

Understatement of Good (but Not Great) News*

Rick Harbaugh
Indiana University
riharbau@indiana.edu

Ted To
U.S. Bureau of Labor Statistics
To_T@bls.gov

VERY PRELIMINARY

Abstract

Is it always wise to boast about good news? When messages are verifiable, the standard result from disclosure games is that good news should be disclosed or the receiver will infer that the sender has nothing favorable to report. We reconsider this conclusion when the receiver has private information about sender type. Because better senders expect the receiver to have more favorable private information about them, we find that the worst sender type with good news has the most incentive to disclose it. Therefore reporting good news can paradoxically make an unfavorable impression. If the good news is sufficiently rare a unique disclosure equilibrium exists in which good news is disclosed. But if the good news is attainable even by mediocre types then two understatement equilibria always exist: a nondisclosure equilibrium in which all types withhold good news and a countersignaling equilibrium in which mediocre types report good news but the best types withhold it. When such equilibria exist, the sender has a legitimate fear of looking overly eager to reveal good news. Having a third party disclose the news, or mandating that the sender disclose the news, eliminates this problem.

Key words: disclosure, persuasion game, verifiable message, countersignaling

JEL codes: D82, C78

*We thank Archishman Chakraborty, Tilman Klumpp, Eric Rasmusen, and participants at the 2003 BLISS conference for helpful comments.

1 Introduction

When statements must be truthful, is all information voluntarily disclosed? For instance, if false statements are illegal, will accounting reports disclose all information of interest to investors? And if food labels must be correct, will they include all nutrition information wanted by buyers? In a sender-receiver game with costless and truthful messages, the standard result is that not only does the sender with the most favorable information disclose it, but even senders with average or below average news must follow their lead or the receiver will skeptically infer that even worse news is being hidden. Such unraveling leads to all information being disclosed (Viscusi, 1978; Grossman and Hart, 1980; Milgrom, 1981; Grossman, 1981; Milgrom and Roberts, 1986), so information asymmetries can be solved simply by requiring that any statements made be truthful.

Despite this logic, nondisclosure appears to be common in practice, and many government agencies mandate release of information rather than just require that any statements be true. Most of the literature addresses this discrepancy by examining why the absence of good news is not always treated skeptically. Answers include that messages are costly (Viscusi, 1978; Jovanovic, 1983; Verrecchia, 1983; Dye, 1986), there are strategic reasons for withholding information (Dye, 1986; Giovannoni and Seidmann, 2002), the sender herself is not always fully informed (Dye, 1985; Farrell, 1986; Okuno-Fujiwara et al., 1990; Shin, 1994, 2003), or the receiver is naive (Dye, 1998), uninformed (Fishman and Hagerty, 2003), or boundedly attentive (Hirshleifer et al., 2002).

An alternative approach to understanding nondisclosure is to examine why the disclosure of good news might be viewed skeptically. This approach is taken by Teoh and Hwang (1991) who analyze a two-period disclosure game in which a firm decides whether or not to immediately disclose news that will eventually be made public anyway. For instance, a firm decides whether to release earnings news immediately or wait until it must be reported in the quarterly statement. Under the assumption that high quality firms are more likely to receive additional good news that investors will observe, they find that holding back on good news is a signal of confidence. It hurts a high quality firm temporarily, but eventually separates it from low quality firms who are less likely to have additional favorable news.

Following Teoh and Hwang (1991), we analyze how an eagerness to show off can reflect unfavorably on the sender. However, instead of assuming that the receiver eventually discovers the good news anyway, we follow a standard disclosure game in which there is only one period and the receiver does not learn of news that is withheld. We consider news that is unambiguously favorable and the best available, but still not impressive. For instance, even if an “A” is the

highest grade possible, it need not be rare. When good news is easy to attain, is boasting about it still a good idea? Or is boasting about minor accomplishments treated with such skepticism that the sender is worse off?

Consider the following example of Los Angeles County restaurant hygiene cards. Starting in 1998, health officials began requiring restaurants to post hygiene report cards at their entrances (see Jin and Leslie, 2003). Based on hygiene scores from health inspectors, the cards feature large letter grades with most grades being an A. It might seem that all restaurants with an A grade benefit unambiguously from voluntarily posting their grade, so that mandating the display of grades is only necessary to ensure that those with lower grades also post their grade. But would a restaurant with an A grade always be eager to boast about its hygiene in this fashion? Consider if diners have their own private information about the restaurant. For instance, they have visited the restaurant previously and know it to be a reputable, clean restaurant. For a restaurant who expects most of its diners to already have a favorable impression, is posting its A always a good idea? Will diners interpret the posting as additional evidence of the restaurant's cleanliness, or as a more ambiguous or even bad sign that the restaurant is anxious to prove that it meets county hygiene standards?

We show that diners can have good reason to be suspicious and that nondisclosure by all or some types can therefore be an equilibrium. Consider the case where most restaurants, even ones that are not exceptionally clean, receive an A and the diners' private information is relatively informative. Suppose that no restaurants are currently disclosing hygiene grades. Who then benefits the most from deviating and disclosing information? If all restaurants with an A benefit equally or if the best restaurants benefit the most, standard equilibrium refinements imply that a restaurant which reports an A will be recognized as relatively high quality and will therefore break the nondisclosure equilibrium. But given that diners already have some information about restaurant quality, it is the worst restaurants within the broad A category which have the strongest incentive to separate themselves from even worse restaurants.¹ We show that diners should therefore make a relatively negative inference about a restaurant which unexpectedly posts its grade, so all restaurants are better off sticking to the equilibrium nondisclosure strategy.

To capture these ideas formally requires two key assumptions. First, the sender's message must be noisy in the sense that it cannot fully reveal the sender's type. In particular, we assume

¹If enforcement against false statements is incomplete (Fischer and Verrecchia, 2002; Korn, 2003), then restaurants below the "A" category have the strongest incentive to claim they are in the "A" category so the problem is aggravated.

that there is a continuum of sender types and that types below some cutoff have bad news and types above it have good news. This inability of the message to fully reveal sender type is common in the literature and does not affect the standard full disclosure result (Grossman and Hart, 1980; Milgrom, 1981; Milgrom and Roberts, 1986).² We make the additional, less standard assumption that the receiver has private, noisy information about sender type beyond the message sent by the sender. For instance, in the restaurant example a diner sees the hygiene grade and also has his own impression from personal experience. The interaction between these two sources of noisy information drives our results.

We find that if there is *any* private receiver information two types of understatement equilibria exist whenever the good message is attainable by sufficiently mediocre types. First, a nondisclosure equilibrium exists in which nobody discloses good news. This equilibrium survives standard refinements for off-equilibrium-path behavior and an additional, stricter refinement that we adopt for sender-receiver games with private receiver information. Second, a partial disclosure equilibrium exists where types with bad news do not disclose it, moderately high types with good news disclose it, and the best types with good news do not disclose it. In this equilibrium, which we refer to as a countersignaling equilibrium for its similarity to such an equilibrium in signaling games (Feltovich et al., 2002), high types show their confidence that they are above having to prove themselves by withholding good news. For instance, by not revealing their good grade, the best restaurants show confidence that diners already have a favorable impression of their cleanliness. In this equilibrium there is no off-equilibrium-path behavior to consider.

These results highlight that the ease of attaining good news is a key factor in allowing nondisclosure. For instance, in the case of restaurant hygiene cards, the system uses broad letter categories instead of numbers and allows most restaurants to receive an A. Similarly, the phenomenon of grade inflation means that a large proportion of moderately serious high school and college students receive primarily A grades. When the best message is attained even by mediocre types, boasting about the message is not necessarily a positive sign so understatement can be an equilibrium.

A complexity that emerges from our analysis is the multiplicity of equilibria. Even when nondisclosure and countersignaling are equilibria, a full disclosure equilibrium coexists. The different equilibria survive standard refinements and cannot typically be Pareto ranked so there is no reason to assume that a particular equilibrium will hold. While low sender types benefit most

²This assumption can be restated as the sender has both verifiable information that can be revealed and unverifiable information that cannot be revealed.

from the nondisclosure equilibrium, medium sender types benefit the most from the disclosure equilibrium and the highest types benefit the most from either the disclosure equilibrium or the countersignaling equilibrium. For receivers, the nondisclosure equilibrium provides the least information but the full disclosure equilibrium and the countersignaling equilibrium can each provide the most information in different situations. Of course, given that there are multiple equilibria, there is no assurance that the sender and receiver will successfully coordinate on one of them.

Comparing our static disclosure game with the two-period game of Teoh and Hwang (1991), the two-period model is appropriate when the sender expects the news to eventually be released independently of the sender’s action. This situation often arises regarding financial statements in which news must be released at regular intervals, but can be released earlier at the firm’s discretion.³ In contrast, the static model is appropriate when the sender’s disclosure decision alone determines whether the news will be released. For instance, a person has a single opportunity to reveal news.⁴ A distinguishing prediction of the two models is the immediate impact of disclosure on the receiver’s expected estimate of the sender’s type. In the main equilibrium of the dynamic game the immediate impact is positive, but the long-run gain from waiting might be even more favorable to the sender. In the nondisclosure equilibrium of the static game the immediate impact is negative, and in the countersignaling equilibrium the immediate impact is positive for intermediate types but negative for higher types.

