False Modesty

When Disclosing Good News Looks Bad*

Dr. Richmond Harbaugh
Indiana University

Dr. Theodore To
Bureau of Labor Statistics

September 2018

Abstract

Is it always wise to disclose good news? Using a new statistical dominance condition, we show that if the receiver has any private receiver information then the weakest senders with good news gain the most from boasting about it. Hence the act of disclosing good news can paradoxically make the sender look bad. Nondisclosure by some or all senders is an equilibrium if standards for the news are sufficiently easy or if prior expectations without the news are sufficiently favorable. Full disclosure is the unique equilibrium if standards are sufficiently difficult or sufficiently fine, or if prior expectations are sufficiently unfavorable. Since the sender has a legitimate fear of looking overly anxious to reveal good news, mandating that the sender disclose the news can help the sender. The model's predictions are consistent with when faculty avoid using titles such as "Dr" or "Professor" in voicemail greetings and course syllabi.

Key words: disclosure, unraveling, countersignaling, persuasion, verifiable message, private receiver information
JEL codes: D82, C78

*For their helpful comments we thank Dan Aaronson, Wouter Dessein, Nick Feltovich, Eric French, Nandini Gupta, Robin Hanson, William Harbaugh, Xinyu Hua, Dmitry Lubensky, Tilman Klumpp, Emre Ozdenoren, Joel Sobel, and Mike Waldman. We also thank conference and seminar participants at the BLISS Conference, the Public Choice Society Conference, the International Industrial Organization Conference, and the Midwest Theory Conference, and at American University, the Chicago Fed, Essex University, George Mason University, and Indiana University. The views expressed are those of the authors and do not necessarily reflect those of the Bureau of Labor Statistics.

Don't be so humble – you're not that great.

– Golda Meir

1 Introduction

If you have good news should you disclose it? According to the classic "unraveling" result, not only is it wise to disclose good news, but it is also necessary to disclose mediocre or even bad news to avoid the perception of having worse news (Viscusi, 1978; Grossman and Hart, 1980; Grossman, 1981; Milgrom, 1981; Milgrom and Roberts, 1986; Okuno-Fujiwara et al., 1990). This result on the power of voluntary disclosure has informed long-running debates over whether to mandate disclosure in consumer product information (Mathios, 2000), financial statements (Greenstone et al., 2006), environmental impact (Powers et al., 2011), and other areas.

However, as suggested by the above quote, inferences about modesty are not always so straightforward. People are sometimes unsure whether to reveal good news, and nondisclosure is often observed in practice (Jin, 2005; Xiao, 2010; Luca and Smith, 2015). Advertisers of high quality products frequently use a "soft sell" approach, donors sometimes make anonymous donations, members of a successful group do not always emphasize their group identity, overachievers are often understated, and offenders sometimes withhold mitigating information rather than "protest too much" or "make excuses."

Most of the literature explains such anomalies by examining why the absence of good news is not treated as evidence of bad news.[1] But what if boasting about good news is itself treated as bad news? Consider whether a restaurant should post its hygiene grade from the local health department. The unraveling result implies that all restaurants should post their grades voluntarily, but many restaurants with even $\boldsymbol{A}$ grades choose not to, and indeed the best restaurants within the $\boldsymbol{A}$ category are less likely to disclose (Bederson et al., 2018). Such nondisclosure has led many communities to require mandatory disclosure rather than rely on voluntary disclosure (Jin and Leslie, 2003). Why is it necessary to require $\boldsymbol{A}$ restaurants to disclose their hygiene grade? Could showing off good news be a bad sign that the restaurant lacks confidence in diners' opinions of it?

These same concerns are faced by individuals. For instance, in environments where titles such as "Dr," "Professor," or "PhD" are common, can use of a title be seen not just as redundant, but as a signal that the person has some reason to fear appearing unqualified? Table 1 summarizes data on the use of titles in voicemail greetings and course syllabi by PhD-holding full-time faculty in 26 economics departments in the same U.S. state (details are in Appendix B). Even though it is costless to use a title, many faculty actively avoid them, for example substituting "instructor" for "professor" on course syllabi. In particular, faculty in the more prestigious universities with

[1] See the literature discussion in Section 2.

Table 1: Faculty usage of formal titles – "Dr.," "Professor," "Ph.D."

	Doctoral Universities	Non-Doctoral Universities
Voicemail title	6	33
Voicemail no title	122	88
Number of faculty	128	121
Syllabus title	1	29
Syllabus no title	123	41
Number of faculty	124	70

doctoral programs are significantly less likely to use a title.[2]

To understand why boasting about good news can be a sign of weakness, we analyze a standard costless disclosure game when there is a continuum of sender types and a finite set of verifiable messages. For instance, there is a pass-fail certificate or a system of letter grades. We show that the classic unraveling result is not robust to allowing noisy private receiver information about sender type.[3] Such information is likely to be present in most disclosure environments, and is equivalent to the sender facing a distribution of different receivers, where better sender types face a higher proportion of receivers with a more favorable prior about the sender. For instance, customers hear more favorable recommendations from friends about a good restaurant than a bad restaurant.

To analyze the interaction of the prior, the sender's message, and the receiver's private information, we develop a new statistical dominance condition that shows when private receiver information implies that the net gain from disclosure is decreasing in the sender's type. This condition ranks the effects of truncations on conditional expectations, and always holds if either the private receiver information is sufficiently strong or sufficiently weak. For intermediate strength the interactions between the information sources can be more complicated, but the condition holds for common functional forms used in examples. The condition implies that skepticism about who discloses is self-confirming in that it gives the weakest sender types who meet a standard for good news the most incentive to show off.

If the standards for good news are sufficiently low or if the prior distribution of sender types is sufficiently favorable, such skepticism implies that there is a *nondisclosure equilibrium* where any type who deviates to disclosure is viewed skeptically. And, analogous to behavior in a (costly) signaling game (e.g., Feltovich et al., 2002), there can be a *countersignaling equilibrium* where lower

[2]The t-statistics are -5.04 and -6.79 for voicemail greetings and course syllabi respectively. The differences are also significant using the non-parametric Mann-Whitney rank test.

[3]Private receiver information has been analyzed in costly signaling games (e.g., Feltovich et al., 2002; Alós-Ferrer and Prat, 2012; Daley and Green, 2014), cheap talk games (Watson, 1996; Chen, 2009; de Barreda, 2010), and Bayesian persuasion games (Kolotilin et al., 2017; Guo and Shmaya, 2017).

types who meet the standard disclose, but higher types who could disclose choose to rely on the private receiver information. These equilibria coexist with the full disclosure (unraveling) equilibrium and survive typical refinements such as the intuitive criterion, D1, and Pareto dominance.

Best response dynamics converge to the different equilibria depending on initial behavior.[4] If enough higher types initially countersignal by not disclosing, the better intermediate types are willing to take a chance and also stop disclosing, which further reduces the gains to disclosure. As more and more medium types join in, eventually the nondisclosure equilibrium is reached in which no types at all disclose. But if initially only the highest types countersignal, then nondisclosure tends to be associated with lower types who do not have good news to disclose. Hence lack of disclosure is more risky for the lower range of high types, and the disclosure region expands. Eventually all types disclose, including the highest types. If initial behavior is intermediate with a large but not too large share of countersignaling high types, play can converge to a countersignaling equilibrium, but only if the private receiver information is strong enough.

When does the classic result of full disclosure due to unraveling still hold? We find three sufficient conditions that ensure that there is unraveling up until some point in any equilibrium, including full unraveling. First, if a standard is sufficiently tough we find that disclosure occurs in any equilibrium since even the most skeptical beliefs about which types disclose still imply that any type meeting the standard does better by disclosing. Second, if prior expectations about sender quality are low enough we also find that disclosure is ensured since even an easy standard is then sufficiently above prior expectations. Third, if standards are sufficiently fine then full unraveling is the unique equilibrium.[5]

Although this paper does not consider optimal information design, the results offer new insight into several issues that have been debated from different perspectives. First, is the long-standing question of when disclosure should be mandatory. The existence of multiple equilibria implies that mandatory disclosure can help reveal information when senders and receivers fail to coordinate on an informative equilibrium. Similarly, having a third party disclose the news can reduce communication problems by allowing the sender to enjoy the benefits of favorable information without the negative inference from disclosing good news.

Second, is the issue of how difficult it should be to meet different standards such as those for school diplomas or other certificates of quality. We show that higher standards are less likely to

[4]The coexistence of multiple stable equilibria can capture the persistence of different cultural traditions for disclosure. For example, professionals in Germany traditionally use full titles, such as "Herr Professor Doktor," while professionals in England are traditionally more understated, including medical doctors who switch from "Dr" to "Mr" upon becoming a member of the Royal College of Surgeons. The multiplicity of equilibria can also capture the strategic uncertainty that firms face about whether it is appropriate to boast in a particular situation – disclosure can help or hurt depending on how it is interpreted, and different interpretations are possible in equilibrium.

[5]Okuno-Fujiwara et al. (1990) establish uniqueness when there is a separate verifiable message available to each type. They allow off the equilibrium path beliefs to be skeptical about which types disclose, even though all types benefit equally from disclosure when sender preferences are state-independent and there is no private receiver information. We show that private receiver information justifies such skepticism.

allow the existence of a nondisclosure or countersignaling equilibrium, so higher standards can paradoxically induce higher certification rates.[6] Moreover, the more favorable the distribution of sender types, the higher the standard that is necessary to ensure disclosure. Hence setting lower standards for groups with an unfavorable distribution and higher standards for groups with a favorable distribution not only divides the conditional distributions in a more informative way, but it can also avoid strategic uncertainty due to multiple equilibria.

Third, the model offers new insight into the question of how fine or coarse standards should be. A large literature shows why an information structure that gives different types the same grade can benefit an information intermediary (e.g., Lizzeri, 1999), the sender (e.g., Kamenica and Gentzkow, 2011), or the receiver (e.g., Harbaugh and Rasmusen, 2016). Our results indicate that, unless disclosure is mandated, the gains from such coarseness may be undermined by the strategic reluctance of senders to disclose good news.

Finally, the model provides necessary and sufficient conditions for the existence of nondisclosure and countersignaling equilibria based on observable properties of the common knowledge distribution of types and standards. In a three-type costly signaling game, Feltovich et al. (2002) predict that higher types are less likely to signal, but do not address the effect of public information. In a two-type costly signaling game, Daley and Green (2014) predict that signaling is less likely when public information about type is more favorable. Here we predict that costless self-promotion is less likely by higher quality senders based on both their actual quality and public measures of quality.

In the following section we review the literature, in Section 3 we provide some simple illustrative examples and in Section 4 we develop a more formal model. For the case of a single standard, we derive conditions for when disclosure may fail or be incomplete in Section 5 and then we allow for multiple standards in Section 6 in order to reassess the classic unraveling result. In Section 7 we conclude the paper.

2 Related literature

Explanations for the failure of unraveling in disclosure games focus on why the absence of good news need not lead a receiver to infer the existence of bad news. Reasons include that messages are costly so disclosure is not always worthwhile (Viscusi, 1978; Jovanovic, 1983; Verrecchia, 1983; Dye, 1986; Levin et al., 2009), there is uncertainty over whether the sender has any news so types with relatively bad news pool with those who have no news (Dye, 1985; Farrell, 1986; Okuno-Fujiwara et al., 1990; Shin, 1994, 2003), and the receiver does not fully understand the game (Dye, 1998; Fishman and Hagerty, 2003; Hirshleifer et al., 2004). Unraveling may also fail when there is not a complete ordering of "good news" (Seidmann and Winter, 1997; Giovannoni and Seidmann, 2007;

[6] We take quality as given, but the same issue arises if quality is generated in part by effort at an earlier point. Costrell (1994) analyzes the tradeoff between forcing higher quality among those who meet the standard against the losses of lower rates of attainment, but this tradeoff assumes that good is news is disclosed.

