms of the level of earnings (which is equal to ROA because we scale by assets). The receiver of news may alternatively focus on performance relative other similar firms. In Table 3, we can extend our analysis to the case in which receivers focus on earnings relative to the industry mean. As described in Section 4.2, we can measure relative segment earnings as segment earnings minus the mean in the segment's SIC2 industry. We measure firm level news as firm earnings minus

the weighted mean performance of the firm’s associated segment industries. The *good relative firm news* indicator is equal to one if firm news exceeds the weighted industry mean. Using these relative performance measures, we again find evidence consistent with strategic distortion of segment news. Column 1 shows that the standard deviation of segment relative earnings is much lower when the firm is underperforming relative to its industry peers than when it is outperforming its peers. Column 2 shows that the standard deviation of segment sales is also significantly lower when the firm is outperforming, although the absolute magnitude of β_1 is smaller than that for segment earnings. This suggests that good times in terms of relative performance may naturally be associated with lower variance in segment news (or that the manager has already strategically manipulated the consistency of segment sales through targeted effort/resource allocation). However, Columns 3 and 4 show that the standard deviation of real segment earnings varies significantly more with overall firm news than predicted given the reported sales data. Under a proportional costs assumption in Column 3, good relative firm news corresponds to an abnormal 18 percent decline in the standard deviation of relative segment earnings. Under an industry-adjusted assumption in Column 4, good relative firm news corresponds to an abnormal 15 percent decline in the standard deviation of relative segment earnings.

Overall, the evidence is supportive of the model prediction that managers strategically allocate shared costs so that standard deviation of segment earning declines more with good firm news than implied by the reported sales data. As a placebo check, we can also compare the behavior of the consistency of segment EBIT within true multi-segment firms with the consistency of matched segments constructed using single-segment firm data. If our results are driven by industry trends among connected segments during good or bad times, we expect to find similar results with industry-matched placebo segments.

We take single-segment firms that have product lines that are comparable to segments in multi-segment firms, and assign them together to mimic multi-segment firms. Specifically, we match each segment-firm-year observation corresponding to a real multi-segment firm to a single segment firm in the same year and SIC2 industry that is the nearest neighbor in terms of Mahalanobis distance for the lagged EBIT, assets, and sales (all unscaled to also match on size). We then regress the log weighted standard deviation of the matched segment earnings, sales, or abnormal earnings on the firm’s performance measures, as in Equation (11). In Table 4, we find that in these artificial multi-segment firms, there is no negative relation between firm performance and the standard deviations of segment earnings or sales. Further, the abnormal standard deviation of segment earnings is not significantly related to the *good firm news* indicator. This is further evidence that managers in real multi-segment firms use their control over cost allocations to distort reported segments earnings to convey consistent good news and inconsistent bad news.

Table 4
Consistency of Matched Segment Earnings

This table reestimates the results in Table 2 using matched segment earnings and sales data. We match each segment-firm-year observation corresponding to a real multi-segment firm to a single segment firm in the same year and SIC2 industry that is the nearest neighbor in terms of the Mahalanobis distance for the lagged unscaled levels of EBIT, assets, and sales. Using the matched data, we calculated the log standard deviations and means of segment earnings, sales, and the abnormal standard deviation of earnings. All other variables are as defined in Table 2. Standard errors are clustered by firm. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	SD Earnings	SD Sales	Abnormal SD Earnings	
	(1)	(2)	(3)	(4)
Good firm news	0.0232 (0.0184)	0.00911 (0.0166)	0.0176 (0.0222)	0.0287 (0.0244)
Cost assumption			Prop	Ind adj
Control for mean	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R ²	0.0119	0.0825	0.211	0.0353
Obs	17192	17191	17192	17192

4.4 Robustness

In this section, we show that our results are robust to reasonable alternative specifications and sample cuts. In Columns 1 through 4 of Table 5 Panel A, we examine the data before and after the passage of SFAS No. 131 in 1997. SFAS No. 131 increased the prevalence of segment reporting among US public firms by requiring that firms disclose segment performance if the segments are evaluated internally as separate units, even if the segments operate in related lines of business. We find significant evidence of distortion of segment earnings as predicted by the model in the periods before and after the policy change.²⁵

In Columns 5 and 6, we limit the sample to firms-years in which all segments are similar in terms of size. In our main analysis, we give more weight to larger segments because they contribute more to overall firm profits and may be more informative of managerial ability. By limiting the sample to firms with similarly-sized segments, we can further check that our results are not driven by relatively small and potentially anomalous segments. After excluding observations in which the size (as measured by assets in the previous year) of the largest segment exceeds the size of the

²⁵Despite regulations mandating segment disclosure, firms may still have some leeway in terms of whether to disclose segment data and which segments to disclose. In relation to our model, it would be very interesting to explore how the choice of which segments to disclose depends on the manager's beliefs about whether overall firm news will be good or bad. In practice, empirical investigation of this question is challenging because firms without major restructuring episodes rarely switch reporting formats year to year and almost always continue to disclose a segment's performance if they disclosed it in the past. Isolated deviations from this norm, e.g., when Valeant Pharmaceuticals switched from reporting five to two segments in 2012, are noted with great suspicion in the financial press. Therefore, we focus our empirical tests on manipulation of consistency after the firm has committed to release performance measures for a set of segments.