The only difference in our model from a standard disclosure game is that we allow the receiver to have some private information.⁵ The approach is therefore similar to that of Feltovich et al.

³In addition to the key assumption that the sender’s news is eventually revealed independently of the sender’s disclosure decision, Teoh and Hwang’s two-period game has additional assumptions that reflect the institutional environment they consider. First, the sender receives a payoff both immediately after the choice to disclose and later after the original news and any additional news is revealed. The equilibrium depends on the rate at which the second payoff is discounted. Second, the sender’s news has a direct effect on sender payoffs beyond the usual indirect effect via receiver estimates of the sender’s type.

⁴Of course in many situations aspects of both games will be present. For instance, there might be a positive probability that the news will be released independently of the sender’s action. If for some reason this probability is increasing in the sender’s type, then worse senders will have more incentive to disclose than better senders.

⁵Private receiver information formally captures both information already present at the time of the sender’s disclosure decision and information that becomes available after the sender has made the disclosure decision but before the receiver’s final estimate of sender type that determines the sender’s payoff. In the latter case even if the information is public to both parties it is still “private receiver information” because the sender has already committed to a strategy. Therefore the Teoh and Hwang (1991) model also has private receiver information.

(2002) who analyze the role of private receiver information in standard signaling games. This paper differs in that we consider a disclosure game with a restricted message space of free and truthful messages, rather than a signaling game with an unrestricted space of increasingly expensive messages that depend on their cost for their credibility. It also differs in that Feltovich et al. (2002) consider the countersignaling equilibrium where medium types show off and high types are understated, whereas in this paper we also investigate the nondisclosure equilibrium where all types are understated.

These results on understatement in disclosure games add to our general understanding of understatement. In addition to the countersignaling model of Feltovich et al. (2002), several models in the signaling literature analyze why high types might choose to be understated in their signals. O'Neill (2002) extends the analysis of countersignaling to the case of multiple receivers who have different information. Other models consider why signals might not be monotonically increasing in type when the costs and benefits of signals are viewed more generally, e.g., there are opportunity costs of education (Spence, 2001), or additional benefits of education from learning about one's own abilities (Hvide, 2003). Understatement in one dimension can also arise when there are multi-dimensional signals, e.g., the combination of high prices and modest advertising can sometimes signal high quality (Orzach et al., 2002), and the combination of high prices and low observable quality can sometimes signal high unobservable quality (Clements, 2002).

There are two empirical implications of our main result that understatement can arise when good news is attainable even by mediocre types. First, the frequency of understatement should be negatively correlated with the difficulty of the good news. For instance, in the restaurant example we expect more understatement when A hygiene grades are easy than when they are difficult. Second, the frequency of understatement should be positively correlated with any public signal that is positively correlated with sender type. This holds because the distribution of types in the model is conditional on all public information. If this conditional distribution is weighted toward higher types because of a favorable public signal, then the good news is no longer that impressive so understatement is more likely. In contrast with many signaling models, the predictions can therefore be readily tested using public information.⁶ In the restaurant example we expect more understatement in expensive locations where most restaurants are already likely

⁶In a signaling model the size of the signal is normally increasing in the sender's type which is the sender's *private* information. Since sender type is not known by the receiver it is typically not known by the econometrician, so empirical tests often use indirect methods to evaluate the theory (Bedard, 2001). Here we predict that understatement is more likely based on *public* signals of the sender's type.

to meet hygiene standards. Or consider the question of when PhDs use the title of “Dr.” In some academic departments or professional schools PhDs are relatively rare, while in others they are quite common. In the former case we expect PhDs to advertise their degree, while in the latter case we expect them to be more understated because showing off can have the opposite effect of making them look insecure.

There are also two policy implications of our main result. First, when understatement equilibria exist the sender has a legitimate fear of looking overly eager to reveal good news. The resulting dilemma over whether to reveal good news is aggravated by the fact that there are multiple equilibria so the receiver might expect different behavior than the sender. Having a third party disclose the news, or mandating that the sender disclose the news, eliminates this problem. By taking the decision to disclose good news out of the hands of the sender, the sender can enjoy the benefits of favorable information without looking overly anxious to disclose it.⁷ For instance, if students are reluctant to brag about their grades, then directly posting their grades can ensure that the information is released, thereby increasing study incentives for students.⁸ A second implication relates to the literature on standard-setting, e.g., setting the difficulty of attaining grades or being certified for quality. This literature typically trades off the gains from forcing higher quality versus the losses of lower rates of attainment (Phillips and Zuckerman, 1994). Our model suggests that higher standards have the additional advantage of being less likely to induce an understatement equilibrium.

2 The model

We consider a disclosure game in which the sender has private information about her type that she can choose to disclose to the receiver. There are three sources of information about sender type in this game. First is a nonverifiable signal $q \in [0, 1]$ observed by the sender alone. The realization of this signal defines the sender’s type. Assume that q is distributed according to the density function $f(q)$ with full support on $[0, 1]$. Second is a signal of “bad news” or “good news” observed by the sender, $s \in \{B, G\}$. We assume that $s = G$ if and only if $q \geq q^*$ for some $q^* \in (0, 1)$. We restrict attention to the case of a binary signal so as to highlight that unraveling

⁷As shown in the example section, in the countersignaling equilibrium very high senders might be better off than in the disclosure or nondisclosure equilibria, so it is not true that all types will necessarily benefit from mandatory disclosure if the countersignaling equilibrium is very likely.

⁸Such posting is legally restricted in the United States but common elsewhere. For instance, first year grades at Oxford are published in the local newspapers.

can fail even at the very first step. Third is a signal observed by the receiver alone, $x \in \{L, H\}$ where $L < H$.⁹ It is the presence of this third source of information that drives our results. We assume that $g(H | q)$ is strictly increasing in q , so that x , q , and s are affiliated,¹⁰ and that $0 < g(H | q) < 1$ for all $q \in (0, 1)$.¹¹

We follow the standard assumption that the sender's payoff is the expected value of the receiver's estimate of the sender's type. The only way the sender can affect this estimate is by disclosing or not disclosing $s = G$ when in fact $s = G$.¹² Denote this decision as $R = d$ if the sender discloses and $R = n$ if the sender does not. We assume that the sender's private information is verifiable. That is, if the sender claims that $s = G$, the receiver can costlessly verify whether or not $s = G$. Moreover, we assume that there is a large penalty attached to falsely disclosing $s = G$ if in fact $s = B$. The game is therefore quite simple in that types $q \in [0, q^*)$ only have one possible strategy (nondisclosure) and types $q \in [q^*, 1]$ have only two possible pure strategies (disclosure or nondisclosure).

We consider only pure strategy equilibria so a strategy is a mapping between types and reports. Since the signal s is deterministic given a sender's type, strategies need be contingent only on type. Let R_q represent the disclosure decision of a sender of type q . Let the function $\mu(q | x, R)$ be a conditional probability distribution representing receiver beliefs about which types q choose disclosure decision R and private information x . Having observed private information x and observed whether the sender disclosed good news or not, the receiver's expectation of the sender's type is

$$\bar{q}_\mu(x, R) = \int_0^1 q \mu(q | x, R) dq$$

⁹We model the extra information as one-dimensional and binary for simplicity. The main results on existence of nondisclosure equilibria only require that, because of some extra information, higher types expect to be evaluated more favorably than lower types.

¹⁰Affiliation is a form of pair-wise correlation (Milgrom and Weber, 1982). Letting $h(x, q)$ represent the joint distribution of x and q , in our context with a binary x it says that $h(H, q')h(L, q) \geq h(H, q)h(L, q')$ for all $q' > q$. This is equivalent to $g(H | q')f(q')g(L | q)f(q) \geq g(H | q)f(q)g(L | q')f(q')$ or $g(H | q')g(L | q) \geq g(H | q)g(L | q')$ or $g(H | q') \geq g(H | q)$. Similarly it can be shown that q and s are affiliated, thereby implying that x , q , and s are all affiliated.

¹¹This latter restriction rules out trivial cases of nondisclosure where high types have no need to disclose since they are already fully separated from lower types.

¹²For instance the sender has a certificate to show the receiver if $s = G$. Depending on the situation it may also be possible to make the verifiable statement that $s = B$, but for simplicity we will not consider this possibility.

so the expected payoff to a sender of type q from making disclosure decision R is

$$E[\bar{q}_\mu(x, R) | q] = \sum_{x \in X} \bar{q}_\mu(x, R)g(x | q).$$

Definition 1 *A pure-strategy perfect Bayesian equilibrium is given by a verifiable type-contingent strategy profile R_q and receiver beliefs $\mu(q | R, x)$ where*

1. $E[\bar{q}_\mu(x, R_q) | q] \geq E[\bar{q}_\mu(x, R') | q]$ for any R'
2. for any R , $\mu(q | x, R)$ is updated using Bayes' rule whenever possible.