Mathis, 2008), e.g., high profits might impress investors but also encourage a competitor (Dye, 1986; Okuno-Fujiwara et al., 1990), or high quality may not optimally position a product (Board, 2009; Celik, 2014). If the sender is risk-averse and receiver preferences are uncertain, then pooling the lowest and highest types helps the sender by mitigating negative reactions from a receiver with contrary preferences (Bond and Zeng, 2018).

The idea that different information sources can interact to encourage nondisclosure was first explored by Teoh and Hwang (1991) who analyze a two-period game in which a firm decides whether to immediately disclose news that will eventually be made public by a different source anyway. Our approach is closest to that of Feltovich et al. (2002) who show how private receiver information in a standard signaling game can lead to equilibria in which no types signal or to a countersignaling equilibrium where only medium types signal. In a countersignaling equilibrium low types find it too expensive to signal, high types show off their confidence by not signaling, and medium types are afraid that they will be confused with low types if they do not signal. Daley and Green (2014) vary the strength of the private receiver information to analyze when full separation or pooling occurs. Alós-Ferrer and Prat (2012) investigate this question by varying the rate at which learning about sender type occurs independent of the signal.

We differ from these approaches in our focus on costless disclosure in a standard one-period disclosure game.[7] In the above signaling models, private receiver information can sometimes overwhelm the effects of single-crossing due to signaling costs so that higher types have less rather than more incentive to signal. But in our costless disclosure game, there are no signaling costs so the information effect must dominate even if it is arbitrarily weak.[8] Hence allowing for private receiver information serves as a robustness check on disclosure models that exclude any private receiver information.

We also differ from the related disclosure and signaling literature in that we consider a continuum of sender types. The insight that worse types have more incentive to disclose good news can be seen even in two or three type models.[9] A richer type space allows us to address additional issues, including how shifts in the standard for good news change the equilibrium set, how multiple

[7]In a disclosure or verifiable message game not all messages can be sent by all types, unlike in signaling and cheap talk games (Sobel, 2009). If signaling is prohibitively expensive for lower types but free for higher types then that is equivalent to our disclosure game. Feltovich et al. (2002) provide an example of countersignaling in a disclosure game based on this equivalence, but their analysis focuses on costly signaling.

[8]Single-crossing can also fail if the opportunity cost of signaling is increasing in type (Spence, 2001; Sadowski, 2016), or the direct benefits from the signal are decreasing in type (Hvide, 2003; Orzach and Tauman, 2005; Chung and Eső, 2013). Single-crossing is side-stepped if there are multi-dimensional signals but a one-dimensional state (Orzach et al., 2002; Bagwell and Overgaard, 2005), or a one-dimensional signal but a multi-dimensional state (Bénabou and Tirole, 2006; Araujo et al., 2007; Bracha and Vesterlund, 2017).

[9]Feltovich et al. (2002) provide an example of countersignaling in a disclosure game with three types. As suggested by a referee, even with just two types private receiver information can give low types more incentive to disclose if the good news signal is noisy and hence sometimes available to bad types. In this case not only is disclosure always an equilibrium, but so is nondisclosure since the receiver can skeptically believe good news is from the low type. A countersignaling-like equilibrium with mixed disclosure by high types exists but is not stable using the dynamics approach in this paper.

standards can encourage disclosure, how distributional changes affect disclosure incentives, and how the cutoff where disclosure stops and countersignaling starts is endogenously determined. This endogeneity also allows us to analyze how nondisclosure by higher types can discourage disclosure by middle types, affecting the stability of different equilibria and the convergence to equilibria from different initial conditions. These convergence results show that nondisclosure and disclosure equilibria are always robust even with arbitrarily weak private receiver information, but that countersignaling equilibria require stronger information.

In a Bayesian persuasion model, Guo and Shmaya (2017) show that with private receiver information the optimal information structure pools the lowest and highest types together, so that the receiver relies on the private information to identify which of the pooled senders are sufficiently good. This countersignaling-like structure maximizes the ex ante probability that the sender's type exceeds a cutoff when disclosure is mandatory. In our model, nondisclosure by higher types is driven instead by the ex post decision of each type to maximize the receiver's estimate of their type.

Separate from our focus on private receiver information, other reasons for understatement are also based on the interaction of multiple information sources. Good news in one area can attract attention to bad news in a related area (Lyon and Maxwell, 2004), good news may only be preliminary (Baliga and Sjostrom, 2001), the sender may have information on only a subset of dimensions of interest (Nanda and Zhang, 2008), a reputation for understatement can be useful for times when there is no good news (Grubb, 2011), understatement can encourage further investigation (Mayzlin and Shin, 2011), and standards for good news may be uncertain and inferred in part from who discloses (Harbaugh et al., 2011). If marketing directly increases consumer awareness of product quality, then success of a product without marketing is an impressive signal (Miklós-Thal and Zhang, 2013). Such demarketing can also serve as a price discrimination strategy when there is uncertainty over consumer preferences (Kim and Shin, 2016). If some receivers are more discerning than others and better senders care more about such receivers, then an equilibrium with inconspicuous signals can arise (Carbajal et al., 2015; Hoffman et al., 2018).

3 Examples

Suppose a sender's quality q has unconditional distribution F with uniform density f on $[0, 1]$ and the sender's payoff is her expected quality as estimated by a receiver. Sender types cannot directly reveal their quality q, but if they are above some standard s they can costlessly disclose this fact. Independent of the disclosure decision, higher quality senders are more likely to face a receiver type who views the sender more favorably. To capture this, let the receiver have some private information represented by a noisy binary signal $x \in \{\ell, \hslash\}$ where $\Pr[\hslash|q]$ is increasing in q so the chance of an $\hslash$ signal is higher for better senders. Let the conditional distributions be $H(q) = F(q|x = \hslash)$ and $L(q) = F(q|x = \ell)$. The sender decides to disclose or not based on the average or expected receiver estimate of the sender's quality.[10] For a sender of type $q \in Q$ denote this expectation, which is the sender's expected payoff, by

$$\bar{q}_Q(q) = \Pr[\hbar|q]E_H[q|q \in Q] + \Pr[\ell|q]E_L[q|q \in Q]. \tag{1}$$

Let D be the set of types who are believed to disclose when disclosure is observed and N the set of types who are believed to not disclose when nondisclosure is observed. If the receiver expects all types $q \geq s$ to disclose, then $D = [s, 1]$ and $N = [0, s)$ and clearly $\bar{q}_D(q) > \bar{q}_N(q)$ for all $q \geq s$ who can disclose. Therefore, such beliefs are confirmed in equilibrium.

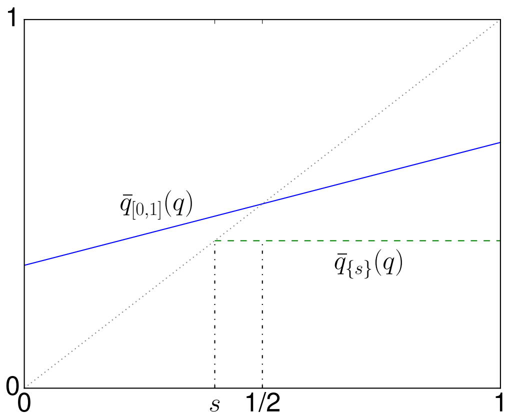
Can withholding good news also arise in equilibrium? First suppose that $\Pr[\hbar|q] = q$ and consider a nondisclosure equilibrium where no types disclose, $N = [0, 1]$. If a sender follows the strategy of not disclosing then the sender's estimated quality is $E_H[q|q \in [0, 1]] = 2/3$ if the receiver observes an $\hbar$ signal and $E_L[q|q \in [0, 1]] = 1/3$ if the receiver observes an ℓ signal, implying the expected payoff from nondisclosure is $\bar{q}_N(q) = q\frac{2}{3} + (1 - q)\frac{1}{3} = 1/3 + q/3$. As seen in Figure 1a, this is increasing in q so higher types receive a higher nondisclosure payoff than lower types due to the private receiver information. Can types who meet the standard s do even better if they deviate and unexpectedly disclose? If the receiver skeptically believes that $D = \{s\}$ so that any unexpected disclosure came from the worst type who can do so, then $\bar{q}_D(q) - \bar{q}_N(q) = s - 1/3 - q/3$. Therefore if the standard is sufficiently low, $s < 1/2$, the marginal sender $q = s$ loses from disclosure and higher types lose even more, so nondisclosure is an equilibrium.[11]

Skepticism about who discloses makes sense because, unlike in a standard costly signaling game, worse types have more incentive to disclose than better types. Since the worst senders meeting the standard gain the least from following the equilibrium strategy of nondisclosure, they will deviate and disclose for a wider range of belief-supportable payoffs for disclosure than other senders. Therefore standard equilibrium refinements imply that the receiver should put more weight on a deviation having come from a worse rather than better type, so skeptical beliefs are appropriate and the equilibrium cannot be refined away. Moreover, in a learning model we will show that if beliefs start out sufficiently skeptical they become more skeptical over time until play converges to the nondisclosure equilibrium.

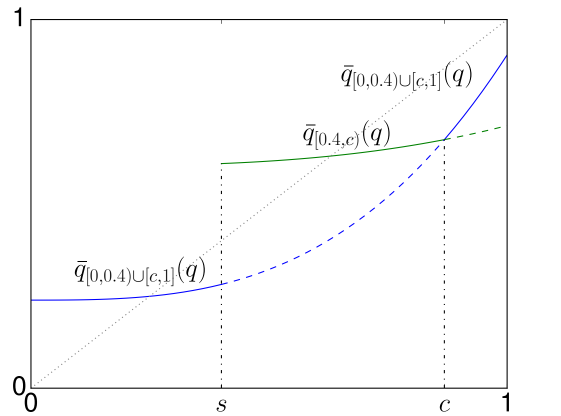
If the standard s is sufficiently low to permit a nondisclosure equilibrium, it also permits a countersignaling equilibrium where the receiver correctly believes that types $D = [s, c)$ disclose and types $N = [0, s) \cup [c, 1]$ cannot or do not disclose for some type $c \in [s, 1)$ who is just indifferent between disclosing or not. Such a c must exist because if c is very low then beliefs about who discloses are so pessimistic that type c strictly prefers nondisclosure, and if c is very high then

[10]The average interpretation is appropriate if the sender faces a population of receivers where proportion $\Pr[\hbar|q] = q$ have a more favorable prior H and the rest have a less favorable prior L.

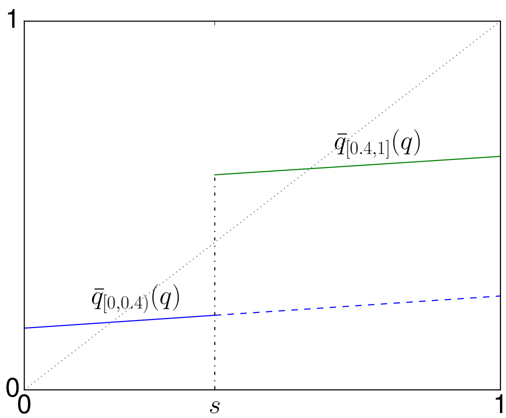
[11]That the net gain to disclosure $\bar{q}_D(q) - \bar{q}_N(q)$ is decreasing in q is simplified here by the assumption that D is a singleton. We will show more generally in Lemmas 1–3 of Appendix A conditions under which private receiver information ensures $\bar{q}_D(q) - \bar{q}_N(q)$ is decreasing.