Table 5
Robustness

This table reestimates Columns 3 and 4 of Table 2 using alternative sample restrictions and control variables. In Panel A, Columns 1 and 2 are restricted to fiscal years ending on or before 1997 (the year when SFAS No. 131 passed) while Columns 3 and 4 are restricted for fiscal years ending after 1997. Columns 5 and 6 are restricted to firm-year observations in which the assets (measured in the previous year) of the largest segment did not exceed the assets of the smallest segment by more than a factor of 2. In Panel B, Columns 1 and 2, we exclude control variables for the absolute mean of segment earnings and sales. In Columns 3 and 4, we restrict the sample to observations in which the minimum level of segment sales in the previous firm-year is above the 25th percentile in the sample in each year. In Columns 5 and 6, we add in control variables for fixed effects for each firm x segment reporting format (a segment reporting format is a consecutive period in which the firm reports the same set of segments). All other variables are as defined in Table 2. Standard errors are clustered by firm. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A						
Abnormal SD Earnings	Year ≤ 1997		Year > 1997		Similarly Sized Segments	
	(1)	(2)	(3)	(4)	(5)	(6)
Good firm news	-0.132*** (0.0230)	-0.0931*** (0.0248)	-0.0760*** (0.0286)	-0.0887*** (0.0333)	-0.138*** (0.0273)	-0.0972*** (0.0311)
Cost assumption	Prop	Ind adj	Prop	Ind adj	Prop	Ind adj
Control for mean	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.164	0.0189	0.203	0.0324	0.159	0.0221
Obs	14275	14275	9001	9001	10818	10818
Panel B						
Abnormal SD Earnings	No Mean Controls		Exclude Low Sales		Firm-Reporting Format FE	
	(1)	(2)	(3)	(4)	(5)	(6)
Good firm news	-0.253*** (0.0193)	-0.134*** (0.0196)	-0.107*** (0.0207)	-0.100*** (0.0229)	-0.0902*** (0.0217)	-0.0559** (0.0244)
Cost assumption	Prop	Ind adj	Prop	Ind adj	Prop	Ind adj
Control for mean	No	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.0134	0.00974	0.166	0.0251	0.673	0.614
Obs	23276	23276	17454	17454	23276	23276

smallest segments by more than a factor of two, we continue to find that the abnormal standard deviation of segment earnings is significantly smaller during good times than bad.

In Columns 1 and 2 of Panel B, we reestimate our baseline specification described by Equation (11), but exclude controls for the absolute values of mean segment earnings and sales within each firm-year. In regressions in which the dependent variable is the standard deviation of segment earnings or sales, these control variables help account for potential scale effects in which the standard deviation of larger numbers naturally tend to be larger. However, our baseline specification tests whether the difference between the actual and predicted standard deviations of earnings is lower when the firm is reporting good news overall. The predictions using segment sales should already account for potential scale effects, implying that we do not need to further control for scale. We find qualitatively similar estimates, slightly larger in absolute magnitude, if we exclude these scale controls.

In Columns 3 and 4, we show that our results are not caused by sales being bounded below by zero. Unlike earnings, sales cannot be negative. The zero lower bound for sales may mechanically limit the standard deviation of segment sales when overall firm news is bad. We find similar results after restricting the sample to observations in which the minimum level of segment sales in the previous firm-year is above the 25th percentile in the sample.

Finally in Columns 5 and 6, we explore how the abnormal consistency of segment earnings varies with firm news, after controlling for firm \times reporting format fixed effects (the set of reported segments remain constant within a reporting format). We do not control for these fixed effects in our baseline specifications because we wish to use both across-firm variation as well as within-firm variation over time. After controlling for firm \times reporting format fixed effects, we continue to find a strong and significant negative relationship between the abnormal standard deviation of segment earnings and the *good firm news* indicator.

So far in the analysis, we have proxied for whether overall firm news exceeds expectations using dummies for whether firm earnings exceeds the level in the previous year or the industry mean in the same year. In Table 6, we find similar results using alternative definitions of good firm news. In Columns 1 and 2, we define firm news to be good if earnings are positive, i.e., the firm is in the black rather than the red. We again find that the abnormal standard deviation of segment earnings is significantly lower when firm news is good, as measured by earnings being positive. Next, we move to two continuous measures of firm news. As discussed previously, the model predicts a discrete drop in the abnormal standard deviation of earnings when firm performance exceeds expectations. However, the empirically-measured relationship between the consistency of segment earnings and firm performance may be more smooth because we use noisy proxies for the expectations of those viewing the segment news disclosures. In Columns 3 through 6, we find that the abnormal standard deviation of segment earnings declines significantly with these continuous measures of firm news.