Condition (1) requires that agents make reports as a best response to the receiver's beliefs. Condition (2) requires that for any information set that can be reached on the equilibrium path, the receiver's beliefs are consistent with Bayes' rule and the equilibrium sender strategy.

Lemma 1 *Let $m(q) = E[\bar{q}_\mu(x, d) | q] - E[\bar{q}_\mu(x, n) | q]$. Let $\Lambda \subset [q^*, 1]$ be a set consisting of all types that choose to disclose. For almost all beliefs that are consistent with Bayes' rule and Λ , $m(q)$ crosses zero at most once.*

Proof. (sketch) Substituting and simplifying,

$$m(q) = g(H | q) (\bar{q}_\mu(H, d) - \bar{q}_\mu(H, n)) + g(L | q) (\bar{q}_\mu(L, d) - \bar{q}_\mu(L, n))$$

so there are five cases to consider: (i) $\bar{q}_\mu(H, d) \geq \bar{q}_\mu(H, n)$ and $\bar{q}_\mu(L, d) \geq \bar{q}_\mu(L, n)$, (ii) $\bar{q}_\mu(H, d) \leq \bar{q}_\mu(H, n)$ and $\bar{q}_\mu(L, d) \leq \bar{q}_\mu(L, n)$, (iii) $\bar{q}_\mu(H, d) \leq \bar{q}_\mu(H, n)$ and $\bar{q}_\mu(L, d) \geq \bar{q}_\mu(L, n)$, (iv) $\bar{q}_\mu(H, d) \geq \bar{q}_\mu(H, n)$ and $\bar{q}_\mu(L, d) \leq \bar{q}_\mu(L, n)$ and (v) $\bar{q}_\mu(H, d) = \bar{q}_\mu(H, n)$ and $\bar{q}_\mu(L, d) = \bar{q}_\mu(L, n)$ where for cases (i-iv) at least one inequality is strict. In case (i) $m(q) > 0$ for all q . In case (ii) $m(q) < 0$ for all q . In case (iii) $m(q)$ is strictly decreasing because $g(H | q)$ is strictly increasing so $m(q)$ crosses zero at most once and from above. In case (iv) $m(q)$ is strictly increasing because $g(H | q)$ is strictly increasing so $m(q)$ crosses zero at most once and from below. In case (v), writing out $\bar{q}_\mu(H, d) - \bar{q}_\mu(H, n)$

$$\int_{\Lambda} q \frac{g(H | q)f(q)}{\int_{\Lambda} g(H | q')f(q')dq'} dq - \int_{[0,1] \setminus \Lambda} q \frac{g(H | q)f(q)}{\int_{[0,1] \setminus \Lambda} g(H | q')f(q')dq'} dq$$

Given Λ , the set of distributions for which both $\bar{q}_\mu(H, d) - \bar{q}_\mu(H, n)$ and $\bar{q}_\mu(L, d) - \bar{q}_\mu(L, n)$ equal zero is of measure zero. ■

As the following proposition shows, there are typically only three types of pure-strategy equilibria that can arise in this game.¹³ There is the standard disclosure equilibrium in which all types with good news disclose it and also two types of equilibria in which at least some types are understated. The first such equilibrium is a nondisclosure equilibrium in which all types with good news withhold it. The second type of understatement equilibrium is motivated by Feltovich et al. (2002). They show that in signaling games with private receiver information non-monotonic or “countersignaling” equilibria can emerge in which medium types send a larger signal than high types. Such an equilibrium has a direct parallel in our disclosure game in which only types within some range disclose good news so for simplicity so we refer to it as a countersignaling equilibrium.

Proposition 1 *Only three types of pure-strategy perfect Bayesian equilibria are possible in this game: (i) a disclosure equilibrium in which $\{q \mid R_q = d\} = [q^*, 1]$, (ii) a nondisclosure equilibrium in which $\{q \mid R_q = d\} = \emptyset$, and (iii) a countersignaling equilibrium in which $\{q \mid R_q = d\} = [q^*, b]$ where $b \in [q^*, 1]$.*

Proof. Let $D_\mu = \{q \mid q \geq q^*, E[\bar{q}_\mu(x, d) \mid q] \geq E[\bar{q}_\mu(x, n) \mid q]\}$ for any given beliefs $\mu(q \mid x, R)$. We will first show that Lemma 1 implies that, for any beliefs, D_μ must be an interval of the form $[q^*, 1], \emptyset, [q^*, b], [b, 1]$ where $q^* \leq b \leq 1$. We will then show that $D_\mu = [b', 1]$ cannot be part of an equilibrium, so only the equilibria in (i), (ii) or (iii) are possible.

From Lemma 1 $m(q) = E[\bar{q}_\mu(x, d) \mid q] - E[\bar{q}_\mu(x, n) \mid q]$ almost always crosses zero at most once for any given Bayesian beliefs $\mu(q \mid x, R)$. If $m(q)$ does cross zero let $b \in [0, 1]$ be the crossing point. If $m(q) > 0$ for all q or $m(q)$ crosses zero from above and $b \in [0, q^*)$ then $D_\mu = \emptyset$. If $m(q) < 0$ for all q or $m(q)$ crosses zero from below and $b \in [0, q^*)$ then $D_\mu = [q^*, 1]$. If $m(q)$ crosses zero from above $b \in [q^*, 1]$ then $D_\mu = [q^*, b]$. Finally, if $m(q)$ crosses zero from below and $b \in [q^*, 1]$ then $D_\mu = [b, 1]$.

Now consider if $D_\mu = [b, 1]$ which means that high types disclose but medium types do not. This implies that high and medium types get a payoff that is strictly greater and less than b . Since disclosing for medium types is costless, medium types have a strict incentive to deviate. Therefore, an equilibrium where all $q \in [b, 1]$ disclose but $q \in [0, b)$ do not cannot be an equilibrium. ■

We now turn attention to when these three possible equilibria exist. Because the set of possible equilibria is so straightforward, it is convenient to use simpler notation. Define the receiver’s expectation of the sender’s type, having observed x and given that she believes the sender to be

¹³With a more complicated distribution of x other equilibria might be possible.

of type $q \in \Lambda$, as

$$\bar{q}_\Lambda(x) = \int_{q' \in \Lambda} q' \frac{g(x | q') f(q')}{\int_{q'' \in \Lambda} g(x | q'') f(q'') dq''} dq'$$

when $\int_\Lambda f(q) dq > 0$ and let $\bar{q}_\Lambda(x) = q$ when $\Lambda = \{q\}$. Therefore, given that the receiver believes the sender to be of type $q \in \Lambda$, a sender of type q gets an expected payoff of

$$E[\bar{q}_\Lambda(x) | q] = \sum_{x \in X} \bar{q}_\Lambda(x) g(x | q).$$

Assumption 1 $\bar{q}_{[a,b]}(x)$ is submodular in (x, a) and supermodular in (x, b) for all $0 < a < b < 1$.¹⁴

This assumption is made for technical reasons and implies that the gap between $\bar{q}_{[a,b]}(H) - \bar{q}_{[a,b]}(L)$ shrinks as the interval $[a, b]$ narrows from either side.¹⁵

Proposition 2 (i) A disclosure equilibrium exists for all $q^* \in (0, 1)$. (ii) There exists $\tilde{q} \in (0, 1)$ such that for all $q^* \in (\tilde{q}, 1)$ disclosure is the unique equilibrium.

Proof. (i) In the disclosure equilibrium $E[\bar{q}_\mu(x, d) | q] = E[\bar{q}_{[q^*, 1]}(x) | q]$ and $E[\bar{q}_\mu(x, n) | q] = E[\bar{q}_{[0, q^*]}(x) | q]$ so clearly deviation is never worthwhile. (ii) In any equilibrium without full disclosure and where $q^* < 1$, senders get expected payoff that is strictly less than 1. To see this, note that the only possible equilibrium where a sender gets a payoff of 1 is where a sender of type $q = 1$ makes a report that differs from all other senders. Senders with $q \in [q^*, 1]$ necessarily get payoff less than 1. Since such senders can always report $R_{q=1}$ and get payoff 1, this cannot be an equilibrium.

Now note that for the case $q^* = 1$, a sender with type $q = 1$ gets a strictly greater payoff under full disclosure than any outcome without full disclosure. In this case sender $q = 1$ has a strict incentive to deviate so that the full disclosure equilibrium is unique. By continuity, it must be the case that there is some $\tilde{q} \in (0, 1)$ such that for $q^* > \tilde{q}$, full disclosure is the unique equilibrium. ■

Lemma 2 Let $\hat{q} = \min\{q | E[\bar{q}_{[0, 1]}(x) | q] = q\}$. Such a $\hat{q} \in (0, 1)$ exists and $E[\bar{q}_{[0, 1]}(x) | q] > q$ for $q < \hat{q}$.

¹⁴By definition (see Topkis, 1998) $\bar{q}_{[a,b]}(x)$ is submodular in (x, a) if and only if for $a' > a$, $\bar{q}_{[a', b]}(H) + \bar{q}_{[a, b]}(L) \leq \bar{q}_{[a, b]}(H) + \bar{q}_{[a', b]}(L)$, and is supermodular in (x, b) if and only if for $b' > b$, $\bar{q}_{[a, b']}(H) + \bar{q}_{[a, b]}(L) \geq \bar{q}_{[a, b]}(H) + \bar{q}_{[a, b']}(L)$.