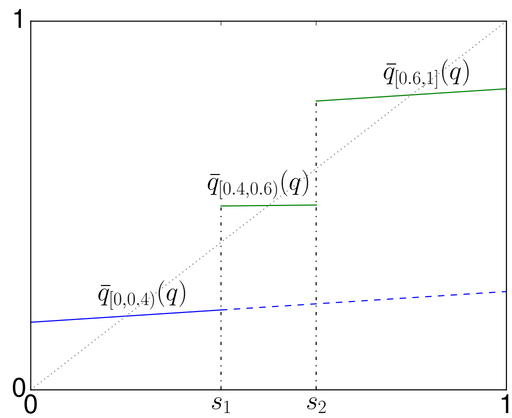
(a) Nondisclosure equilibrium: f uniform, $\Pr(H|q) = q$, $s = 2/5$



(b) Countersignaling equilibrium: f uniform, $\Pr(H|q) = q^3$, $s = 2/5$



(c) Unique disclosure equilibrium: $f(q) = 2 - 2q$, $\Pr(H|q) = q$, $s = 2/5$



(d) Unique unraveling equilibrium: f uniform, $\Pr(H|q) = q$, $s_1 = 2/5$, $s_2 = 3/5$

Figure 1: Expected payoffs as a function of q for different equilibria

beliefs are so optimistic that type c strictly prefers disclosure. Figure 1b shows a countersignaling equilibrium for the same example except the private receiver information better differentiates high types, $\Pr[\hbar|q] = q^3$. Types who just meet the standard $s = 2/5$ benefit the most from disclosure, and the net gain from disclosure $\bar{q}_D(q) - \bar{q}_N(q)$ falls until type $c = 0.87$ is indifferent. Higher types then benefit more from countersignaling and relying on the receiver's prior information to stochastically differentiate them from the lowest types. This example also has nondisclosure and disclosure equilibria with different sender types benefiting more from the different equilibria, so no equilibrium is Pareto dominant.[12] As shown in the next section, a wide range of initial behavior converges to this countersignaling equilibrium.[13]

If the standard s is sufficiently high, or if the prior distribution is sufficiently unfavorable, disclosure is the unique equilibrium. In this case even the most skeptical belief that $D = \{s\}$ makes all types want to disclose, so skepticism is not sustainable in equilibrium. In the example of Figure 1a if $s > 1/2$ instead of $s = 2/5$ then in the unique equilibrium all types $[s, 1]$ disclose. Or if the prior distribution is $f(q) = 2 - 2q$ as in Figure 1c then disclosure is the unique equilibrium even for $s = 2/5$. For instance, if there is some public information about the sender that makes the receiver think the sender is unlikely to be of high quality, then even the skeptical belief that a type who discloses is only $q = 2/5$ is still relatively favorable so all types disclose.[14]

Disclosure is also ensured if multiple standards divide up the type space sufficiently finely. In Figure 1d the prior distribution is again uniform, but now there are two standards, $s_1 = 2/5$ and $s_2 = 3/5$. Based on the prior analysis types $q \geq 3/5$ always disclose since the standard they meet is so high. But given that they disclose, in any equilibrium where lower types do not disclose the sender is at best in the region $[0, 3/5]$. Conditional on being in this region, the standard $s_1 = 2/5$ is now relatively high, so the remaining types $q \in [2/5, 3/5)$ disclose even though they would not always disclose with a single standard $s = 2/5$. For instance, if there is just a pass grade indicating $q \geq 2/5$ then disclosure of the pass grade is not ensured, but if there is an A grade indicating $q \geq 3/5$ and a B grade indicating $q \in [2/5, 3/5)$ then disclosure of both grades is ensured.

In the following two sections we formalize and generalize these ideas. We provide conditions under which private receiver information implies that the net gain from disclosure is decreasing in the sender's type. We then use this result to show when refinements permit and even require pessimistic beliefs that support a nondisclosure equilibrium, and when best response dynamics converge to pessimistic beliefs that support nondisclosure or countersignaling. We then show how finer certification standards and tougher certification standards can be used to ensure uniqueness

[12]For the receiver, countersignaling equilibria sometimes provide more information to the receiver than a full disclosure equilibrium, e.g., in this example $c = 0.98$ is also a countersignaling equilibrium and is the most informative equilibrium.

[13]In the previous example with $\Pr[\hbar|q] = q$ a countersignaling equilibrium also exists but, based on arguments in the next section, it is not stable and play will converge to either nondisclosure or disclosure. Countersignaling with uniform f and $s = 2/5$ is stable for $\Pr[\hbar|q] = q^n$ for $n \geq 2$, as seen in Figure 2b for the case where $n = 3$.

[14]However, nondisclosure is still an equilibrium if the standard is sufficiently low, which is $s < 1/3$ in this example.

of disclosure equilibria.

4 The model

In this sender-receiver game the sender has quality q distributed according to the smooth distribution F with density f which has support on $[0,1]$. The sender knows the realized value of q and sends a type-restricted message v to the receiver that is potentially informative about q. The receiver does not know q but has his own noisy binary signal $x \in \{\ell, \hbar\}$ where $\Pr[\hbar|q]$ is smooth and strictly increasing.[15] The conditional distributions $H(q)$ and $L(q)$ given these signals have respective densities $h(q)$ and $l(q)$.

We assume that $H(q)$ dominates $L(q)$ in the sense that $E_H[q|q \in [a,b]] - E_L[q|q \in [a,b]]$ is strictly decreasing in a and strictly increasing in b,[16] so that the receiver's information has more impact as the interval $[a,b]$ expands. Lemma 2 in Appendix A provides two sufficient conditions for such dominance – that $h(q)$ is increasing and $l(q)$ is decreasing, or that the generalized failure rate ratio $(h(q)/(H(q')-H(q)))/(l(q)/(L(q')-L(q)))$ is increasing in q for all $q, q' \in [0,1]$.[17] The first condition requires that the receiver's information is in a sense sufficiently strong relative to the prior,[18] while the second condition can hold even with weak receiver information and a strong prior. Indeed, as shown in Lemma 3 in Appendix A, it holds for any logconcave F if the receiver's information is sufficiently weak in the sense that $\Pr[\hbar|q]$ is sufficiently flat.

The sender first learns her type q and then sends the message v. After learning x and hearing v the receiver then takes an action α. Following standard assumptions in the sender-receiver game literature, we assume that the receiver's payoff $u^R(q,\alpha)$ is maximized when the receiver's action α equals the receiver's estimate of the sender's type, e.g., $u^R(q,\alpha) = -(q-\alpha)^2$, and that the sender's state-independent payoff, u^S, is strictly increasing in α. We will present the model for $u^S = \alpha$ since all of the proofs extend trivially to the strictly increasing case.[19] Our nondisclosure and uniqueness results extend to more general u^R as long as the sender's expectation of the receiver's action α is increasing in the sender's quality q.[20]

[15] A binary signal simplifies the analysis and highlights that even weak private receiver information can allow for equilibria with nondisclosure.

[16] We refer to these as decreasing Mean Residual Life (MRL) differences and increasing Mean Time to Failure (MTTF) differences respectively. See Shaked and Shanthikumar (2007) for related dominance conditions and orderings. Our results extend partially to ratios rather differences as noted in Footnote 30. Jewitt (2004) shows that the MRL-MTTF gap $E[q|q \in [c,1]] - E[q|q \in [0,c]]$ is increasing (decreasing) in c for f strictly decreasing (increasing), with application to binary signaling games (e.g., Bénabou and Tirole, 2006, 2011). Our new dominance condition implies the MRL-MTTF gap is increasing faster for the dominated distribution, which is also relevant for such games.

[17] Neither condition implies the other and at least one is satisfied in the above examples. For $a = 0$, the assumption on the (reverse) hazard rate ratio is equivalent to geometric dominance (Noe, 2017).

[18] If the prior is very uninformative the condition can be met even with weak receiver information. For instance, it holds for F uniform even if $\Pr[\hbar|q]$ is arbitrarily close to being flat.

[19] We are not analyzing ex ante maximization of sender payoffs as in the Bayesian persuasion literature, so the exact shape of u^S is not relevant.

[20] This holds by Theorem 2 of Athey (2002) if $u^R(q,\alpha)$ satisfies the single-crossing property and x and q are affiliated,

We consider only pure strategy equilibria so a strategy is a mapping between types and messages. Let the conditional cumulative distribution function $\mu(q|x, v)$ represent receiver beliefs about the sender's type given the message v and private information x. Our equilibrium concept is that of a pure-strategy perfect Bayesian equilibrium.

Definition 1 *A* pure-strategy perfect Bayesian equilibrium *is given by a verifiable message profile $v(q)$, a receiver action profile $\alpha(x, v)$, and receiver beliefs $\mu(q|x, v)$ where:*

i) For all q, $v(q) \in \arg\max_{v'} E[u^S(\alpha(x, v'))|q]$;

ii) For all x and v, $\alpha(x, v) = \arg\max_{\alpha'} E_\mu[u^R(q, \alpha')|x, v]$;

iii) $\mu(q|x, v)$ is updated from the sender's strategy and F using Bayes' rule whenever possible.

Condition i) requires that the sender's message is a best response to the receiver's expected actions. Condition ii) requires that the receiver's action is a best response to the sender's message. Condition iii) 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. Since we are considering pure strategy equilibria, if Q is the set of types who are believed to have sent message v when it is observed, then $\mu(q|x, v) = \int_{t\in[0,q]\cap Q} f(t|x)dt / \int_{t\in Q} f(t|x)dt$, implying $\bar{q}_Q(q)$ is as defined in equation (1).

5 Single standard

We start with the case of a single pass-fail standard as introduced in the examples. For this standard, we assume that there is a "blank" message v_0 for nondisclosure and message v_1 for disclosure. Message v_0 is always sent by types $q \in [0, s)$ and either v_0 or v_1 may be sent by types $q \in [s, 1]$.[21] For instance, a person has a certificate to prove that she passed an exam (but nothing to prove that she failed it). In Section 6 we consider multiple standards.

As in the example, let N be the set of types who are believed to have sent v_0 when it is observed and D be the set of types who are believed to have sent v_1 when it is observed. The net gain from disclosure $\bar{q}_D(q) - \bar{q}_N(q)$ is monotonic for any D and N by the assumption that $\Pr[\hbar|q]$ is monotonic, so sender best responses for any receiver beliefs involve: (i) no types disclose, (ii) types in an interval $[s, c)$ disclose for some $c < 1$, (iii) types in an interval $[d, 1]$ disclose for some $d \in [s, 1]$. Case (i) corresponds to nondisclosure, case (ii) corresponds to countersignaling, and case (iii, $d = s$) corresponds to full disclosure. In the following, we examine the conditions under which each of these cases can be an equilibrium. Case (iii, $d > s$) cannot arise in equilibrium since, with

which holds for $\Pr[\hbar|q]$ increasing.

[21] The results are not qualitatively affected if instead sender types $q \geq s_1$ cannot disclose with probability ε for small $\varepsilon > 0$. Senders with no message to disclose are forced to pool with others who do not disclose.

the corresponding beliefs, every type will prefer to disclose. Hence (i), (ii) and (iii, $d = s$) are the only pure strategy equilibrium strategies that are possible.