Table 6
Alternative definitions of good firm news

This table explores alternative measures firm-level news. Columns 1 and 2 measure good firm news using an indicator for whether firm earnings exceeds zero. Columns 3 and 4 use a continuous measure of firm news equal to the difference between firm earnings in the current and previous year. Columns 5 and 6 use a continuous measure of firm news equal to the difference between firm earnings in the current year and the weighted industry mean, calculated as described in Section 4.2. All other variables are as defined in Tables 2 and 3. Standard errors are clustered by firm. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Abnormal SD:	Earnings				Relative Earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
Firm earnings > 0	-0.694*** (0.0398)	-0.546*** (0.0454)				
Δ Firm earnings (cont)			-0.495*** (0.110)	-0.293** (0.114)		
Firm relative earnings (cont)					-0.669*** (0.0955)	-0.497*** (0.100)
Cost assumption	Prop	Ind adj	Prop	Ind adj	Prop	Ind adj
Control for mean	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.197	0.0332	0.179	0.0226	0.0739	0.0294
Obs	23276	23276	23276	23276	23275	23276

Finally, we explore how the results vary using an alternative “average costs” assumption to predict how total costs within each firm-year would be associated with segments in the absence of strategic distortions. We use information on the average fraction of total costs assigned to each segment over time. For each segment, we calculate the average fraction of total costs that are allocated to the segment, based upon reported segment earnings and sales over the entire period in which a firm-reporting format exists in the data. For example, suppose a firm has three segments A, B, and C. On average, over a five year period, segment A is allocated 0.5 of total costs, segment B is allocated 0.2 of total costs, and segment C is allocated 0.3 of total costs. We can then predict segment earnings in each year, assuming that, absent strategic cost allocations, segments A, B, and C would be assigned 0.5, 0.2, and 0.3, respectively, of total costs in each year. We then test whether the difference between the standard deviation of real segment earnings and these predicted earnings varies with whether the firm is releasing good news overall. In other words, we test whether cost allocations differ from the within-firm mean over time in a manner predicted by the model.

Table 7 shows the results using this average costs assumption. Because our predictions use information on average cost allocations, our regression coefficients are identified from within-firm variation in firm-level news over time. To focus on this within-firm variation, we show specifications that include firm-reporting format fixed effects in Columns 2 and 4. In addition, we also present results using continuous measures of overall firm performance in Columns 3 and 4 to allow for greater

Table 7**Predicted earnings using an average costs assumption**

This table presents results using an average costs assumption to estimate the predicted segment earnings benchmark. For each segment, we calculate the average fraction of total costs that are allocated to the segment based upon reported segment earnings and sales data over the entire period in which a firm-reporting format exists in the data. We predict that, absent strategic distortions, each segment would be associated with this average fraction of total costs for each segment-year in the data. We then estimate the abnormal standard deviation of segment earnings as the difference between the the standard deviation of real segment earnings and these predicted earnings. To use within-firm variation over time, we introduce firm-reporting format fixed effects in Columns 2 and 4. To allow for greater within-firm variation in firm-level news, we use a continuous measure of overall firm performance in Columns 3 and 4. All other variables are as defined in Table 2. Standard errors are clustered by firm. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Abnormal SD Earnings (Average Costs Assumption)			
	(1)	(2)	(3)	(4)
Good firm news	-0.0859*** (0.0145)	-0.0725*** (0.0206)		
Δ Firm earnings (cont)			-0.664*** (0.0857)	-0.665*** (0.135)
Control for mean	Yes	Yes	Yes	Yes
Firm-reporting format FE	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
R ²	0.0883	0.512	0.0900	0.514
Obs	23247	23247	23247	23247

variation in firm-level news over time. In all specifications, we continue to find that the abnormal standard deviation of segment earnings is lower when firm-level news is good. These results support our theory of consistent good news and inconsistent bad news and show that our findings are robust to a variety of different cost allocation benchmarks.

5 Conclusion

These results show that selective news distortion to affect the consistency and hence persuasiveness of news can be an important strategy in sender-receiver environments. Theoretically we show when a more precise signal is more persuasive in that it moves the posterior distribution and posterior estimate more strongly away from the prior, and when more consistent news implies that the mean of the news is a more precise signal. We then show that the most persuasive strategy for a sender is to increase or decrease consistency in good or bad times by focusing on distorting the least or most favorable news respectively. These incentives for selective distortion lead to the testable implication that reported news is more consistent when it is favorable than unfavorable.

We test the model on firms that report earnings data for multiple segments. We show that earnings are more consistent when the firm is doing well than when the firm is doing poorly. This same pattern does not arise in the consistency of reported sales across firms. Similarly the same pattern does not arise in a sample of stand-alone firms matched to segments in multi-segment firms. Therefore the evidence supports the interpretation that multi-segment firms shift the allocation of costs across segments to manipulate the consistency of reported earnings.