¹⁵In the limit as $a \rightarrow b$ or $b \rightarrow a$ clearly $\bar{q}_{[a,b]}(H) - \bar{q}_{[a,b]}(L)$ shrinks to zero even without this assumption. But if mass is concentrated in one region of the distribution, then this difference might not shrink monotonically as a or b moves over the region.

Proof. Let $h(q) = E[\bar{q}_{[0,1]}(x) | q] - q$. Since $h(q)$ is continuous on $[0, 1]$, $h(0) = E[\bar{q}_{[0,1]}(x) | q = 0] > 0$, and $h(1) = E[\bar{q}_{[0,1]}(x) | q = 1] - 1 < 0$, there exists $q \in (0, 1)$ such that $h(q) = 0$. If \hat{q} is the smallest such q then, since $h(0) > 0$, $h(q) > 0$ for $q < \hat{q}$. ■

The following proposition shows that a countersignaling equilibrium exists for q^* low enough. The proof involves finding a cutoff $\hat{q} > q^*$ where the sender is indifferent between disclosure and nondisclosure. Consider the relative payoffs from nondisclosure and disclosure as this cutoff varies. For very high cutoff values there are so few nondisclosing high types to pool with that nondisclosure is too risky and the payoff from disclosure is higher. If for low cutoff values the payoff from nondisclosure is higher then there must be at least one intermediate cutoff value where the gains are equal. For $\hat{q} = q^*$, the payoff from nondisclosure is $E[\bar{q}_{[0,1]}(x) | q^*]$ and the payoff from disclosure is q^* . So if $E[\bar{q}_{[0,1]}(x) | q^*] > q^*$, the existence of an intermediate cutoff where the marginal type is indifferent is ensured. The proof then checks that types less than this cutoff all prefer disclosure and types above it all prefer nondisclosure.¹⁶

Proposition 3 *A countersignaling equilibrium exists for all $q^* \in (0, \hat{q})$.*

Proof. Let q' be the marginal type willing to disclose in that $\{q | R_q = d\} = [q^*, q']$. Then $E[\bar{q}_\mu(x, d) | q] = E[\bar{q}_{[q^*, q']}(x) | q]$ and $E[\bar{q}_\mu(x, n) | q] = E[\bar{q}_{[0, q^*] \cup (q', 1]}(x) | q]$. Let her net gain from countersignaling be $w(q') = E[\bar{q}_{[0, q^*] \cup (q', 1]}(x) | q'] - E[\bar{q}_{[q^*, q']}(x) | q']$. Note that $w(q^*) = E[\bar{q}_{[0, 1]}(x) | q^*] - q^*$. From the definition of \hat{q} , $E[\bar{q}_{[0, 1]}(x) | q^*] - q^* = 0$ when $q^* = \hat{q}$ and from Lemma 2 $w(q^*) > 0$ when $q^* < \hat{q}$. Since $w(1) = E[\bar{q}_{[0, q^*]}(x) | 1] - E[\bar{q}_{[q^*, 1]} | 1] < 0$ and the expectation functions are continuous on $[q^*, 1]$, by the intermediate value theorem there exists $\hat{q} \in [q^*, 1]$ such that $w(\hat{q}) = 0$.

Now consider such a \hat{q} and suppose that the receiver believes that only types $[q^*, \hat{q}]$ disclose. If $E[\bar{q}_{[q^*, \hat{q}]}(x) | q] \geq E[\bar{q}_{[0, q^*] \cup (\hat{q}, 1]}(x) | q]$ for $q \in [q^*, \hat{q}]$ and $E[\bar{q}_{[0, q^*] \cup (\hat{q}, 1]}(x) | q] \geq E[\bar{q}_{[q^*, \hat{q}]}(x) | q]$ for $q \in [q^*, \hat{q}]$, then a countersignaling equilibrium exists. From Lemma 1, $E[\bar{q}_{[q^*, \hat{q}]}(x) | q] - E[\bar{q}_{[0, q^*] \cup (\hat{q}, 1]}(x) | q]$ crosses zero at most once, and by the definition of \hat{q} , this happens at \hat{q} . So if the crossing is from above, then a countersignaling equilibrium exists. For the crossing to be from below, from the proof of Lemma 1 case (iv) must hold, or $\bar{q}_{[q^*, \hat{q}]}(H) \geq \bar{q}_{[0, q^*] \cup (\hat{q}, 1]}(H)$ and $\bar{q}_{[q^*, \hat{q}]}(L) \leq \bar{q}_{[0, q^*] \cup (\hat{q}, 1]}(L)$ where at least one of these inequalities is strict. Note that there exists some η such that $\eta \bar{q}_{[q^*, \hat{q}]}(x) + (1 - \eta) \bar{q}_{[0, q^*] \cup (\hat{q}, 1]}(x) = \bar{q}_{[0, 1]}(x)$. Since $\bar{q}_{[0, 1]}(x)$ is a weighted average of

¹⁶Note that with no private receiver information there can be a trivial partial disclosure equilibrium where \hat{q} is such that everyone $q \geq q^*$ is indifferent to disclosure or nondisclosure. Such an equilibrium depends on all types $q \in [q, \hat{q}]$ disclosing and all types $q \in (\hat{q}, 1]$ not disclosing despite each type being indifferent.

$\bar{q}_{[q^*, \hat{q}]}(x)$ and $\bar{q}_{[0, q^*] \cup (\hat{q}, 1]}(x)$, it follows that these conditions are equivalent to $\bar{q}_{[q^*, \hat{q}]}(H) \geq \bar{q}_{[0, 1]}(H)$ and $\bar{q}_{[q^*, \hat{q}]}(L) \leq \bar{q}_{[0, 1]}(L)$ where at least one of these inequalities is strict. Together these imply

$$\bar{q}_{[0, 1]}(H) - \bar{q}_{[0, 1]}(L) < \bar{q}_{[q^*, \hat{q}]}(H) - \bar{q}_{[q^*, \hat{q}]}(L).$$

But by the assumption that $\bar{q}_{[a, b]}(x)$ is submodular in (x, a) and supermodular in (x, b) ,

$$\begin{aligned} \bar{q}_{[0, 1]}(H) - \bar{q}_{[0, 1]}(L) &\geq \bar{q}_{[q^*, 1]}(H) - \bar{q}_{[q^*, 1]}(L) \\ &\geq \bar{q}_{[q^*, \hat{q}]}(H) - \bar{q}_{[q^*, \hat{q}]}(L), \end{aligned}$$

a contradiction. So we are left with the conclusion that the crossing must be from above rather than below, so types $q \in [q^*, \hat{q})$ strictly prefer disclosure and types $q \in (\hat{q}, 1]$ strictly prefer nondisclosure. ■

This result can be compared with Feltovich et al. (2002) who consider countersignaling in a three-type game with a costly signal and find conditions under which mediums signal while lows and highs do not. The above result considers a continuum of types in which types above some cutoff have good news and shows a simple condition under which mediocre types above the cutoff reveal the good news while higher types do not. Comparing the two games, countersignaling behavior arises more readily in the disclosure game for several reasons. First, in the signaling game the low type might break the countersignaling equilibrium through costly mimicry of the medium type, but in the disclosure game types below q^* simply do not have good news to reveal and cannot pay to attain it. Second, in the signaling game high types can choose to break the countersignaling equilibrium either by pooling with medium types rather than low types or by sending a higher signal that is too expensive for medium types to mimic. In the disclosure game this latter option does not arise. Finally, in the signaling game there are only three types so any countersignaling equilibrium must take the particular form of the medium type signaling and high and low types not signaling. In the disclosure game with a continuum of types the lower cutoff is fixed by who has good news, but the upper cutoff is flexible so the exact range of types which disclose can vary depending on the information conditions.

We now consider the nondisclosure equilibrium in which no types disclose good news. So far we have not had to consider off-equilibrium-path behavior because in both the disclosure equilibrium and the countersignaling equilibrium both disclosure and nondisclosure can be observed in equilibrium so the receiver never faces unexpected behavior. However in a nondisclosure equilibrium we must consider what the receiver believes when disclosure is observed even though no types are expected to disclose in equilibrium.

The receiver could be agnostic and assume that any observed deviation was completely accidental so that for any $q \in [q^*, 1]$ the likelihood of deviation is determined by the receiver's initial prior, f , and Bayes rule. If all types have the same incentive to deviate then such an assumption is appropriate (Banks and Sobel, 1987). In standard disclosure games without private receiver information the costlessness of the message means that all types do in fact have the same incentive to deviate. However, in our model the incentive varies with type because the receiver considers his own private information in estimating the type of a sender and this information is expected to be more favorable for higher types.