Since the disclosure equilibrium always exists by the same argument as for the example, our focus is on the existence of nondisclosure and countersignaling equilibria. For existence of a nondisclosure equilibrium, consider the most pessimistic beliefs about who unexpectedly discloses, so $D = \{s\}$ and $N = [0, 1]$ and $\bar{q}_D(q) - \bar{q}_N(q) = s - \bar{q}_{[0,1]}(q)$. Let

$$\hat{q} = \min\{q : q = \bar{q}_{[0,1]}(q)\}, \tag{2}$$

where existence of $\hat{q}$ follows from the fact that $\bar{q}_{[0,1]}(q)$ is continuous and has range in $[0, 1]$. If $s \leq \hat{q}$ then type s will not want to disclose and all higher types have even less incentive to disclose since $\bar{q}_{[0,1]}(q)$ is increasing in q, so nondisclosure is an equilibrium.

Regarding existence of a countersignaling equilibrium, in such an equilbrium $N = [0, s) \cup [c, 1]$ and $D = [s, c)$. Suppose that $s < \hat{q}$, which, as shown above, holds if s is sufficiently small. By the definition of $\hat{q}$ it must be that $s < \bar{q}_{[0,1]}(s)$. Therefore, for c sufficiently close to s, it must be that $\bar{q}_D(c) < \bar{q}_N(c)$. Similarly, since $\bar{q}_{[s,1]}(c) > \bar{q}_{[0,s)}(c)$, it must be that $\bar{q}_D(c) > \bar{q}_N(c)$ for c sufficiently close to 1. By continuity, these imply that there exists some $c \in (s, 1)$ such that

$$\bar{q}_{[s,c)}(c) = \bar{q}_{[0,s)\cup[c,1]}(c). \tag{3}$$

Given the indifference of type c, for this to be an equilibrium we need to show that types below c disclose and types above c do not disclose. It is sufficient to show that $\bar{q}_{[s,c)}(q) - \bar{q}_{[0,s)\cup[c,1]}(q)$ is decreasing in q. Noting the equality (3), this holds by Lemma 1 and can seen in Figure 1b. Therefore if s is sufficiently small that a disclosure equilibrium exists, then so does a countersignaling equilibrium.

Now consider when disclosure is the unique equilibrium. For sufficiently large s, the sender's expected payoff from non-disclosure in any candidate non-disclosure or countersignaling equilibrium is strictly bounded above by s but for any s, the sender's disclosure payoff is bounded below by s. Hence if s is sufficiently close to 1 then every sender $q \geq s$ prefers to disclose and neither nondisclosure nor countersignaling can be an equilibrium. Since disclosure, nondisclosure and countersignaling are the only possible pure strategy equilibria with a single standard, disclosure is unique.

Note that a sufficiently high s for a given F is equivalent to a sufficiently large $F(s)$ for a given s, i.e., a sufficiently tough standard is equivalent to a sufficiently unfavorable distribution. Similarly a sufficiently weak standard is equivalent to a sufficiently favorable prior distribution. The following proposition collects these results.

Proposition 1 (Existence) *(i) A nondisclosure equilibrium exists if s or $F(s)$ is sufficiently small. (ii) A countersignaling equilibrium exists if a nondisclosure equilibrium exists. (iii) A* *disclosure equilibrium always exists and is unique if s or $\boldsymbol{F}(s)$ is sufficiently large. (iv) No other pure strategy equilibria are possible.*

Robustness

The nondisclosure equilibrium is based on skeptical beliefs about which sender types unexpectedly disclose. Are these beliefs reasonable based on "forward induction" arguments about which types have the strongest incentive to deviate? The Intuitive Criterion restricts the receiver to put zero probability on a type having deviated if it would not benefit from deviation under the most favorable possible beliefs about who deviates. But in this game every type would be happy to disclose if they would be thought of as the highest type, so beliefs are unrestricted and hence skeptical beliefs survive.

The D1 condition requires, in our context, that if one type benefits from a deviation for a set of rationalizable receiver best responses that is a subset of that for some other type, zero weight should be put on the former type (Cho and Kreps, 1987; Cho and Sobel, 1990; Ramey, 1996). In a nondisclosure equilibrium higher types expect to be evaluated more favorably than lower types because of the private receiver information, so they have less incentive to deviate than lower types. Therefore, not only does D1 have no power to refine away the nondisclosure equilibrium, it actually reinforces it by dictating that *out-of-equilibrium actions must be viewed skeptically.*[22] Without private receiver information this argument, and the best response dynamics analyzed below, do not support skepticism. In that case the incentive to deviate from nondisclosure is the same for each type, so the receiver should maintain his original priors concentrated on the range of types who can send the verifiable message, leading all senders to deviate from the nondisclosure equilibrium. The proof, and all subsequent proofs, is in Appendix A.

Proposition 2 (Refinements) *The nondisclosure, countersignaling and disclosure equilibria are all robust to the Intuitive Criterion and D1.*

Nondisclosure on and off the equilibrium path is reasonably inferred to be from higher types, so D1 does not refine away any of the equilibria. However, in our model with a continuum of types, the cutoff for nondisclosure by high types is endogenously determined, so there is still the question of whether this cutoff is stable. When higher types countersignal, do they induce more intermediate types to also not disclose? Or does disclosure by intermediate types induce some of the weaker high types to start disclosing? Do such adjustments settle down so that countersignaling is a stable equilibrium, or does behavior converge instead to full disclosure or nondisclosure?

To understand these questions, we now investigate when the different equilibria can arise as the stable outcome of best response dynamics. Suppose that in each "period" the sender's strategy

[22] Daley and Green (2014) apply D1 to a finite-type signaling game with private receiver information by focusing on sets of receiver beliefs rather than rationalizable best responses. In our context without signaling costs such an approach would seem to also generate skeptical beliefs, but we leave this question open.

is a best response to receiver beliefs, where these beliefs in any period after the first period equal the previous best response strategy of the sender. For example, if prior period best responses have $q \in \boldsymbol{D}$ disclosing and $q \in \boldsymbol{N}$ not disclosing then in the current period, the receiver expects that $q \in \boldsymbol{D}$ disclose and $q \in \boldsymbol{N}$ do not. If it is a best response for no types to disclose, we assume for simplicity that beliefs in the next period have $\boldsymbol{D} = \{s\}$.[23] Recall that the monotonicity of $\bar{q}_D(q) - \bar{q}_N(q)$ implies that for *any* receiver beliefs there are three cases for sender best responses. These cases can be ranked in terms of "modesty" by who discloses: (i) nondisclosure is most modest, (ii) countersignaling behavior is less modest and decreasingly modest for higher c and hence a larger disclosure range $[s, c]$, and (iii) full disclosure by all $q \in [s, 1]$) is least modest.[24] Thus consider $c \in [s, 1]$ to be our inverse measure of modesty where $c = s$ represents the highest level of modesty and $c = 1$ represents the least.

With this measure of modesty we have the following convergence result based on initial play by the sender that is a best response to any receiver beliefs. For the result on countersignaling, we say that private receiver information is "too weak" if $Pr[\hbar|q]$ is sufficiently close to being constant as formalized in Lemma 3.

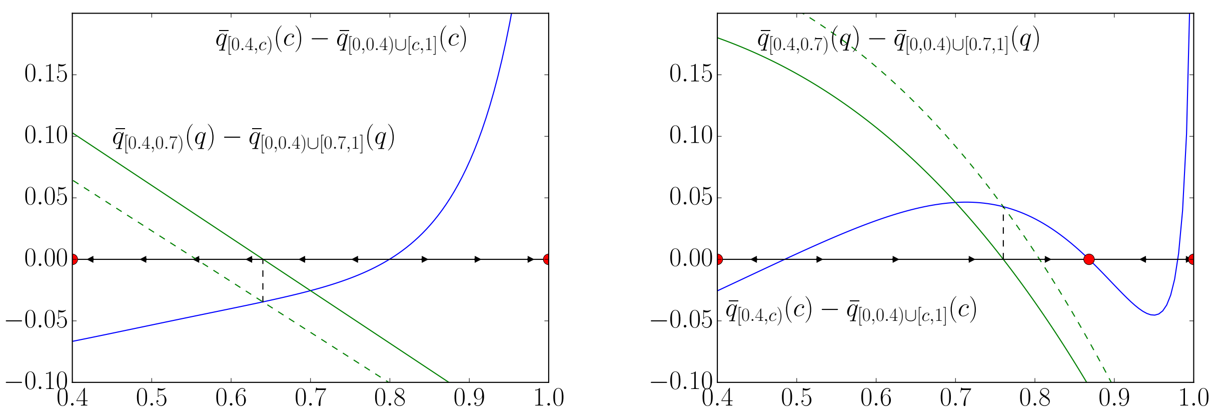
Proposition 3 (Convergence) *If nondisclosure, countersignaling, and disclosure equilibria coexist, then for initial behavior $\boldsymbol{D} = [s, c)$, (i) play converges to the nondisclosure equilibrium if initial behavior is sufficiently modest, i.e., c is sufficiently close to s; (ii) play converges to a countersignaling equilibrium only if the private receiver information is not too weak and initial behavior is intermediate; (iii) play converges to the disclosure equilibrium if initial behavior is sufficiently immodest, i.e., c is sufficiently close to 1.*

The convergence to different equilibria can be seen in Figure 2a based on the example of Figure 1a. The blue line shows the net gain to disclosure for type c as c varies between $s = 0.4$ and 1 when types $\boldsymbol{D} = [s, c)$ disclose. If initial behavior is sufficiently modest in the sense that the countersignaling region is large and $\boldsymbol{D}$ is small, then the corresponding beliefs will induce more medium types to countersignal in the next period, thereby leading to even more pessimistic beliefs about who discloses and less disclosure, until no types at all disclose. For example, the dashed line shows the first step of this convergence when $c = 0.7$. If the receiver forms beliefs accordingly, then the gain from disclosure is $\bar{q}_{[s,\ 0.7)}(q) - \bar{q}_{[0,s)\cup[\ 0.7,1]}(q)$ which is negative for all types $q > 0.64$ as seen from the green line, so in the next period the countersignaling region expands and $c = 0.64$. With these beliefs, by the same logic types $q > 0.56$ will not disclose so $c = 0.56$, and then finally $c = 0.44$ after which beliefs are so pessimistic that all types stop disclosing and the nondisclosure

[23] As with equilibrium refinements, when disclosure is not observed, the receiver must still form expectations in the event that disclosure is subsequently observed. We need only assume that beliefs in the next period are sufficiently skeptical that nondisclosure remains the best response. Unchanged beliefs or $\boldsymbol{D} = \{s\}$ are the most straightforward such beliefs.

[24] For the best response dynamic, we do not consider disclosure by types $[d, 1]$ where $d > s$ since the best response to such beliefs is $\boldsymbol{D} = [s, 1]$ (see the second paragraph in Section 5).

Figure 2: Convergence to nondisclosure, countersignaling, and disclosure



(a) Nondisclosure and disclosure: f uniform, $s = 2/5$, $\Pr[\hbar|q] = q$

(b) Nondisclosure, countersignaling and disclosure: f uniform, $s = 2/5$, $\Pr[\hbar|q] = q^3$

equilibrium is reached. Conversely, if initial behavior is such that the countersignaling region is small, $c > 0.8$, then beliefs in the next period are sufficiently optimistic about disclosure that c increases as more and types disclose until the disclosure equilibrium is reached.