6 Appendix

Proof of Lemma 1. Since $E[q]$ is arbitrary, assume WLOG that $y = 0$. Then, using the symmetry of g ,

$$\begin{aligned}
E[q|x, \rho] &= \frac{\int qg(-q, \rho)f(q)dq}{\int g(-q, \rho)f(q)dq} = \frac{\int_{-\infty}^0 qg(-q, \rho)f(q)dq + \int_0^{\infty} qg(-q, \rho)f(q)dq}{\int_{-\infty}^0 g(-q, \rho)f(q)dq + \int_0^{\infty} g(-q, \rho)f(q)dq} \\
&= \frac{\int_0^{\infty} qg(q, \rho)(f(q) - f(-q))dq}{\int_0^{\infty} g(q, \rho)(f(q) + f(-q))dq} \\
&= \frac{\int_0^{\infty} qg(q, \rho)(f(q) - f(-q))\frac{f(q)+f(-q)}{f(q)+f(-q)}dq}{\int_0^{\infty} g(q, \rho)(f(q) + f(-q))dq} \\
&= \int_0^{\infty} v(q)dZ(q)
\end{aligned} \tag{12}$$

where

$$v(q) = q \frac{f(q) - f(-q)}{f(q) + f(-q)} \tag{13}$$

and

$$Z(q|\rho) = \frac{\int_0^q g(\bar{q}|\rho)(f(\bar{q}) + f(-\bar{q}))d\bar{q}}{\int_0^{\infty} g(\bar{q}|\rho)(f(\bar{q}) + f(-\bar{q}))d\bar{q}}. \tag{14}$$

Integrating (12) by parts,

$$E[q|x, \rho] - E[q|x, \rho'] = \int_0^{\infty} v'(q) (Z(q|\rho') - Z(q|\rho)) dq. \tag{15}$$

By assumption $g(q, \rho') \succ_{UV} g(q, \rho)$, so for $\rho' > \rho$ and $q > 0$, $g(q, \rho') \succ_{MLR} g(q, \rho)$ and hence $g(q, \rho')(f(q) + f(-q)) \succ_{MLR} g(q, \rho)(f(q) + f(-q))$. MLR dominance implies FOSD dominance, so $Z(q|\rho') < Z(q|\rho)$ for all $q > 0$. Hence from (15) $E[q|x, \rho] < E[q|x, \rho']$ if $v' > 0$ and $E[q|x, \rho] > E[q|x, \rho']$ if $v' < 0$.

If $E[q] > 0$ then by symmetry and quasiconcavity $f(q) > f(-q)$ for $q > 0$, so taking the derivative of the log of $v(q)$, and noting that $f(-q) = f(q + 2E[q])$, the sign of v' is positive if

$$\frac{1}{q} + \frac{f'(q) - f'(q + 2E[q])}{f(q) - f(q + 2E[q])} - \frac{f'(q) + f'(q + 2E[q])}{f(q) + f(q + 2E[q])} > 0 \tag{16}$$

which, since $q > 0$, holds if

$$f(q + 2E[q])f'(q) - f(q)f'(q + 2E[q]) > 0 \tag{17}$$

or

$$\frac{d}{dq} \ln f(q) > \frac{d}{dq} \ln f(q + 2E[q]) \tag{18}$$

which holds for $E[q] > 0$ by logconcavity of f . So $E[q|x, \rho] < E[q|x, \rho']$ for $E[q] > y = 0$.

If instead $E[q] < 0$ then $f(q) < f(-q)$ for $q > 0$, and following similar steps $v' < 0$ so $E[q|x, \rho] > E[q|x, \rho']$. ■

Proof of Lemma 2. Recall from (2) that

$$\begin{aligned} g(\bar{x} - q, s) &= \int \prod_{i=1}^n \phi(x_i - q, \sigma_\varepsilon) dH(\sigma_\varepsilon) \\ &= \int \frac{1}{(\sigma_\varepsilon \sqrt{2\pi})^n} e^{-\frac{ns^2 + n(\bar{x} - q)^2}{2\sigma_\varepsilon^2}} dH(\sigma_\varepsilon) \end{aligned} \quad (19)$$

where we have used the independence of σ_ε and q .