When, as in our game, different types have different incentives to deviate, the literature on refinements argues that receivers will use “forward induction” to infer who might have deviated so beliefs should reflect relative incentives to deviate. The most common refinements that restrict beliefs on this basis are the Intuitive Criterion, divinity, and D1. The Intuitive Criterion states that the receiver should put zero probability on a type having deviated if it would not benefit from deviation under the most favorable possible beliefs about who deviates. Divinity states that if one type benefits from deviation for a smaller set of rationalizable payoffs than another type, less weight should be put on the first type. And D1 states that in such a situation zero weight should be on the first type (Banks and Sobel, 1987).

In standard signaling games which satisfy the single-crossing property, the Intuitive Criterion can sometimes rule out a pooling equilibrium because if someone deviates and signals the receiver might infer that a low type would never have unexpectedly signaled because it would not be worthwhile. If there is no private receiver information in such games, divinity can always eliminate pooling equilibria if there are only two types because the receiver can infer that the higher type can signal at lower cost. By an extension of this logic, D1 can always eliminate any pooling equilibria for any number of types. Because of the increasing power to rule out pooling equilibria by ruling out the possibility that unexpected signaling was from low types, in standard signaling games divinity is a stronger refinement than the Intuitive Criterion and D1 is a stronger refinement than divinity.

In our model the presence of private receiver information and the absence of signaling costs change the role of refinements considerably. Because there are no costs to making a report, the Intuitive Criterion clearly does not restrict any type from disclosing. In fact every type would be very happy to disclose if they would be thought of as the highest type by doing so. Regarding divinity and D1, in a nondisclosure equilibrium higher types expect to be evaluated more favorably than lower types because of the private receiver information, so they must be given

a larger payoff to induce them to deviate. Therefore, agnostic beliefs are not permitted under divinity while skeptical beliefs are, and skeptical beliefs are not just permitted under D1 but are actually *required*.¹⁷

Definition 2 Let $u^*(q)$ be a vector of Perfect Bayesian Equilibrium payoffs for the sender. For any $R \in \{n, d\}$, define

$$J(R) = \{q \mid u^*(q) > 1\}.$$

A pure strategy Perfect Bayesian Equilibrium satisfies the Intuitive Criterion if and only if there does not exist $R \in \{n, d\}$ and $q \in [0, 1]$ such that

$$u^*(q) < \min\{[0, 1] \setminus J(R)\}$$

for $\{[0, 1] \setminus J(R)\} \neq \emptyset$.

As defined, $J(R)$ is the set of types that could never do better by deviating from the equilibrium. According to the Intuitive Criterion, out of equilibrium beliefs must therefore put zero probability on the event that some $q \in J(R)$ deviated. As a result, a type $q \in [0, 1]$ agent who deviates can never expect to get less than $\min\{[0, 1] \setminus J(R)\}$. If, for some agent, this is greater than her expected equilibrium payoff then the equilibrium fails to survive the Intuitive Criterion. Notice immediately that for our application, every type can potentially do better by disclosing so that $J(R)$ is always empty. In this case, the Intuitive Criterion has no power to refine away the nondisclosure (pooling) equilibrium.

Since D1 is stronger than divinity we will simplify the presentation by only considering D1. The definition of D1 requires defining the sets of rationalizable gross payoffs that would give the sender a greater net payoff than candidate equilibrium payoff for deviating.

$$D(q, R) = \{\pi \mid L \leq \pi \leq H \text{ and } u^*(q) < \pi\}$$

$$D^0(q, R) = \{\pi \mid L \leq \pi \leq H \text{ and } u^*(q) = \pi\}$$

where $u^*(q)$ is the equilibrium payoff of a type q sender.

¹⁷Feltovich et al. (2002) find that in the presence of private receiver information D1 can lose its power to ensure a unique equilibrium in signaling games. However, D1 still implies a unique equilibrium if the private receiver information is not too important and signaling costs are decreasing in type at a sufficient rate. In disclosure games the role of signaling costs is not present so the effect of private receiver information always dominates.

Definition 3 *The criterion D1 puts probability zero on any type q disclosing if there is some other type q' such that*

$$\{D(q, R) \cup D^0(q, R)\} \subsetneq D(q', R).$$

The argument is that if there are a wider range of receiver payments rationalizable by some beliefs for which type q' would deviate than type q , then type q' must be the type that deviated. In signaling games this refinement implies that the receiver has a very favorable impression of types who deviate from a non-signaling equilibrium since high types have the lowest signaling costs. It implies the opposite in our disclosure game with private receiver information.

Lemma 3 *In a candidate nondisclosure equilibrium, (i) the Intuitive Criterion places no restriction on beliefs and (ii) D1 implies skeptical beliefs.*

Proof. (i) If beliefs display extreme optimism, $\mu(q | x, d) = 0$ for all $q < 1$, then clearly all types prefer to deviate since $E_\mu[q' | q, d] < 1$ for all $q \in [q^*, 1]$.

(ii) By affiliation of x and q , $E_\mu[q' | q, d]$ is increasing in q so higher types expect higher payoffs in equilibrium. Therefore the set of payoffs that is larger for deviation than for following the equilibrium strategy is decreasing in sender type. ■

These results are clearly supportive of an understatement equilibrium as will be seen in Proposition 4. However, in the presence of private receiver information, these refinements do not fully capture the intuition of forward induction. If the support of receiver beliefs is not a singleton, the sender has to consider her expected payoff from deviating based on receiver beliefs, and this payoff will vary for different types of senders. So an alternative criterion that better captures the intuition of forward induction is whether the set of beliefs supporting deviation is larger for one type than for another.

To do this we first restrict attention to beliefs that are not too unreasonable. As shown in Lemma 1, for any possible receiver beliefs the set of types who are better off from disclosing form an interval that is a subset of $[q^*, 1]$. It seems that as a minimal restriction the receiver should not be permitted beliefs that could never correspond to any pattern of sender behavior regardless of what the sender thought about receiver beliefs. Therefore we restrict sender beliefs to those that are consistent with the strategy profile $\{q | R_q = d\} = [a, b] \subseteq [q^*, 1]$.

Based on this restriction, we define a specialized refinement “expected D1” to examine the nondisclosure equilibrium in our environment with private receiver information. Define the sets of permissible receiver beliefs about who unexpectedly discloses that would give the sender a greater

expected payoff than the expected candidate equilibrium payoff for nondisclosure as

$$C(q) = \{[a, b] \mid [a, b] \subseteq [q^*, 1] \text{ and } E[\bar{q}_{[a,b]}(x) \mid q] > u^*(q)\} \quad (1)$$

$$C^0(q) = \{[a, b] \mid [a, b] \subseteq [q^*, 1] \text{ and } E[\bar{q}_{[a,b]}(x) \mid q] = u^*(q)\} \quad (2)$$

where $u^*(q)$ is the equilibrium payoff of a type q sender. Expected D1 puts probability zero on any type q disclosing if there is some other type q' such that

$$\{C(q) \cup C^0(q)\} \subsetneq C(q').$$

Lemma 4 *In a candidate nondisclosure equilibrium, expected D1 implies skeptical beliefs if $q^* \in (0, \hat{q})$.*

Proof. It is sufficient to show that for all $q' > q^*$, (i) there does not exist $[a, b] \subseteq [q^*, 1]$ such that $E[\bar{q}_{[a,b]}(x) \mid q^*] < E[\bar{q}_{[0,1]}(x) \mid q^*]$ and $E[\bar{q}_{[a,b]}(x) \mid q'] > E[\bar{q}_{[0,1]}(x) \mid q']$ and (ii) there does exist $[a, b] \subseteq [q^*, 1]$ such that $E[\bar{q}_{[a,b]}(x) \mid q^*] > E[\bar{q}_{[0,1]}(x) \mid q^*]$ and $E[\bar{q}_{[a,b]}(x) \mid q'] < E[\bar{q}_{[0,1]}(x) \mid q']$.

(i) Consider any $[a, b] \subseteq [0, 1]$. Note that submodularity of $\bar{q}_{[a,b]}(x)$ in (a, x) implies for $a > 0$,

$$\bar{q}_{[a,1]}(H) + \bar{q}_{[0,1]}(L) < \bar{q}_{[0,1]}(H) + \bar{q}_{[a,1]}(L)$$

and supermodularity of $\bar{q}_{[a,b]}(x)$ in (b, x) implies for $b < 1$,

$$\bar{q}_{[a,1]}(H) + \bar{q}_{[a,b]}(L) > \bar{q}_{[a,b]}(H) + \bar{q}_{[a,1]}(L).$$

Together these imply for $[a, b] \subseteq [0, 1]$,

$$\bar{q}_{[a,b]}(H) - \bar{q}_{[0,1]}(H) < \bar{q}_{[a,b]}(L) - \bar{q}_{[0,1]}(L). \quad (3)$$

Therefore for $[a, b] \subseteq [q^*, 1]$ and all $q' > q^*$,

$$\begin{aligned} E[\bar{q}_{[a,b]}(x) \mid q'] - E[\bar{q}_{[0,1]}(x) \mid q'] &= g(L \mid q')\bar{q}_{[a,b]}(L) + g(H \mid q')\bar{q}_{[a,b]}(H) - g(L \mid q')\bar{q}_{[0,1]}(L) - g(H \mid q')\bar{q}_{[0,1]}(H) \\ &= g(L \mid q')(\bar{q}_{[a,b]}(L) - \bar{q}_{[0,1]}(L)) + g(H \mid q')(\bar{q}_{[a,b]}(H) - \bar{q}_{[0,1]}(H)) \\ &< g(L \mid q^*)(\bar{q}_{[a,b]}(L) - \bar{q}_{[0,1]}(L)) + g(H \mid q^*)(\bar{q}_{[a,b]}(H) - \bar{q}_{[0,1]}(H)) \\ &= E[\bar{q}_{[a,b]}(x) \mid q^*] - E[\bar{q}_{[0,1]}(x) \mid q^*] \end{aligned} \quad (4)$$

where the inequality follows from (3) and the fact that $g(L \mid q)$ is decreasing in q and $g(H \mid q)$ is increasing in q .¹⁸

¹⁸For more general distributions of x this result follows from Topkis (1998) Corollary 3.9.1(a) as long as x is affiliated with q .