Countersignaling plays an important role in the dynamics of disclosure in this example, but play converges to either disclosure or nondisclosure. With stronger information, countersignaling equilibria can be stable as seen in Figure 2b based on the example of Figure 1b. Play converges to the nondisclosure equilibrium if initial play is sufficiently modest that types up to $c < 0.48$ disclose, and play converges to the disclosure equilibrium if initially all types up to $c > 0.98$ or higher disclose. Within this broad range, play converges to the middle countersignaling equilibrium where only types $[0.4, 0.87)$ disclose. The dashed line shows the first step of this convergence for the case where initially $c = 0.7$. With corresponding beliefs by the receiver, $\bar{q}_{[s,0.7)}(q) - \bar{q}_{[0,s)\cup[0.7,1]}(q) > 0$ for all types $q < 0.76$, so in the next period types $c = 0.76$. By this process beliefs and behavior converge to the middle countersignaling equilibrium.[25]

6 Multiple standards

We now allow for $N \in \mathbb{N}^+$ verifiable messages that disclose a subinterval of the sender's typespace, e.g., a system of certificates or letter grades. In particular, we assume that the typespace is partitioned into $N+1$ nonempty subintervals by a set of strictly increasing standards $\{s_1, s_2, \ldots, s_N\}$ and that the sender can send the verifiable message $v = v_j$ if and only if $q \in [s_j, s_{j+1})$ for $j =$

[25] Best responses may "overshoot" the countersignaling equilibrium if $\bar{q}_D(q) - \bar{q}_N(q)$ is too flat in the neighborhood of the equilibrium, though this is not possible if $h(q)$ is rising and $l(q)$ is falling as in this example.

$1, 2, \ldots, N$.[26,27] Message v_0 is always sent by types $q \in [0, s_1)$ and may also be sent by higher types. Therefore a message profile is $v(q) \in \{v_0, v_j\}$ for $q \in [s_j, s_{j+1})$ and $j = 0, 1, \ldots, N$. We refer to sending v_0 as "not disclosing" and to sending any other message v as "disclosing."

As demonstrated above, the model admits to multiple equilibria. In addition to disclosure, non-disclosure, and countersignaling equilibria, for $N \geq 2$ there can be equilibria with non-disclosure of lower standards but not higher standards, and countersignaling equilibria with disclosure of intermediate standards. Given the multitude of equilibria, we focus on sufficient conditions for when a type must disclose and for when nondisclosure exists. As with a single standard, it is always an equilibrium for all types who can disclose to disclose.

Extending the argument for the single standard case, we expect that nondisclosure arises when s_j is relatively low so revealing good news is not so impressive. To check this intuition, we look for sufficient conditions on s_j such that an equilibrium exists in which v_j and any worse news is not disclosed. Consider the minimum value of q such that the expected payoff for a sender of type q is equal to q when news of v_j and lower is not disclosed. Generalizing (2), let

$$\hat{q}_j = \min\{q : q = \bar{q}_{[0,s_{j+1})}(q)\}, \tag{4}$$

where the existence of $\hat{q}_j$ follows from the fact that $\bar{q}_{[0,s_{j+1})}(q)$ is continuous in q and has range in $[0, s_{j+1}]$.[28] If the receiver skeptically believes that a sender who deviates from nondisclosure is of the lowest type who could deviate, then the highest payoff from disclosure of news v_k for $k \leq j$ is s_k. Therefore, nondisclosure is clearly an equilibrium if $s_j < \hat{q}_j$.

The following proposition summarizes these results, with details of the proof in Appendix A.

Proposition 4 (Existence – multiple standards) *(i) An equilibrium with nondisclosure by types $q \in [0, s_{j+1})$ exists if standard j is sufficiently low, $s_j \leq \hat{q}_j$. (ii) A full disclosure equilibrium always exists.*

Under our dominance condition, similar convergence results for disclosure and nondisclosure equilibria with a single standard apply to multiple standards. For example, given $c_j < s_{j+1}$ but sufficiently close to s_{j+1} for $j \geq k$ (i.e., behavior is sufficiently immodest within each interval), the best response dynamic will converge to disclosure for $j \geq k$. Similarly for nondisclosure. Of course, with multiple standards, many other stable outcomes are possible.

In the classic disclosure model without private receiver information, full disclosure is the unique equilibrium due to "unraveling." Since types with the best news will always reveal it, types with

[26]In an alternative formulation with the same qualitative implications a sender of type q can send message $v \in \{v_0, v_j, v_{j+1}, \ldots, v_N\}$ if and only if $q \geq s_j$

[27]Following convention, we define $s_0 = 0$ and $s_{N+1} = 1$ and ignore the open/closed set distinction in the notation for $[s_N, s_{N+1}]$.

[28]Since $E[\bar{q}_{[0,s_{j+1})}(x)|q]$ is strictly increasing in j, it follows from Theorem 1 of Milgrom and Roberts (1994) that $\hat{q}_j$ is strictly increasing in j.

the next best news will also reveal it, and so on until all news has been revealed. With only binary news, it was shown that unraveling in our model can fail because even the types with the best available news might not reveal it. We are interested in conditions under which the best types will always reveal their news and, when there are multiple levels of news, how far unraveling will continue.

Consider the set of beliefs where the receiver believes that for $q \in [s_{j+1}, 1]$ the sender discloses but for $q \in [s_1, s_{j+1})$ the sender may either disclose or not disclose. Define $\check{q}_j$ for $j = 1, 2, \ldots, N$ to be the maximal payoff for nondisclosure under any such beliefs.[29] Now consider unraveling. If $s_N > \check{q}_N$ then types with the best news v_N will disclose. Since $q \in [s_N, 1]$ are known to disclose, the attractiveness of nondisclosure by types with news v_{N-1} decreases. Thus types $q \in [s_{N-1}, s_N)$ will always disclose under the weaker condition that $s_{N-1} > \check{q}_{N-1}$. If these types disclose then this same logic applies to types with news v_{N-2}, etc. Because the $\check{q}_j$ are nondecreasing in j, unraveling implies that the standard for impressiveness becomes less strict as unraveling progresses from the best news down. For instance, if a PhD is sufficiently rare that it is disclosed, then it becomes more likely that an MA is disclosed, in which case it is also more likely that a BA is disclosed.

The following proposition uses these arguments to show when an equilibrium must involve a certain degree of disclosure. Unlike the classic unraveling results, this proposition does not imply that full unraveling or even any unraveling is assured. Instead, it gives conditions under which different levels of news are sufficiently impressive that they are always disclosed. In particular, a given level of news will be disclosed if it is sufficiently impressive conditional on higher levels of news being disclosed because they too are sufficiently impressive.

Proposition 5 (Unraveling due to high standards) *Types $q \in [s_j, 1]$ disclose in any equilibrium if standards s_k for $k \geq j$ are sufficiently high, $s_k \geq \check{q}_k$.*

This proposition shows that full unraveling is the unique equilibrium if the verifiable news is sufficiently impressive. The following result extends the unraveling argument to show that full unraveling is the unique equilibrium if the verifiable information is sufficiently fine. When the verifiable messages separate the different types sufficiently well, the highest types have an incentive to disclose their (exceptionally) good news v_N even if they are thought of as being only of type s_N rather than from the range $[s_N, 1]$. As seen in Figure 1d for $N = 2$, given that the highest types disclose v_N, the next highest types have an incentive to disclose v_{N-1} even under skeptical beliefs if s_{N-1} is sufficiently close to s_N, etc. If the difference between standards is sufficiently close for all the verifiable messages, i.e., the message space is sufficiently fine, then the unraveling continues until all news is disclosed.

[29] These critical values are derived and given by equation (A.22) in the proofs of Propositions 1 and 5 in Appendix A.

Proposition 6 (Unraveling due to fine standards) *Given $N \geq 2$ and s_1, types $q \in [s_j, 1]$ disclose in any equilibrium if standards are sufficiently fine, i.e., $\max_{k \geq j}\{s_{k+1} - s_k\}$ is sufficiently small.*

While Proposition 4 (i) shows that if standards are set low enough then nondisclosure is always an equilibrium, Propositions 5 and 6 show that if standards are set high enough or if standards are fine enough then nondisclosure cannot be an equilibrium. The following proposition shows how the distribution of sender types affects the potential for nondisclosure equilibria.

Proposition 7 (Unraveling due to low expectations) *(i) Types $q \in [s_j, 1]$ disclose in any equilibrium if prior expectations are sufficiently unfavorable, i.e., $F(s_j)$ is sufficiently large. (ii) If expectations are less favorable ($F(q)$ is MLR dominated by $G(q)$), the necessary and sufficient conditions for disclosure in any equilibrium are weaker ($\check{q}_j^F \leq \check{q}_j^G$ and $\hat{q}_j^F \leq \hat{q}_j^G$).*

Part (i) implies that full unraveling is the unique equilibrium if the prior distribution F is sufficiently unfavorable, i.e., $F(s_1)$ is sufficiently large, as seen in Example 1c and Proposition 1 for the case of $N = 1$. By similar arguments, nondisclosure is an equilibrium if F is sufficiently favorable, i.e., $F(s_N)$ is sufficiently small. Part (ii) shows that if the common knowledge distribution F is more favorable, then the conditions for the uniqueness of disclosure equilibria become stricter and the conditions for the existence of nondisclosure equilibria become less strict.

This result shows how different public information affects disclosure. For example, faculty members who are known to work at elite universities will tend to face priors (G) that have greater mass in the right tail relative to the priors (F) for faculty members at less research-focused universities. In general if there is any public signal that is affiliated with q, such as work place, then the conditional distribution given a higher realization of this signal implies MLR dominance of the conditional distribution given a lower realization. Hence the more favorable is public information about the sender the higher are $\check{q}_j$ and $\hat{q}_j$, so the less likely it is that $s_j > \check{q}_j$ and the more likely it is that $s_j \leq \hat{q}_j$. Looking back at Figure 1, the more favorable distribution in Figure 1a permitted nondisclosure and countersignaling equilibria, while the less favorable distribution in Figure 1c had a unique disclosure equilibrium.

7 Conclusion

A large body of research concludes that costless disclosure of good news should benefit the sender, but in practice senders often withhold good news. In this paper we consider a disclosure game when the receiver also has private information about sender quality. We show that the presence of any private receiver information, even if only slightly informative, implies that equilibria with nondisclosure by some or all types exist unless the standard for good news is restricted to sufficiently high quality senders or the prior distribution of types is sufficiently unfavorable. When there are multiple standards for news, such as letter grades, we find that the standard unraveling result holds if standards are sufficiently high or fine, but need not hold generally. These results provide support for mandatory or third-party disclosure of information as a way to reduce the damage that "false modesty" can have on communication.

A Proofs

Lemma 1 (Lower disclosure incentives for higher types) *Suppose* $E_H[q|q \in [a,b]] - E_L[q|q \in [a,b]]$ *is strictly decreasing in* a *and strictly increasing in* b *for all* $a < b$*. Then, for any* $a < b$*,* $\frac{d}{dq'}\left(\bar{q}_{[a,b)}(q') - \bar{q}_{[0,a)\cup[b,1]}(q')\right) < 0$ *for all* q' *if there exists* q *such that* $\bar{q}_{[a,b)}(q) = \bar{q}_{[0,a)\cup[b,1]}(q)$.

Proof: Letting $D = [a,b)$ and $N = [0,a) \cup [b,1]$, note that $\bar{q}_D(q) - \bar{q}_N(q)$ equals

$$\Pr[\hslash|q](E_H[q|q \in D] - E_H[q|q \in N]) + (1 - \Pr[\hslash|q])\left(E_L[q|q \in D] - E_L[q|q \in N]\right) \tag{A.1}$$

so $\frac{d}{dq}\Pr[\hslash|q] > 0$ implies the sign of $\frac{d}{dq}\left(\bar{q}_D(q) - \bar{q}_N(q)\right)$ equals that of

$$\left(E_H[q|q \in D] - E_H[q|q \in N]\right) - \left(E_L[q|q \in D] - E_L[q|q \in N]\right). \tag{A.2}$$

By assumption $\bar{q}_D(q) = \bar{q}_N(q)$ for some q, so from (A.1)

$$E_H[q|q \in D] - E_H[q|q \in N] \tag{A.3}$$

and

$$E_L[q|q \in D] - E_L[q|q \in N] \tag{A.4}$$

have opposite signs or are equal to zero.