For given q the likelihood ratio $g(\bar{x} - q, s')/g(\bar{x} - q, s)$ is increasing in \bar{x} if $g(\bar{x} - q, s)$ is log-supermodular in (\bar{x}, s) which, since log-supermodularity is preserved by integration, holds if $g(\bar{x} - q, s, \sigma_\varepsilon)$ is log-supermodular in $(\bar{x}, s, \sigma_\varepsilon)$ (Lehmann (1955) - see discussion of Lemma 2 in Athey (2002)). This holds if all the cross-partials are non-negative (Topkis, 1978). Checking,

$$\begin{aligned} \frac{d}{d\bar{x}} \frac{d}{ds} \ln g(\bar{x} - q, s, \sigma_\varepsilon) &= 0 \\ \frac{d}{d\sigma_\varepsilon} \frac{d}{ds} \ln g(\bar{x} - q, s, \sigma_\varepsilon) &= \frac{2ns}{\sigma_\varepsilon^3} \\ \frac{d}{d\sigma_\varepsilon} \frac{d}{d\bar{x}} \ln g(\bar{x} - q, s, \sigma_\varepsilon) &= \frac{2n(\bar{x} - q)}{\sigma_\varepsilon^3} \end{aligned} \quad (20)$$

which are all non-negative for $\bar{x} \geq q$. Hence $g(\bar{x} - q, s')/g(\bar{x} - q, s)$ is increasing in \bar{x} for $\bar{x} \geq q$ and, by symmetry, $g(\bar{x} - q, s')/g(\bar{x} - q, s)$ is decreasing in \bar{x} for $\bar{x} \leq q$. This establishes that $g(\bar{x} - q, s') \succ_{UV} g(\bar{x} - q, s)$, so we can apply Lemma 1 where $y = \bar{x}$ and $\rho = s$. ■

Lemma 3 *If $g(y - q|\rho)$ is a symmetric quasiconcave density where $g(y - q|\rho') \succ_{UV} g(y - q|\rho)$ for $\rho' > \rho$ and f is a symmetric logconcave density, then (i) $F(q|y, \rho) \leq F(q|y, \rho')$ for all $q < y$ if $y \geq E[q]$ and (ii) $F(q|y, \rho) \geq F(q|y, \rho')$ for all $q > y$ if $y \leq E[q]$.*

Proof. (i) Assume WLOG that $E[q] = 0$. By UV dominance, for $q < q' < y$, $g(q' - y|\rho)/g(q' - y|\rho') > g(q - y|\rho)/g(q - y|\rho')$ or $g(q' - y|\rho)/g(q - y|\rho) > g(q' - y|\rho')/g(q - y|\rho')$. So for any $a < y$, integrating over $q \in [a, y]$,

$$\frac{\int_a^y g(\bar{q} - y|\rho) f(\bar{q}) d\bar{q}}{g(q - y|\rho)} > \frac{\int_a^y g(\bar{q} - y|\rho') f(\bar{q}) d\bar{q}}{g(q - y|\rho')} \quad (21)$$

or

$$\frac{g(q - y|\rho)}{\int_a^y g(\bar{q} - y|\rho) f(\bar{q}) d\bar{q}} < \frac{g(q - y|\rho')}{\int_a^y g(\bar{q} - y|\rho') f(\bar{q}) d\bar{q}}. \quad (22)$$

Integrating over q for $q \leq a$ yields

$$\frac{\int_{-\infty}^a g(q - y|\rho) f(q) dq}{\int_a^y g(\bar{q} - y|\rho) f(\bar{q}) d\bar{q}} < \frac{\int_{-\infty}^a g(q - y|\rho') f(q) dq}{\int_a^y g(\bar{q} - y|\rho') f(\bar{q}) d\bar{q}} \quad (23)$$

or, for any $q < y$,

$$\frac{F(q|y, \rho)}{F(y|y, \rho) - F(q|y, \rho)} < \frac{F(q|y, \rho')}{F(y|y, \rho') - F(q'|y, \rho')} \quad (24)$$

or

$$F(q|y, \rho)F(y|y, \rho') < F(q|y, \rho')F(y|y, \rho). \quad (25)$$

Since $g(y - q) = g(q - y)$ and also $f(q) = f(-q)$ by $E[q] = 0$,

$$\begin{aligned} F(y|y, \rho) &= \frac{\int_{-\infty}^y g(q - y|\rho)f(q)dq}{\int_{-\infty}^y g(q - y|\rho)f(q)dq + \int_y^{\infty} g(q - y|\rho)f(q)dq} \\ &= \frac{\int_{-\infty}^y g(y - q|\rho)f(q)dq}{\int_{-\infty}^y g(q - y|\rho)(f(q) + f(q - 2y))dq} \\ &= \frac{\int_{-\infty}^y v(q)g(q - y|\rho)(f(q) + f(q - 2y))dq}{\int_{-\infty}^y g(q - y|\rho)(f(q) + f(q - 2y))dq}, \end{aligned} \quad (26)$$

where

$$v(q) = \frac{f(q)}{f(q) + f(q - 2y)}. \quad (27)$$

Note that $g(y - q|\rho)(f(q) + f(q - 2y)) \succ_{MLR} g(y - q|\rho')(f(q) + f(q - 2y))$ since $g(y - q|\rho) \succ_{MLR} g(y - q|\rho')$ for $q < y$. Therefore, letting

$$Z(q|y, \rho) = \frac{g(q - y|\rho)(f(q) + f(q - 2y))}{\int_0^{\infty} g(q - y|\rho)(f(q) + f(q - 2y))dq}, \quad (28)$$