(ii) Let $[a', b']$ be such that $E[\bar{q}_{[a', b']}(x) | q^*] = E[\bar{q}_{[0, 1]}(x) | q^*]$. Such an $[a', b'] \subseteq [q^*, 1]$ exists if $q^* < \hat{q}$ by the intermediate value theorem since $E[\bar{q}_{[a', 1]}(x) | q^*] > E[\bar{q}_{[0, 1]}(x) | q^*]$ for all $a' > 0$ and by Lemma 1 $E[\bar{q}_{[a', a']}(x) | q^*] < E[\bar{q}_{[0, 1]}(x) | q^*]$ for a' sufficiently close to q^* . Since $E[\bar{q}_{[a', b']}(x) | q^*] = E[\bar{q}_{[0, 1]}(x) | q^*]$ then $E[\bar{q}_{[a', b']}(x) | q'] < E[\bar{q}_{[0, 1]}(x) | q']$ by (4). Therefore, by continuity, for all $q' > q^*$ there exists $[a, b] \subseteq [q^*, 1]$ such that $E[\bar{q}_{[a, b]}(x) | q^*] > E[\bar{q}_{[0, 1]}(x) | q^*]$ and $E[\bar{q}_{[a, b]}(x) | q'] < E[\bar{q}_{[0, 1]}(x) | q']$. ■

Note that this refinement implies that out of equilibrium receiver beliefs put all weight on the worst type q^* . If we did not restrict receiver beliefs to be consistent with $\{q | R_q = d\} = [a, b] \subseteq [q^*, 1]$ then the refinement would not necessarily imply that all weight be put on the worst type, but it would not rule out such beliefs. In particular, part (ii) of the proof shows that there always exists beliefs such that the worst type is willing to deviate and no other type is willing to deviate. So beliefs that put all weight on this type cannot be ruled out.

Based on Lemmas 2, 3, and 4 we find that the only requirement for a nondisclosure equilibrium to exist and survive these refinements is that q^* is low enough.

Proposition 4 *A nondisclosure equilibrium surviving the Intuitive Criterion, D1, and expected D1 exists [if and only if?] for all $q^* \in [0, \hat{q}]$.*

Proof. From Lemmas 3 and 4 beliefs that put all weight on type q^* survive all three refinements. By Lemma 2 $E[\bar{q}_{[0, 1]}(x) | q^*] > q^*$ for $q^* < \hat{q}$, so type $q = q^*$ will not deviate under these beliefs. Since $E[\bar{q}_{[0, 1]}(x) | q]$ is strictly increasing in q , if type $q = q^*$ does not deviate no other type will deviate. ■

Propositions 3 and 4 show that if standards are set low enough an equilibrium with understatement exists. The following proposition shows how the distribution of sender types affects the potential for understatement. In particular it shows that if there is any common knowledge information that makes the conditional distribution more favorable, then the sufficient conditions for the existence of nondisclosure and countersignaling equilibria in Propositions 3 and 4 are more readily satisfied.

Proposition 5 *Let y be a random variable affiliated with q . If the realization of y is common knowledge then \hat{q} is increasing in y .*

Proof. Let $\bar{q}_\Lambda(x, y)$ be the expected value of q given y and x following the earlier definition of $\bar{q}_\Lambda(x)$. Since y is affiliated with q so $\bar{q}_{[0, 1]}(L, y)$ and $\bar{q}_{[0, 1]}(H, y)$ are increasing in y (Milgrom and

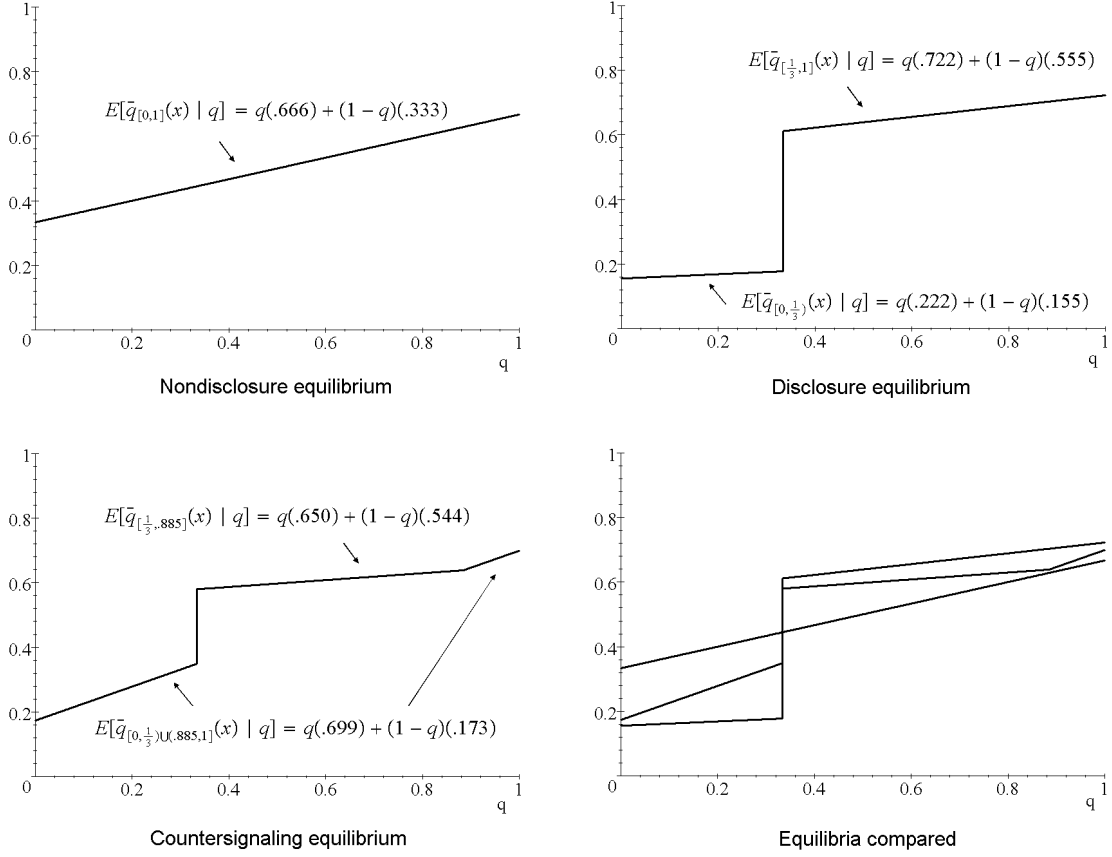


Figure 1: Expected sender payoffs, $f(q) \sim U[0, 1]$, $q^* = \frac{1}{3}$, $g(H | q) = q$.

Weber, 1982). For $y' > y$, $E[\bar{q}_{[0,1]}(x, y') | q] - E[\bar{q}_{[0,1]}(x, y) | q]$ equals

$$g(L | q)\bar{q}_{[0,1]}(L, y') + g(H | q)\bar{q}_{[0,1]}(H, y') - g(L | q)\bar{q}_{[0,1]}(L, y) - g(H | q)\bar{q}_{[0,1]}(H, y)$$

or

$$g(L | q) (\bar{q}_{[0,1]}(L, y') - \bar{q}_{[0,1]}(L, y)) + g(H | q) (\bar{q}_{[0,1]}(H, y') - \bar{q}_{[0,1]}(H, y)) > 0.$$

Therefore $\hat{q}(y) = \min\{q | E[\bar{q}_{[0,1]}(x, y) | q] = q\}$ is increasing in y . ■

3 Examples

To see these results more clearly assume types are distributed uniformly on $[0, 1]$ and $q^* = \frac{1}{3}$. Figure 1 shows expected sender payoffs for the three different equilibria when the probability of

a favorable receiver signal is $g(H | q) = q$. In the nondisclosure equilibrium no types disclose so the receiver must rely only on his private information. The expected payoff is $E[\bar{q}_{[0,1]}(x) | q] = g(H | q)\bar{q}_{[0,1]}(H) + g(L | q)\bar{q}_{[0,1]}(L) = q(.667) + (1 - q)(.333) = .333(1 + q)$ which is increasing in q because the probability that the receiver has favorable private information is increasing in q . In the disclosure equilibrium types $q < q^*$ have nothing favorable to report while types $q \geq q^*$ disclose good news. Therefore the former receive expected payoff $E[\bar{q}_{[0, \frac{1}{3}]}(x) | q]$ and the latter expected payoff $E[\bar{q}_{[\frac{1}{3}, 1]}(x) | q]$. These payoffs are also increasing in q , but most of the increase is from the discrete jump at $q = q^*$ due to the disclosure of good news. Finally, there is a unique countersignaling equilibrium in which types $q > .885$ choose to not disclose the good news that they are above q^* , while types $q \in [\frac{1}{3}, .885]$ disclose it. Therefore low and high types $q \in [0, \frac{1}{3}) \cup (.885, 1]$ receive expected payoff $E[\bar{q}_{[0, \frac{1}{3}) \cup (.885, 1]}(x) | q]$ while intermediate types $q \in [\frac{1}{3}, .885]$ receive expected payoff $E[\bar{q}_{[\frac{1}{3}, .885]}(x) | q]$.