Suppose (A.3) is non-negative and (A.4) is non-positive. By the law of total expectations, (A.3) non-negative implies $E_H[q|q \in D] \geq E_H[q] \geq E_H[q|q \in N]$, and (A.4) non-positive implies $E_H[q|q \in N] \geq E_L[q] \geq E_L[q|q \in D]$. Together, these imply

$$E_H[q|q \in D] - E_L[q|q \in D] \geq E_H[q] - E_L[q]. \tag{A.5}$$

But this contradicts the assumption that $E_H[q|q \in [a,b]] - E_L[q|q \in [a,b]]$ is strictly decreasing in a and strictly increasing in b. Hence it must be that (A.3) is negative and (A.4) is positive, implying (A.2) is negative. ■

Lemma 2 (Decreasing MRL Differences and Increasing MTTF Differences) *Suppose* H *and* L *are smooth distribution functions where the respective densities* h *and* l *are strictly* *positive on* $[0,1]$.[30] *Then, for all* $0 \le a < q < b \le 1$, *the gap* $E_H[q|q \in [a,b]] - E_L[q|q \in [a,b]]$ *is strictly decreasing in* a *and strictly increasing in* b *if (i)* H *is strictly convex and* L *is strictly concave or (ii)* H *and/or* L *is logconcave and the generalized hazard rate ratio*

$$\frac{h(q)}{H(q') - H(q)} / \frac{l(q)}{L(q') - L(q)} \tag{A.6}$$

is strictly increasing in $q \neq q'$ *for all* $q, q' \in [0,1]$.

Proof: (i) Lemma 1(iii) of Harbaugh and Rasmusen (2016) implies that $\frac{d}{da}E_H[q|q \in [a,b]] \le 1/2 \le \frac{d}{db}E_L[q|q \in [a,b]]$ and $\frac{d}{db}E_H[q|q \in [a,b]] \ge 1/2 \ge \frac{d}{da}E_L[q|q \in [a,b]]$. As can be seen from the proof of the lemma, the inequalities are strict for $a \neq b$. Hence

$$\begin{aligned} &\frac{d}{da}\left(E_H[q|q \in [a,b]] - E_L[q|q \in [a,b]\right) \le 1/2 - 1/2 = 0 \\ &\frac{d}{db}\left(E_H[q|q \in [a,b]] - E_L[q|q \in [a,b]\right) \ge 1/2 - 1/2 = 0, \end{aligned} \tag{A.7}$$

with strict inequalities for $a \neq b$.

(ii) Differentiating and simplifying,

$$\begin{aligned} \frac{d}{da}E_H[q|q \in [a,b]] &= \frac{h(a)}{H(b) - H(a)}\left(E_H[q|q \in [a,b]] - a\right) \\ \frac{d}{db}E_H[q|q \in [a,b]] &= \frac{h(b)}{H(b) - H(a)}\left(b - E_H[q|q \in [a,b]]\right) \end{aligned} \tag{A.8}$$

and similarly for $E_L[q|q \in [a,b]]$. By application of l'Hospital's rule, $\lim_{a \to b} dE_H[q|q \in [a,b]]/da = \lim_{b \to a} dE_H[q|q \in [a,b]]/db = \frac{1}{2}$ for $h(a), h(b) > 0$, and similarly for $E_L[q|q \in [a,b]]$.

Again differentiating and simplifying,

$$\begin{aligned} \frac{d^2}{(da)^2}E_H[q|q \in [a,b]] &= \frac{h'(a)}{h(a)}\frac{d}{da}E_H[q|q \in [a,b]] + \left(2\frac{d}{da}E_H[q|q \in [a,b]] - 1\right)\frac{h(a)}{H(b) - H(a)} \\ \frac{d^2}{(db)^2}E_H[q|q \in [a,b]] &= \frac{h'(b)}{h(b)}\frac{d}{db}E_H[q|q \in [a,b]] + \left(1 - 2\frac{d}{db}E_H[q|q \in [a,b]]\right)\frac{h(b)}{H(b) - H(a)} \end{aligned} \tag{A.9}$$

[30] The proof can be extended to allow h and l to be zero at the endpoints. In some applications MRL and MTTF ratios rather than differences may be important. The former is decreasing if $\frac{\frac{d}{da}E_H[q|q\in[a,b]]}{E_H[q|q\in[a,b]]} < \frac{\frac{d}{da}E_L[q|q\in[a,b]]}{E_L[q|q\in[a,b]]}$ and the latter is increasing if $\frac{\frac{d}{db}E_H[q|q\in[a,b]]}{E_H[q|q\in[a,b]]} > \frac{\frac{d}{db}E_L[q|q\in[a,b]]}{E_L[q|q\in[a,b]]}$. If $E_H[q|q \in [a,b]] > E_L[q|q \in [a,b]]$, as implied by our conditions and other conditions, a decreasing MRL difference implies a decreasing MRL ratio, and an increasing MTTF difference is implied by an increasing MTTF ratio.

and similarly for $E_L[q|q \in [a,b]]$. Note that

$$\begin{aligned} \frac{d^2}{(da)^2} E_H[q|q \in [a,b]]\Big|_{a=b} &= \frac{h'(b)}{h(b)}\frac{1}{2} > \frac{l'(b)}{l(b)}\frac{1}{2} = \frac{d^2}{(da)^2} E_L[q|q \in [a,b]]\Big|_{a=b} \\ \frac{d^2}{(db)^2} E_H[q|q \in [a,b]]\Big|_{b=a} &= \frac{h'(a)}{h(a)}\frac{1}{2} > \frac{l'(a)}{l(a)}\frac{1}{2} = \frac{d^2}{(db)^2} E_L[q|q \in [a,b]]\Big|_{b=a} \end{aligned} \tag{A.10}$$

where the inequalities hold by $H \succ_{MLR} L$, which is implied by (A.6).[31]

Let $\Delta_{[a,b]} = E_H[q|q \in [a,b]] - E_L[q|q \in [a,b]]$ and recall that $d\Delta_{[a,b]}/da = 1/2 - 1/2 = 0$ at $a = b$. From (A.10) we know that $d^2\Delta_{[a,b]}/(da)^2 > 0$ at $a = b$ so $d\Delta_{[a,b]}/da < 0$ for $a < b$ in a neighborhood of b. By way of contradiction, suppose that for some $a < b$, $d\Delta_{[a,b]}/da = 0$; at such an a, define $C = dE_H[q|q \in [a,b]]/da = dE_L[q|q \in [a,b]]/da$ where $C > 0$. In this case,

$$\begin{aligned} \frac{d^2\Delta_{[a,b]}}{(da)^2} &= \left(\frac{h'(a)}{h(a)} - \frac{h'(a)}{l(a)}\right) C + (2C-1)\left(\frac{h(a)}{H(b)-H(a)} - \frac{l(a)}{L(b)-L(a)}\right) \\ &> C\left(\frac{h'(a)}{h(a)} - \frac{l'(a)}{l(a)}\right) + C\left(\frac{h(a)}{H(b)-H(a)} - \frac{l(a)}{L(b)-L(a)}\right) \\ &\propto \left(\frac{h'(a)}{h(a)} + \frac{h(a)}{H(b)-H(a)}\right) - \left(\frac{l'(a)}{l(a)} + \frac{l(a)}{L(b)-L(a)}\right) \\ &= \frac{d}{da}\ln\left(\frac{h(a)}{H(b)-H(a)} \Big/ \frac{l(a)}{L(b)-L(a)}\right) > 0 \end{aligned} \tag{A.11}$$

where the first inequality uses the constraint implied by logconcavity (or decreasing hazard rate) that $C < 1$ and the implication from MLR dominance that $\frac{h(a)}{H(b)-H(a)} < \frac{l(a)}{L(b)-L(a)}$, and the last inequality follows from (A.6) for $q = a$ and $q' = b$ and from the log function being strictly increasing. But $d^2\Delta_{[a,b]}/da^2 > 0$ for any $a < b$ such that $d\Delta_{[a,b]}/da = 0$ cannot, by continuity, permit $d\Delta_{[a,b]}/da < 0$ for $a < b$ in a neighborhood of b as already established. So existence of an $a < b$ such that $d\Delta_{[a,b]}/da = 0$ is a contradiction and therefore $d\Delta_{[a,b]}/da < 0$ for all $a < b$.

Following similar steps, and using the implication of logconcavity that $\frac{d}{db}E[q|q \in [a,b] < 1$ (Bagnoli and Bergstrom, 2005) and the implication of MLR dominance that $\frac{h(b)}{H(b)-H(a)} > \frac{l(b)}{L(b)-L(a)}$, yields $d\Delta_{[a,b]}/db > 0$ for all $b > a$ if

$$\frac{d}{db}\ln\left(\frac{h(b)}{H(b)-H(a)} \Big/ \frac{l(b)}{L(b)-L(a)}\right) > 0, \tag{A.12}$$

which, by rearranging the denominators, also follows from (A.6) for $q = b$ and $q' = a$. ■

Lemma 3 (Weak private receiver information) *The generalized hazard rate ratio condition (A.6) holds for any logconcave F if $Pr[\hbar|q]$ is linear with a positive slope sufficiently close to zero.*

[31] This is seen by evaluating (A.6) at $a = b$, which requires application of l'Hospital's rule twice.

Proof: Let $\Pr[\hbar|q] = k + mq$. For presentational simplicity assume WLOG $k = 1/2$. From (A.11), the condition (A.6) for $q = a$ and $q' = b$ can be written as

$$\left(\frac{h'(a)}{h(a)} - \frac{l'(a)}{l(a)}\right) + \left(\frac{h(a)}{H(b) - H(a)} - \frac{l(a)}{L(b) - L(a)}\right) > 0 \tag{A.13}$$

or substituting and simplifying,

$$\frac{m}{1/4 - m^2a^2} + \frac{f(a)\int_a^b f(t)\,(m(a-t))dt}{\int_a^b f(t)\,(1/2 - mt))\,dt \int_a^b f(t)(1/2 + mt)dt} > 0\ , \tag{A.14}$$

which can be written as,

$$\frac{m}{1/4 - m^2a^2} + \frac{mf(a)\,((F(b) - F(a))\,a - (F(b) - F(a))\,E[q|q \in [a, b]]))}{(F(b) - F(a))^2\left(\frac{1}{2} - mE[q|q \in [a, b]]\right)\left(\frac{1}{2} + mE[q|q \in [a, b]]\right)} > 0, \tag{A.15}$$

or

$$\frac{m}{1/4 - m^2a^2} - \frac{mf(a)\,(E[q|q \in [a, b]] - a)}{(F(b) - F(a))\left(\frac{1}{4} - m^2E[q|q \in [a, b]]^2\right)} > 0. \tag{A.16}$$

Note that the LHS equals zero at $m = 0$. Taking its derivative with respect to m and evaluating at $m = 0$ yields,

$$4 - 4\frac{f(a)\,(E[q|q \in [a, b] - a)}{F(b) - F(a)}. \tag{A.17}$$

As seen from (A.8), this equals

$$4 - 4\frac{d}{da}E[q|q \in [a, b]], \tag{A.18}$$

which is greater than zero for logconcave F since $\frac{d}{da}E[q|q \in [a, b] < 1$, so condition (A.6) holds. The corresponding steps based on (A.8), and the implication of logconcave F that $\frac{d}{db}E[q|q \in [a, b] < 1$, establish that the condition holds for $q = b$ and $q' = a$. ■

Proof of Proposition 2: For the countersignaling and signaling equilibria, since both disclosure and nondisclosure can be observed in equilibrium, there are no unreached information sets so the Intuitive Criterion and D1 have no power to refine them away.