MLR dominance implies FOSD dominance, so $Z(q|\rho) < Z(q|\rho')$ for all $q < 0$. Since

$$F(y|y, \rho) - F(y|y, \rho') = \int_0^q v'(q) (Z(q|y, \rho') - Z(q|y, \rho)) dq \quad (29)$$

therefore the sign of $F(y|x, \rho) - F(y|x, \rho')$ equals the sign of the slope of $v(q)$. From the derivative of (27), this is the same sign as

$$\ln f(q) - \frac{d}{dq} \ln f(q - 2y) < 0, \quad (30)$$

where the inequality follows by logconcavity of f for $y > E[q] = 0$. This establishes $F(y|y, \rho) < F(y|y, \rho')$, so (25) implies $F(q|y, \rho) < F(q|y, \rho')$ for $q < y$.

(ii) The case of $y < E[q]$ follows by symmetry. ■

Lemma 4 *If $g(y - q|\rho)$ is a symmetric quasiconcave density where $g(y - q|\rho') \succ_{UV} g(y - q|\rho)$ for $\rho' > \rho$ and f is a symmetric logconcave density, then (i) $\int_{-\infty}^a F(q|y, \rho)da \leq \int_{-\infty}^a F(q|y, \rho')dq$ for all a if $y \geq E[q]$ and (ii) $\int_{-\infty}^a F(q|y, \rho)dq \geq \int_{-\infty}^a F(q|y, \rho')dq$ for all a if $y \leq E[q]$.*

Proof. (i) The likelihood ratio of the posterior densities is

$$\begin{aligned}\frac{f(q|y, \rho)}{f(q|y, \rho')} &= \frac{f(q)g(q-y|\rho)}{\int f(q)g(q-y|\rho)dq} / \frac{f(q)g(q-y|\rho')}{\int f(q)g(q-y|\rho')dq} \\ &= \frac{g(q-y|\rho)}{g(q-y|\rho')} / \frac{\int f(q)g(q-y|\rho)dq}{\int f(q)g(q-y|\rho')dq},\end{aligned}\tag{31}$$

so

$$\frac{d}{dq} \frac{f(q|y, \rho)}{f(q|y, \rho')} \propto \frac{d}{dq} \frac{g(q-y|\rho)}{g(q-y|\rho')},\tag{32}$$

and hence $g(q-y|\rho') \succ_{UV} g(q-y|\rho)$ implies $f(q|y, \rho') \succ_{UV} f(q|y, \rho)$. Since $\lim_{q \rightarrow -\infty} F(q|y, \rho') = \lim_{q \rightarrow -\infty} F(q|y, \rho) = 0$ and $\lim_{q \rightarrow \infty} F(q|y, \rho') = \lim_{q \rightarrow \infty} F(q|y, \rho) = 1$, therefore $F(q|y, \rho')$ crosses $F(q|y, \rho)$ once from above. By Lemma 1 above $E[q|y, \rho] > E[q|y, \rho']$ or, integrating by parts,

$$\int_{-\infty}^{\infty} F(q|y, \rho)dq < \int_{-\infty}^{\infty} F(q|y, \rho')dq.\tag{33}$$

Since $F(q|y, \rho')$ crosses $F(q|y, \rho)$ once from above this implies for all a that

$$\int_{-\infty}^a F(q|y, \rho)dq < \int_{-\infty}^a F(q|y, \rho')dq.\tag{34}$$

(ii) Follows by the same logic as (i). ■

Proof of Proposition 3. The naive receiver case follows from the discussion in the text. For the sophisticated receiver case let \tilde{x}^* be the conjectured equilibrium strategy in the statement of the proposition. First suppose $\bar{x} \leq E[q]$, in which case $\tilde{x}^* = (x_1 - d/2, x_2, \dots, x_{n-1}, x_n + d/2)$. If news such that $\tilde{x}_n - \tilde{x}_{n-1} \geq d/2$ and $\tilde{x}_2 - \tilde{x}_1 \geq d/2$ is observed the news is on the path and the receiver believes accordingly that $x = (\tilde{x}_1 + d/2, \tilde{x}_2, \dots, \tilde{x}_{n-1}, \tilde{x}_n - d/2)$. Hence any deviation to $\tilde{x}' \neq \tilde{x}^*$ that is on the path implies the sender is either not distorting in the most s increasing direction or is not using the entire d distortion budget to increase s , so $s(\tilde{x}^{*-1}(\tilde{x}')) < s(\tilde{x}^{*-1}(\tilde{x}^*)) = s(x)$. News such that $\tilde{x}_n - \tilde{x}_{n-1} < d/2$ or $\tilde{x}_2 - \tilde{x}_1 < d/2$ is off the equilibrium path. Let $X(\tilde{x})$ be the feasible set of x that satisfy (6) for an observed \tilde{x} that is off the equilibrium path. Let \underline{x} be the $x \in X(\tilde{x})$ that implies the lowest s , which is unique by (4). Suppose the receiver puts all weight on \underline{x} . Since \tilde{x}^* is feasible and implies at least a high s as \underline{x} by definition of \underline{x} , no type benefits by deviating. So \tilde{x}^* is a PBE strategy.