Regarding which equilibrium is most likely, it is sometimes argued that if one equilibrium Pareto-dominates another, then the “good equilibrium” is more likely. In the example of Figure 1 the equilibria cannot be Pareto-ranked since types $q < \frac{1}{3}$ have the opposite ranking of expected payoffs than types $q \geq \frac{1}{3}$. Even types $q \geq \frac{1}{3}$ may not always agree on which equilibrium offers the highest payoffs as seen in Figure 2 where the probability of a favorable receiver signal is $g(H | q) = q^3$. In this example the private receiver information is quite good at separating high and low types, so very high types are nearly indifferent between pooling with all types in the nondisclosure equilibrium or pooling with types $q \geq \frac{1}{3}$ in the disclosure equilibrium. In fact they are so unlikely to be confused with low types that they are even better off in the pictured countersignaling equilibrium in which they pool with low types but are separated from medium types $q \in [\frac{1}{3}, 0.831]$.¹⁹ For other types the disclosure equilibrium offers the highest expected payoffs. Note also that the expected payoffs from countersignaling are higher than from nondisclosure for all types except a small set of types close to $q = .831$ who receive slightly higher payoffs from nondisclosure.

Another way to choose between equilibria is to consider equilibrium refinements. Since both the disclosure and countersignaling equilibria involve some types disclosing and some types not disclosing, off-equilibrium-path beliefs are not relevant. So the only question is whether the

¹⁹There are three countersignaling equilibria in this example. The pictured equilibrium is the middle one. It might have desirable “dynamic stability” properties if receiver beliefs are based on previous sender actions and sender actions are based on previous sender payoffs, but the issue of dynamic stability is quite complicated in this model.

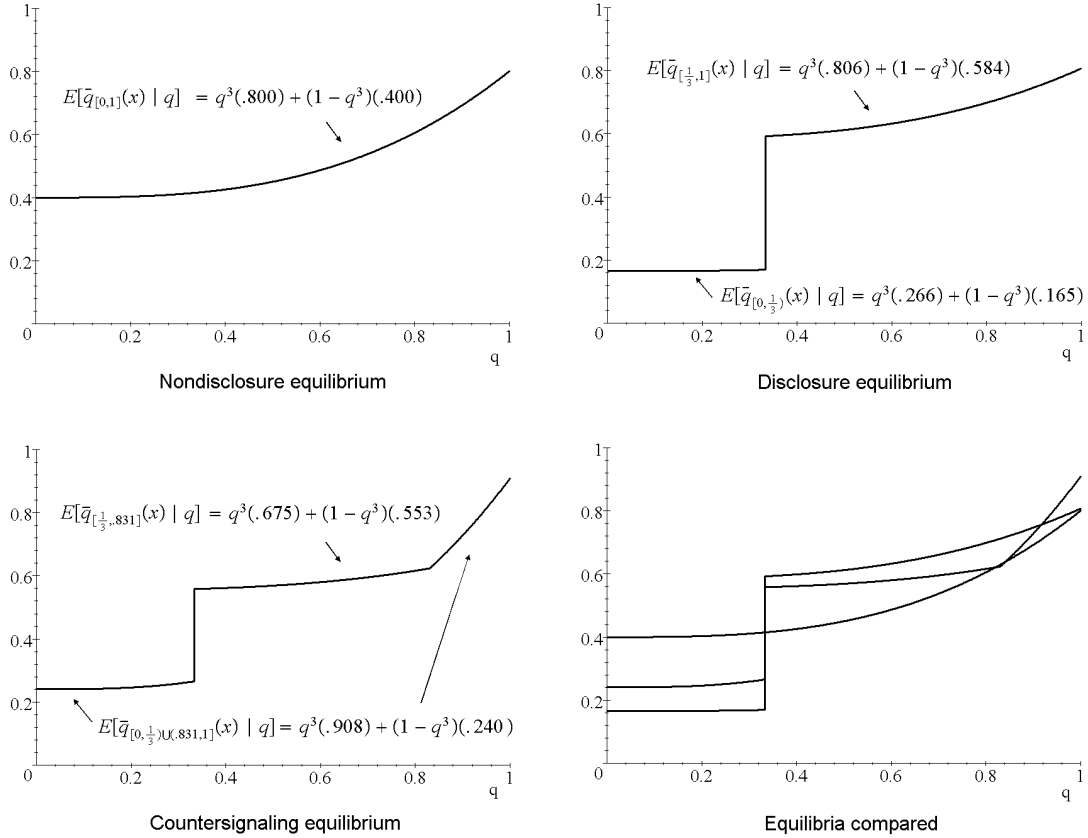


Figure 2: Expected sender payoffs, $f(q) \sim U[0, 1]$, $q^* = \frac{1}{3}$, $g(H | q) = q^3$.

nondisclosure equilibria can be refined away. As shown in Lemma 3, refinements which tend to eliminate low types from consideration as unexpected signalers in signaling games have the opposite effect in our disclosure game with no signaling costs and private receiver information. The Intuitive Criterion does not rule out type $q = q^*$ as the possible source of a deviation, and D1 implies that the receiver should assume that type $q = q^*$ was the source. The reason can be seen in Figures 1 and 2 where the return from nondisclosure is increasing in sender type. Among those who can deviate, type $q = q^*$ receives the lowest payoff from nondisclosure so it has the most incentive to deviate. Therefore, as shown in Proposition 4, the Intuitive Criterion and D1 cannot eliminate the nondisclosure equilibrium.

Standard refinements such as D1 do not fully capture the intuition of forward induction in our model. Private receiver information implies that the payoff from nondisclosure is increasing in

type, but it also implies that the payoff from deviating is increasing in type as long as the receiver has any uncertainty about who might have deviated. Therefore it is not clear who has the most incentive to deviate. Some insight into the question can be seen from Figures 2 and 3. Note that in each case the gap between the payoff for disclosure and nondisclosure is decreasing in q over the disclosure range $[q^*, 1]$ (disclosure equilibrium) and the disclosure range $[q^*, \hat{q}]$ (countersignaling equilibrium). Therefore the marginal type $q = q^*$ has the strongest incentive to disclose in each case. Lemma 4 shows that this pattern is more general - the marginal type $q = q^*$ always has the strongest incentive to disclose for any interval beliefs $[a, b] \subseteq [q^*, 1]$ about who discloses. Therefore lower types are always willing to deviate for a wider range of such beliefs than higher types, so the receiver should react skeptically to any deviation and no type wishes to deviate from the nondisclosure equilibrium. For instance, consider if the receiver believes that an unexpected disclosure came from type $q \in [q^*, b]$ where $b \in [q^*, 1]$. In the example of Figure 1, the marginal type $q = q^*$ will deviate for $b > 0.549$, while the best type $q = 1$ will deviate only if $b > 0.910$. In the example of Figure 2, the marginal type $q = q^*$ will deviate if $b > 0.497$, while the best type $q = 1$ is doing so well under the nondisclosure equilibrium that she will deviate only if $b > 0.991$.

All three types of equilibria exist in these two examples, but as shown in Proposition 2 the disclosure equilibrium is unique if good news is sufficiently rare, and as shown in Propositions 3 and 4 countersignaling and nondisclosure equilibria exist if good news is sufficiently common. In particular, they exist if $q^* \leq \hat{q}$ where $\hat{q} = \min\{q \mid E[\bar{q}_{[0,1]}(x) \mid q] = q\}$. Solving for the example of Figure 1, $\hat{q} = .5$, and solving for the example of Figure 2, $\hat{q} = .432$. Figure 3 shows these respective cases where the middle line in each figure is $E[\bar{q}_{[0,1]}(x) \mid q]$. Clearly for q low enough $E[\bar{q}_{[0,1]}(x) \mid q] > q$ so a sender of type q is better off under nondisclosure than being identified as type q . Therefore, for $q^* < \hat{q}$, marginal type $q = q^*$ prefers nondisclosure if receiver beliefs are pessimistic. And if the marginal type prefers nondisclosure, all higher types prefer nondisclosure also so nondisclosure is an equilibrium.