Now consider the nondisclosure equilibrium. Regarding the Intuitive Criterion, every type could benefit if the receiver believed them to be of type 1. Therefore, the Intuitive Criterion places no restrictions on out-of-equilibrium beliefs and nondisclosure remains an equilibrium. Regarding D1, the question is whether the skeptical beliefs $\mu(s|x, v_1) = 1$ are permissible. Under this refinement beliefs must put zero weight on any type which is willing to deviate for a strictly smaller range of actions by the receiver than another type when the actions must be a best response for some admissible beliefs. In our context where the receiver's only action is to estimate the sender's type, this means that beliefs must put zero weight on any type which is willing to deviate for a strictly smaller set of possible type estimates given the message. Since the estimate $E[\bar{q}_{[0,1]}(x)|q]$ for nondisclosure is strictly increasing in q and since $E[\bar{q}_{[0,1]}(x)|q=s] \geq s$ by the condition $s \leq \hat{q}$, the set of type estimates in $[s,1]$ that dominates this estimate is either empty or is the interval $[E[\bar{q}_{[0,1]}(x)|q],1]$. In the former case nondisclosure is an equilibrium for any beliefs. In the latter case, this set is largest for type $q=s$ since $E[\bar{q}_{[0,1]}(x)|q]$ is increasing in q, so D1 implies skeptical beliefs where $\mu(s|x,v_1)=1$. ■

Proof of Proposition 3: (i) For c sufficiently close to 1, $\bar{q}_{[s,c)}(c) > \bar{q}_{[0,s)\cup[c,1]}(c)$ since $\bar{q}_{[s,c)}(q) > \bar{q}_{[0,s)\cup[c,1]}(q)$ for all q as established in Proposition 1(iii). And for c sufficiently close to s, $\bar{q}_{[s,c)}(c) < \bar{q}_{[0,s)\cup[c,1]}(c)$ by continuity since $\bar{q}_{\{s\}}(s) < \bar{q}_{[0,1]}(s)$ by the existence of a [strict] nondisclosure equilibrium. Therefore there must be some $c > s$ such that $\bar{q}_{[s,c)}(c) = \bar{q}_{[0,s)\cup[c,1]}(c)$ and that $\frac{d}{dc}\bar{q}_{[s,c)}(c) > \frac{d}{dc}\bar{q}_{[0,s)\cup[c,1]}(c)$. Label the first such point c_1. Suppose initial behavior (rather than receiver beliefs) is more modest than this, $D=\emptyset$ or $D=[s,c)$ for $c<c_1$. In the former case beliefs are assumed to remain sufficiently pessimistic that the sender's best response is nondisclosure in the next period, and hence in each subsequent period. In the latter case beliefs in the next period are $D=[s,c)$ and $N=[0,s)\cup[c,1]$. Since $c<c_1$ then $\bar{q}_{[s,c)}(c) < \bar{q}_{[0,s)\cup[c,1]}(c)$. If $\bar{q}_{[s,c)}(q) < \bar{q}_{[0,s)\cup[c,1]}(q)$ for all q then convergence to nondisclosure is complete. If $\bar{q}_{[s,c)}(q) > \bar{q}_{[0,s)\cup[c,1]}(q)$ for some q then by continuity $\bar{q}_{[s,c)}(q) = \bar{q}_{[0,s)\cup[c,1]}(q)$ for some q, implying by Lemma 1 that $\frac{d}{dq}\bar{q}_{[s,c)}(q) < \frac{d}{dq}\bar{q}_{[0,s)\cup[c,1]}(q)$ for all q. Then, since $\bar{q}_{[s,c)}(c) < \bar{q}_{[0,s)\cup[c,1]}(c)$, the q such that $\bar{q}_{[s,c)}(q) = \bar{q}_{[0,s)\cup[c,1]}(q)$ is strictly lower than c. So the set of types who will disclose in the next period is more modest. This continues until no types disclose and convergence on nondisclosure is complete.

Jumping to (iii), for c sufficiently close to 1, $\bar{q}_{[s,c)}(c) > \bar{q}_{[0,s)\cup[c,1]}(c)$ as noted above. Consider any such c' such that $\bar{q}_{[s,c)}(c) > \bar{q}_{[0,s)\cup[c,1]}(c)$ for all $c > c'$. Suppose initial behavior is less modest than this, $D=[s,c)$ for $c \geq c'$. Beliefs in the next period are $D=[s,c)$ and $N=[0,s)\cup[c,1]$. Note that $\bar{q}_{[s,c)}(c) > \bar{q}_{[0,s)\cup[c,1]}(c)$ since $c \geq c'$. If $\bar{q}_{[s,c)}(q) > \bar{q}_{[0,s)\cup[c,1]}(q)$ for all q then in the next period $D=[s,1]$ and $N=[0,s)$. So convergence to disclosure is complete. If $\bar{q}_{[s,c)}(q) = \bar{q}_{[0,s)\cup[c,1]}(q)$ for some q then, by Lemma 1, $\frac{d}{dq}\bar{q}_{[s,c)}(q) < \frac{d}{dq}\bar{q}_{[0,s)\cup[c,1]}(q)$ for all q. Then, since $\bar{q}_{[s,c)}(c) > \bar{q}_{[0,s)\cup[c,1]}(c)$, the q such that $\bar{q}_{[s,c)}(q) = \bar{q}_{[0,s)\cup[c,1]}(q)$ is strictly higher than c. So the set of types who will disclose in the next period is less modest. This continues until all types disclose and convergence on disclosure is complete.

(ii) By the above arguments in (i) and (iii), when types $q > c$ countersignal the countersignaling region $[0,s)\cup[c,1]$ expands if $\bar{q}_{[s,c)}(c) < \bar{q}_{[0,s)\cup[c,1]}(c)$ and shrinks if $\bar{q}_{[s,c)}(c) > \bar{q}_{[0,s)\cup[c,1]}(c)$. Hence for countersignaling by types $q > c$ to be a stable equilibrium, $\bar{q}_{[s,c)}(c) - \bar{q}_{[0,s)\cup[c,1]}(c)$ must cross zero from above at c. In the limiting case where $\Pr[\hbar|q]$ is constant, the slope of $\bar{q}_{[s,c)}(c) - \bar{q}_{[0,s)\cup[c,1]}(c)$

reduces to $\frac{d}{dc}E[q|q \in [s,c]] - \frac{d}{dc}E[q|q \notin [s,c]]$. Note that by the law of iterated expectations,

$$\begin{aligned}\frac{d}{dc}E[q|q &\in [s,c]]\Pr[q \in [s,c]] + E[q|q \in [s,c]]\frac{d}{dc}\Pr[q \in [s,c]] + \\ \frac{d}{dc}E[q|q &\notin [s,c]](1 - \Pr[q \in [s,c]]) - E[q|q \notin [s,c]]\frac{d}{dc}\Pr[q \in [s,c]] = 0.\end{aligned} \tag{A.19}$$

In equilibrium for the limiting case, $E[q|q \in [s,c]] = E[q|q \notin [s,c]]$ so from (A.19),

$$\frac{d}{dc}E[q|q \in [s,c]] - \frac{d}{dc}E[q|q \notin [s,c]] = -\frac{\frac{d}{dc}E[q|q \notin [s,c]]}{\Pr[q \in [s,c]]} > 0, \tag{A.20}$$

where the inequality follows since an increase in c takes away mass that is above the mean of $E[q|q \notin [s,c]]$. Since this inequality is strict, it holds as long as $\Pr[\hbar|q]$ is sufficiently close to constant. Hence $\bar{q}_{[s,c)}(c) - \bar{q}_{[0,s)\cup[c,1]}(c)$ cannot cross zero from above at an equilibrium c if the private receiver information is too weak. ■

Proof of Proposition 4: (i) Consider the outcome in which news $v \le v_j$ is not disclosed while news $v > v_j$ is disclosed. First consider senders $q \in [s_k, s_{k+1})$ for $k \le j$. Assume that following an unexpected disclosure of v_k for $k \le j$, the receiver skeptically believes that the sender must be of type s_k. This yields the lowest possible out-of-equilibrium payoff of s_k. Since $\bar{q}_{[0,s_{j+1})}(q)$ is continuous in q, it follows that $q \le \bar{q}_{[0,s_{j+1})}(q)$ for any $q \le \hat{q}_j$. Therefore, since $s_k \le s_j \le \hat{q}_j$, $\bar{q}_{[0,s_{j+1})}(s_k) \ge s_k$ so that types $q < s_{j+1}$ have no incentive to disclose.

Now consider senders $q \in [s_k, s_{k+1})$ for $k > j$. Following the argument from (i), all $q \ge s_{j+1}$ are strictly better off disclosing.

(ii) Since $\bar{q}_{[s_j,s_{j+1})}(q) > \bar{q}_{[0,s_1)}(q)$ for any $j = 1, 2, \ldots, N$, if the receiver believes that all types disclose, every type is better off disclosing. ■

Proof of Proposition 5: For $0 < q' \le q'' \le 1$, define:

$$\tilde{q}(q', q'') = \sup_{Q}\{\bar{q}_Q(q'') : [0,q') \subset Q \subset [0,q'')\}. \tag{A.21}$$

This is the maximum nondisclosure payoff for a sender of quality $q = q''$ when the receiver believes that all $q \in [0,q')$ do not disclose and all $q \in [q'',1]$ disclose. Since $\bar{q}_Q(q)$ is increasing in q, this is an upper-bound on the maximum non-disclosure payoff for any sender $q \le q''$. Note that $\tilde{q}(q',q'')$ is continuous in q' since $F(q,x)$ is continuous in q, is nonincreasing in q' since higher q' implies a tighter restriction on Q, and is increasing in q'' since $\bar{q}_Q(q)$ is increasing in $q = q''$ and since higher

q'' implies a weaker restriction on $\boldsymbol{Q}$. Let

$$\check{q}_j = \begin{cases} q : q = \tilde{q}(q, s_{j+1}) & \text{if } j = 1 \\ \tilde{q}(s_1, s_{j+1}) & \text{if } j > 1 \end{cases} \tag{A.22}$$

where $\tilde{q}$ is given by (A.21). Since $\tilde{q}(\cdot, \cdot)$ is continuous and nonincreasing in its first argument, the fixed point $\check{q}_j$ exists and is unique. It is straightforward to see that $\check{q}_j < \check{q}_k$ whenever $j < k$. Note $\check{q}_j$ is the maximal possible nondisclosure payoff when the receiver believes that every $q \geq s_{j+1}$ disclose.

Starting with the highest types, if $s_N > \check{q}_N$ then types $q \in [s_N, 1]$ strictly prefer to disclose v_N by the definition of $\tilde{q}_N$. In this case if $s_{N-1} > \tilde{q}_{N-1}$ then types $q \in [s_{N-1}, s_N)$ strictly prefer to disclose v_{N-1} by the definition of $\tilde{q}_{N-1}$. The unraveling continues until types $q \in [s_j, s_{j+1})$ disclose v_j for $j > 1$. For the case where $j = 1$, we know that $s_1 > \check{q}_1 \geq \tilde{q}(s_1, s_2)$ since $\tilde{q}(q', q'')$ is nonincreasing in q' so that types $q \in [s_1, s_2)$ strictly prefer to disclose v_1. ■

Proof of Proposition 6: Let $\varepsilon = \min_{q \geq s_1}\{q - \tilde{q}(s_1, q)\}$. Since $[0, s_1)$ has positive mass, $\varepsilon > 0$. Starting with the highest types, suppose $1 - s_N < \varepsilon$. By the definition of ε and $\check{q}_N$, this implies $1 - s_N < 1 - \check{q}_N$, or $s_N > \check{q}_N$. By Proposition 5, news v_N is disclosed. Now suppose $s_N - s_{N-1} < \varepsilon$, which similarly implies $s_N - s_{N-1} < s_N - \check{q}_{N-1}$, or $s_{N-1} > \check{q}_{N-1}$. So by Proposition 5 news v_{N-1} is also disclosed. Continuing this process for the difference $s_{N-1} - s_{N-2}$, etc. down to the difference $s_{j+1} - s_j$, Proposition 5 implies news v_j is disclosed as long as $s_{k+1} - s_k < \varepsilon$ for $k \geq j$. ■

Proof of Proposition 7: (i) The question is whether, if the mass of F is sufficiently concentrated below a given s_j, it is assured that $\check{q}_j < s_j$. If $F(s_j)$ is sufficiently close to 1, then $\tilde{q}(s_1, s_{j+1}) < s_j$ since there is full support, since $s_1 > 0$, and since nearly all of the mass is below s_j. Thus $\check{q}_j < s_j$ for $j > 1$. Similarly for $j = 1$ the fixed point $q = \tilde{q}(q, s_2)$ must be less than s_1 so $\check{q}_1 < s_1$.