To show that \tilde{x}^* survives D1, we need to show that beliefs putting all weight on \underline{x} are not eliminated by the refinement. We adopt standard D1 notation for best responses by a receiver to our case of values of s inferred by the receiver. For subsets I of $X(\tilde{x})$ let $S(I, \tilde{x})$ be the set of s generated by beliefs concentrated on I : $s(I, \tilde{x}) = \cup_{\{\mu: \mu(I)=1\}} s(\mu, \tilde{x})$. Define D_x as the set of s that cause type x to deviate, $D_x = \{s \in S(X(\tilde{x}), \tilde{x}) : U^*(x) < U(\bar{x}, s)\}$ and D_x^o as the set of s that make x indifferent, $D_x^o = \{s \in S(X(\tilde{x}), \tilde{x}) : U^*(x) = U(\bar{x}, s)\}$. The D1 refinement requires that for any x if there exists x' such $D_x \cup D_x^o \subseteq D_{x'}$ that zero weight be put on type x in the

receiver's beliefs. For any $x \neq \underline{x}$, $U^*(x) > U^*(\underline{x})$ so by continuity there exists a feasible $x' \neq \underline{x}$ such that $U^*(x) > U^*(x')$, so type x' is strictly willing to deviate if type x is weakly willing to deviate. Therefore $D_x \cup D_x^\circ \subseteq D_{x'}$ so type x must have zero weight. For type $x = \underline{x}$, no such x' can exist. Hence D1 not only permits beliefs putting all weight on \underline{x} , but such beliefs are the only beliefs surviving D1.

Now suppose $\bar{x} \geq E[q]$, in which case any report is an equilibrium report for some realization of x , so the only question is whether any types benefit from \tilde{x}^* given equilibrium beliefs. If a report is observed such that some extreme news is pooled, $\tilde{x}_1 = \tilde{x}_2 = \dots = \tilde{x}_j$ and $\tilde{x}_k = \dots = \tilde{x}_{n-1} = \tilde{x}_n$ for some $j > 1$ and/or $k < n$, then the report is not invertible as beliefs imply a range of possible x and a corresponding distribution over s . Otherwise, even if some non-extreme news is identical, the report is inferred as fully separating and hence invertible.

First consider when x is such that \tilde{x}^* is separating. It is not feasible to deviate to an equilibrium report with pooling since by construction the sender pools the most extreme news when feasible. And, by the same argument as above, deviating to another invertible report leads the receiver to believe that s is higher than from the equilibrium strategy. So no type whose equilibrium strategy is separating benefits by deviating.

Now consider when x is such that \tilde{x}^* involves some pooling of extreme news, so equilibrium beliefs imply a distribution over possible values of s . Since U is monotonically decreasing in s , a distribution that is first-order stochastic dominated is better for the sender. For any report \tilde{x} the least favorable case in the support of receiver beliefs is that $x = (\tilde{x}_1 - d/2, \tilde{x}_2, \dots, \tilde{x}_{n-1}, \tilde{x}_n + d/2)$, which is the same as separating. Therefore deviating to any report that is for a separating type never benefits the sender. So the remaining case is deviating to another pooling report. Clearly the sender will never benefit from deviating to a higher \tilde{x}_i if $\tilde{x}_i > \bar{x}$ or a lower \tilde{x}_i if $\tilde{x}_i < \bar{x}$, or ever benefit from using less than the full distortion budget on each side, so deviations on either side of the mean by a total of $d/2$ can be considered separately. Consider deviations from the upper pooling region. Deviating to another pooling report implies reducing some report or reports that are pooled in equilibrium to below other pooling reports, and hence reducing the pooled reports less. The effect on s from reducing the pooled reports is least for the case in the support of receiver equilibrium beliefs where $x = (\tilde{x}_1 - d/(2j), \dots, \tilde{x}_j - d/(2j), \tilde{x}_{j+1}, \dots, \tilde{x}_{k-1}, \tilde{x}_k + d/(2(n+1-k)), \dots, \tilde{x}_n + d/(2(n+1-k)))$. But even in this case, by (4) the effect on s is larger than from reducing any reports below the pool, so such a deviation hurts the sender. By the same argument, deviations from the lower pooling region must also hurt the sender, so no type benefits from deviating from \tilde{x}^* . ■

References

- [1] Athey, Susan (2002). "Monotone comparative statics under uncertainty," *Quarterly Journal of Economics* 117(1), 187–223.