A key prediction of the model is that understatement is in some sense more likely when there is favorable public information about the sender's type. In particular, Proposition 5 says that if the distribution of sender types is known to be more favorable based on any public information then \hat{q} should be higher. To see this consider if there is an additional binary signal $y \in \{l, h\}$ where $\Pr[h \mid q] = q$ and y is independent of x conditional on q . If $y = h$ ($y = l$) is observed by both the sender and receiver, then the distribution of types conditional on this information is weighted upwards (downwards), so $E[\bar{q}_{[0,1]}(x, h) \mid q] > E[\bar{q}_{[0,1]}(x) \mid q] > E[\bar{q}_{[0,1]}(x, l) \mid q]$, thereby implying $\hat{q}(h) > \hat{q}(l)$ as shown in Proposition 5. Figure 3 shows \hat{q} for the two examples when

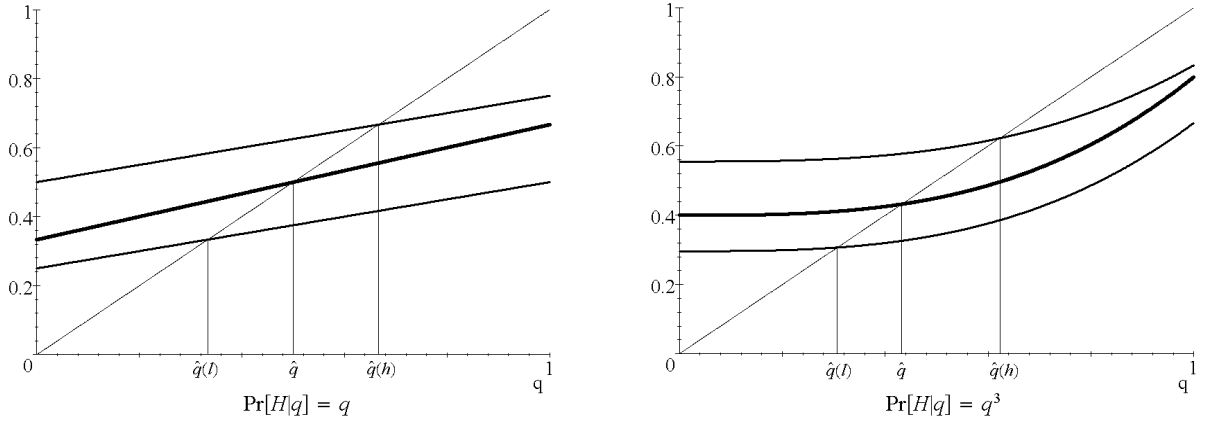


Figure 3: Expected sender payoffs $E[\bar{q}_{[0,1]}(x) | q]$ (thick line), $E[\bar{q}_{[0,1]}(x, h) | q]$ (top medium line) and $E[\bar{q}_{[0,1]}(x, l) | q]$ (bottom medium line), $f(q) \sim U[0, 1]$, $q^* = \frac{1}{3}$

$y = h$ (top line in each figure) and $y = l$ (bottom line in each figure). In the $g(x | h) = q$ example $\hat{q}(h) = 0.666$ and $\hat{q}(l) = 0.333$ and in the $g(x | h) = q^3$ example $\hat{q}(h) = 0.622$ and $\hat{q}(l) = .307$. Since $\hat{q}(l) < \frac{1}{3}$, nondisclosure is not an equilibrium with $g(x | h) = q^3$ for $q^* = \frac{1}{3}$ when there is unfavorable public information. Because the receiver has such a low opinion of the sender, senders will disclose good news even if the receiver views disclosure skeptically.²⁰

4 Conclusion

A large body of research in accounting, finance, and economics concludes that disclosing good news always benefits the sender. In this paper we consider a standard disclosure game assuming that good news is noisy and that the receiver also has private information about sender type. The former assumption is common in the disclosure literature while the latter is not. We show that the presence of *any* private receiver information, no matter how weak, implies that understatement equilibria with nondisclosure by some or all types exist unless good news is restricted to sufficiently high quality senders. Since private receiver information is present in most real world situations, models that abstract away from such information may lead to predicted behavior that is substantially different than observed behavior.

²⁰Note that low $q^* > \hat{q}$ does not rule out the existence of a countersignaling equilibrium in general, although in this example it does.

References

- Banks, Jeffrey S. and Joel Sobel**, “Equilibrium Selection in Signaling Games,” *Econometrica*, 1987, *55*, 647–662.
- Bedard, Kelly**, “Human Capital Versus Sorting Models: University Access and High School Drop-outs,” *Journal of Political Economy*, 2001, *109*, 749–775.
- Clements, Matthew T.**, “Low quality as a signal of high quality,” 2002. Working paper.
- Dye, Ronald A.**, “Disclosure of Nonproprietary Information,” *Journal of Accounting Research*, 1985, *23*, 123–145.
- , “Proprietary and Nonproprietary Disclosures,” *Journal of Business*, 1986, *59*, 351–366.
- , “Investor Sophistication and Voluntary Disclosures,” *Review of Accounting Studies*, 1998, *3*, 261–287.
- Farrell, Joseph**, “Voluntary Disclosure: Robustness of the Unraveling Result, and Comments on its Importance,” in Ronald E. Grieson, ed., *Antitrust and Regulation*, Lexington, MA, and Toronto: Lexington Books, 1986, pp. 91–103.
- Feltovich, Nick, Rick Harbaugh, and Ted To**, “Too Cool for School? Signalling and Countersignalling,” *Rand Journal of Economics*, Winter 2002, *33* (4), 630–649.
- Fischer, Paul E. and Robert E. Verrecchia**, “Reporting Bias,” *The Accounting Review*, 2002, *75*, 229–245.
- Fishman, Michael J. and Kathleen M. Hagerty**, “Mandatory Versus Voluntary Disclosure in Markets with Informed and Uninformed Customers,” *Journal of Law Economics and Organization*, 2003, *19*, 45–63.
- Giovannoni, Francesco and Daniel J. Seidmann**, “Secrecy and Contrary Preferences,” April 2002. working paper.
- Grossman, Sanford J.**, “The Information Role of Warranties and Private Disclosure About Information Quality,” *Journal of Law and Economics*, 1981, *24*, 461–483.
- and **Oliver B. Hart**, “Disclosure Laws and Takeover Bids,” *Journal of Law and Economics*, 1980, *35*, 323–334.

- Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh**, “Disclosure to a Credulous Audience: The Role of Limited Attention,” 2002. working paper.
- Hvide, Hank**, “Education and the Allocation of Talent,” *Journal of Labor Economics*, 2003.
- Jin, Ginger Zhe and Phillip Leslie**, “The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards,” *Quarterly Journal of Economics*, 2003, *forthcoming*.
- Jovanovic, Boyan**, “Truthful Disclosure of Information,” *Bell Journal of Economics*, 1983, *13*, 36–44.
- Korn, Evelyn**, “Voluntary Disclosure of Partially Verifiable Information,” 2003. Working paper.
- Milgrom, Paul**, “Good News and Bad News: Representation Theorems and Applications,” *Bell Journal of Economics*, 1981, *12*, 380–391.
- Milgrom, Paul R. and John Roberts**, “Relying on the Information of Interested Parties,” *Rand Journal of Economics*, 1986, *17*, 18–32.
- and **Robert J. Weber**, “A Theory of Auctions and Competitive Bidding,” *Econometrica*, 1982, *50*, 1089–1122.
- Okuno-Fujiwara, Masahiro, Andrew Postlewaite, and Kotaro Suzumura**, “Strategic Information Revelation,” *Review of Economic Studies*, 1990, *57*, 25–47.
- O’Neill, Barry**, “Nuclear Weapons and the Pursuit of Prestige,” 2002. Working paper.
- Orzach, Ram, Per Baltzer Overgaard, and Yair Tauman**, “Modest Advertising Signals Strength,” *Rand Journal of Economics*, 2002, *33*, 340–358.
- Phillips, Damon J. and Ezra W. Zuckerman**, “A Simple Model of Educational Standards,” *American Economic Review*, 1994, *84*, 956–971.
- Shin, Hyun Song**, “News Management and the Value of Firms,” *Rand Journal of Economics*, 1994, *25*, 58–71.
- , “Disclosures and Asset Returns,” *Econometrica*, 2003, *71*, 105–134.
- Spence, Michael**, “Signaling in Retrospect and the Informational Structure of Markets,” *American Economic Review*, 2001, *92*, 434–459.

Teoh, Siew Hong and Chuan Yang Hwang, “Nondisclosure and Adverse Disclosure as Signals of Firm Value,” *Review of Financial Studies*, 1991, 4 (2), 283–313.

Topkis, Donald M., *Supermodularity and Complementarity*, New Jersey: Princeton University Press, 1998.

Verrecchia, Robert E., “Discretionary Disclosure,” *Journal of Accounting and Economics*, 1983, 5, 179–194.

Viscusi, W. Kip, “A Note on ‘Lemons’ Markets with Quality Certification,” *Bell Journal of Economics*, 1978, 9, 277–279.