- [2] Babbage, Charles (1830). *Reflections on the Decline of Science in England, and on Some of Its Causes*. London.
- [3] Berger, Philip G. and Rebecca Hann (2003). “The impact of SFAS No. 131 on information and monitoring,” *Journal of Accounting Research* 41(2), 163–223.
- [4] Berger, Philip G. and Rebecca N. Hann (2007). “Segment profitability and the proprietary and agency costs of disclosure,” *The Accounting Review* 82(4), 869–906.
- [5] Chakraborty, Archishman and Rick Harbaugh (2010). “Persuasion by cheap talk,” *American Economic Review* 100(5), 2361–2382.
- [6] Chambers, Christopher P. and Paul J. Healy (2012). “Updating toward the signal,” *Economic Theory* 50, 765–786.
- [7] Chevalier, Judith and Glenn Ellison (1999). “Career concerns of mutual fund managers,” *Quarterly Journal of Economics* 114(2), 389–432.
- [8] Cho, In-Koo and David M. Kreps (1987). “Signaling games and stable equilibria,” *Quarterly Journal of Economics* 102(2), 179–221.
- [9] Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers (1997), “Measuring mutual fund performance with characteristic-based benchmarks,” *Journal of Finance* 52 (3), 1035–1058.
- [10] Dawid, A.P. (1973). “Posterior expectations for large observations,” *Biometrika* 60(3), 664–667.
- [11] Egger, Matthias, George Davey Smith, Martin Schneider, and Christoph Minder (1997). “Bias in meta-analysis detected by a simple, graphical test,” *British Medical Journal* 315, 629–634.
- [12] Epstein, Marc J., and Krishna G. Palepu (1999). “What financial analysts want,” *Strategic Finance* 80(10), 48–52.
- [13] Finucan, H.M. (1973). “The relative position of the two regression curves,” *Journal of the Royal Statistical Society. Series B (Methodological)* 35(2), 316–322.
- [14] Gentzkow, Matthew, and Jesse M. Shapiro (2006). “Media bias and reputation,” *Journal of Political Economy* 114(2), 280–316.
- [15] Givoly, Dan, Carla Hayn, and Julia D’Souza (1999). “Measurement errors and information content of segment reporting,” *Review of Accounting Studies* 4, 15–43.

- [16] Holmstrom, Bengt and Paul Milgrom (1991). “Multitask principal-agent analyses: incentive contracts, asset ownership, and job design,” *Journal of Law, Economics, & Organization* 7, 24–52.
- [17] Jacob, John (1996). “Taxes and transfer pricing: income shifting and the volume of intrafirm transfers,” *Journal of Accounting Research* 34 (2) 301–312.
- [18] Kamenica, Emir and Matthew Gentzkow (2011). “Bayesian persuasion,” *American Economic Review* 101(6), 2590–2616.
- [19] Kirschenheiter, Michael and Nahum D. Melumad (2002). “Can ‘big bath’ and earnings smoothing co-exist as equilibrium financial reporting strategies?” *Journal of Accounting Research* 40(3), 761–796.
- [20] Lehmann, E. L. (1955). “Ordered families of distributions,” *Annals of Mathematical Statistics* 26(3), 399–419.
- [21] Meyer, Jack (1987). “Two-moment decision models and expected utility maximization,” *American Economic Review* 77(3), 421–430.
- [22] Milgrom, Paul R. (1981). “Good news and bad news: Representation theorems and applications,” *Bell Journal of Economics* 12 (2), 380–391.
- [23] O’Hagan, A. (1979). “On outlier rejection phenomena in Bayes inference,” *Journal of the Royal Statistical Society. Series B (Methodological)* 41(3), 358–367.
- [24] O’Hagan, A. (1981). “A moment of indecision,” *Biometrika* 68(1), 329–330.
- [25] Oreskes, Naomi and Erik M. Conway (2010). *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. Bloomsbury Press, New York.
- [26] Scharfstein, David S. and Jeremy C. Stein (2000). “The dark side of internal capital markets: Divisional rent-seeking and inefficient investment,” *Journal of Finance* 55(6), 2537–2564.
- [27] Stein, Jeremy C. (1989). “Efficient capital markets, inefficient firms: A model of myopic corporate behavior,” *Quarterly Journal of Economics* 104(4), 655–669.
- [28] Stone, Daniel F. (2015). “A few bad apples. Communication in the presence of strategic ideologues,” working paper.
- [29] Student (1908). “The probable error of a mean,” *Biometrika* 6(1), 1–25.

- [30] Subramanyam, K. R. (1996). “Uncertain precision and price relations to information,” *The Accounting Review* 71(2), 207–220.
- [31] Topkis, Donald (1978). “Minimizing a submodular function on a lattice,” *Operations Research* 26, 305–321.
- [32] Whitt, Ward (1985). “Uniform conditional variability ordering of probability distributions,” *Journal of Applied Probability* 22(3), 619–633.
- [33] You, Haifeng (2014). “Valuation-driven profit transfer among corporate segments,” *Review of Accounting Studies* 19(2), 805–838.