(ii) Let $\bar{q}_{Q,F}(q)$ and $\bar{q}_{Q,G}(q)$ be the expected estimates of q for distributions F and G respectively. Regarding $\check{q}_j$, MLR dominance implies that $\bar{q}_{Q,F}(q) \geq \bar{q}_{Q,G}(q)$ for any $\boldsymbol{Q}$. Therefore $\sup_Q\{\bar{q}_{Q,F}(q) : [0, q') \subset Q \subset [0, q'')\} \geq \sup_Q\{\bar{q}_{Q,G}(q) : [0, q') \subset Q \subset [0, q'')\}$, so $\tilde{q}_F(q', q'') \geq \tilde{q}_G(q', q'')$, which proves the result for $j > 1$. For $j = 1$, since $\tilde{q}(q, q'') - q$ is continuous in q and $\tilde{q}(q', q'') \in [0, 1]$ for all q, the conclusion follows directly from Theorem 1 of Milgrom and Roberts (1994). Similarly, regarding $\hat{q}_j$, $\bar{q}_{[0,s_{j+1}),F}(q) - q$ is continuous in q and $\bar{q}_{[0,s_{j+1}),F}(q) \geq \bar{q}_{[0,s_{j+1}),G}(q)$ and $\bar{q}_{[0,s_{j+1}),F}(q) \in [0, 1]$ for all q. So again the conclusion follows directly from Theorem 1 of Milgrom and Roberts (1994). ■

B Empirical example

Table 1 in the introduction indicates whether full-time, tenure-track faculty use the title "Dr," "PhD," or "Professor" and when they go by their names alone. This decision arises in many contexts including curricula vitæ, business cards, office doors, web sites, email signatures, etc. We look at office voicemail greetings and class syllabi since a sufficiently large sample is obtainable and the choice is likely to be under the control of the faculty.

To minimize the impact of different traditions in different disciplines we focus on economics departments, and to minimize regional variation we look at all state universities in California. In particular, based on faculty lists from department websites in the summer of 2004, we consider tenure-track faculty (assistant, associate, and full professors whom we refer to collectively as "faculty") with PhDs at 26 universities in the University of California and California State University systems with economics departments, excluding one department where the department chair was the only listed faculty member. Based on whether or not the economics department has a doctoral program, we divide the sample into eight "doctoral universities" and 18 "non-doctoral universities."

We start with a sample of 430 faculty with a primary position in one of the economics departments, 226 at doctoral universities and 204 at non-doctoral universities. For voicemail greetings we called at odd hours and on holidays when the faculty member was unlikely to be present. Excluding cases where voicemail was not working, was automated without a personal greeting, or was recorded by staff, we obtained valid voicemail greetings data for 128 of the faculty in doctoral universities and 121 in non-doctoral universities. For course syllabi we used the first listed undergraduate class on faculty web pages with a syllabus, and chose the format most likely to be handed out in class if multiple formats were listed, e.g., the .pdf format over the .html format. We obtained syllabi for 124 of the faculty at doctoral universities and 70 of the faculty at non-doctoral universities.

Based on the model we hypothesize that title usage is more likely when titles are less common. All of the economics faculty in our sample are tenure-track faculty with PhDs, but they are not immediately distinguishable to students and other observers from faculty without PhDs and from part-time instructors. As shown in Table 1, faculty at doctoral universities are significantly less likely to use titles in voicemail and syllabi. This patterns supports the hypothesis if titles are less common at non-doctoral universities.[32] Available data from annual *Common Data Set* reports for each university are consistent with this. For the 11 non-doctoral universities with available data, the average percent of full-time faculty with a PhD or the highest degree in their field was 80.1% in 2004. For part-time faculty the comparable number was 24.5%. The doctoral universities do not collect this data individually, but those that report a percentage use an estimate from the University of California system that 98% of faculty have PhDs or the highest degree in their field. For the 13 non-doctoral universities with available data, the percent of all faculty that were full-time faculty was 55.6% in 2004. For the seven doctoral universities with such data, the same figure was 80.0%.

[32] The model also predicts that title usage is more likely by younger faculty and by women since they are less represented among faculty. As shown in an earlier working paper version, these patterns hold.

References

Alós-Ferrer, C. and J. Prat (2012), “Job market signaling and employer learning,” *Journal of Economic Theory*, 147(5):1787–1817.

Araujo, A., D. Gottlieb and H. Moreira (2007), “A Model of Mixed Signals with Applications to Countersignaling and the GED,” *RAND Journal of Economics*, 38(4):1020–1043.

Athey, S. (2002), “Monotone Comparative Statics and Uncertainty,” *Quarterly Journal of Economics*, 117:187–223.

Bagnoli, M. and T. Bergstrom (2005), “Log-concave Probability and its Applications,” *Economic Theory*, 26(2):445–469.

Bagwell, K. and P. B. Overgaard (2005), “Look How Little I'm Advertising!” Working paper.

Baliga, S. and T. Sjostrom (2001), “Optimal design of peer review and self-assessment schemes,” *RAND Journal of Economics*, 32:27–51.

de Barreda, I. M. (2010), “Cheap Talk with Two-sided Private Information,” Working paper.

Bederson, B. B., G. Z. Jin, P. Leslie, A. J. Quinn and B. Zou (2018), “Incomplete Disclosure: Evidence of Signaling and Countersignaling,” *American Economic Journal - Microeconomics*, 10(1):41–66.

Bénabou, R. and J. Tirole (2006), “Incentives and Prosocial Behavior,” *American Economic Review*, 96:1652–1678.

— (2011), “Laws and Norms,” NBER Working Paper No. 17579.

Board, O. (2009), “Competition and Disclosure,” *Journal of Industrial Economics*, 57(1):197–213.

Bond, P. and Y. Zeng (2018), “Silence is safest: non-disclosure when the audience's preferences are uncertain,” Unpublished.

Bracha, A. and L. Vesterlund (2017), “Mixed signals: Charity reporting when donations signal generosity and income,” *Games and Economic Behavior*, 104:24–42.

Carbajal, J. C., J. Hall and H. Li (2015), “Inconspicuous conspicuous consumption,” Working paper.

Celik, L. (2014), “Information Unraveling Revisited: Disclosure of Horizontal Attributes,” *Journal of Industrial Economics*, 62(1):113–136.

Chen, Y. (2009), “Communication with Two-sided Asymmetric Information,” Working paper.

Cho, I.-K. and D. M. Kreps (1987), "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102:179–221.

Cho, I.-K. and J. Sobel (1990), "Strategic Stability and Uniqueness in Signaling Games," *Journal of Economic Theory*, 50:381–413.

Chung, K.-S. and P. Eső (2013), "Persuasion and Learning by Countersignaling," *Economics Letters*, 121(3):487–491.

Costrell, R. M. (1994), "A Simple Model of Educational Standards," *American Economic Review*, 84:956–971.

Daley, B. and B. Green (2014), "Market Signaling with Grades," *Journal of Economic Theory*, 151(1):114–145.

Dye, R. A. (1985), "Disclosure of Nonproprietary Information," *Journal of Accounting Research*, 23:123–145.

— (1986), "Proprietary and Nonproprietary Disclosures," *Journal of Business*, 59:351–366.

— (1998), "Investor Sophistication and Voluntary Disclosures," *Review of Accounting Studies*, 3:261–287.

Farrell, J. (1986), "Voluntary disclosure: robustness of the Unraveling Result, and Comments on its Importance," in R. E. Grieson, ed., "Antitrust and Regulation," pp. 91–103, Lexington Books, Lexington, MA, and Toronto.

Feltovich, N., R. Harbaugh and T. To (2002), "Too Cool for School? Signalling and Countersignalling," *RAND Journal of Economics*, 33(4):630–649.

Fishman, M. J. and K. M. Hagerty (2003), "Mandatory versus Voluntary Disclosure in Markets with Informed and Uninformed Customers," *Journal of Law Economics and Organization*, 19:45–63.

Giovannoni, F. and D. J. Seidmann (2007), "Secrecy, Two-sided Bias and the Value of Evidence," *Games and Economic Behavior*, 59(2):296–315.

Greenstone, M., P. Oyer and A. Vissing-Jorgensen (2006), "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments," *Quarterly Journal of Economics*, 121(2):399–460.

Grossman, S. J. (1981), "The Information Role of Warranties and Private Disclosure about Information Quality," *Journal of Law and Economics*, 24:461–483.

Grossman, S. J. and O. B. Hart (1980), "Disclosure Laws and Takeover Bids," *Journal of Law and Economics*, 35:323–334.

Grubb, M. D. (2011), "Developing a Reputation for Reticence," *Journal of Economics and Management Strategy*, 20(1):225–268.

Guo, Y. and E. Shmaya (2017), "The Interval Structure of Optimal Disclosure," Working Paper.

Harbaugh, R., J. W. Maxwell and B. Roussillon (2011), "Label Confusion: The Groucho Effect of Uncertain Standards," *Management Science*, 57(9):1512–1527.

Harbaugh, R. and E. Rasmusen (2016), "Coarse Grades: Informing the Public by Witholding Information," *American Economic Journal: Microeconomics*, forthcoming.

Hirshleifer, D., S. S. Lim and S. H. Teoh (2004), "Disclosure to an Audience with Limited Attention," Working paper.

Hoffman, M., C. Hilbe and M. A. Nowak (2018), "The signal-burying game can explain why we obscure positive traits and good deeds," *Nature Human Behaviour*, p. 1.

Hvide, H. (2003), "Education and the Allocation of Talent," *Journal of Labor Economics*, 21:945–976.

Jewitt, I. (2004), "Notes on the Shape of Distributions," Working paper, Nuffield College, Oxford.

Jin, G. Z. (2005), "Competition and Disclosure Incentives: An Empirical Study of HMOs," *RAND Journal of Economics*, 36(1):93–112.

Jin, G. Z. and P. Leslie (2003), "The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics*, 118:409–451.

Jovanovic, B. (1983), "Truthful Disclosure of Information," *Bell Journal of Economics*, 13:36–44.

Kamenica, E. and M. Gentzkow (2011), "Bayesian Persuasion," *American Economic Review*, 101(6):2590–2615.

Kim, J. and D. Shin (2016), "Price discrimination with demarketing," *The Journal of Industrial Economics*, 64(4):773–807, ISSN 1467-6451, doi:10.1111/joie.12129.

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk and M. Li (2017), "Persuasion of a Privately Informed Receiver," *Econometrica*, 85(6):1949–1964.

Levin, D., J. Peck and L. Ye (2009), "Quality Disclosure and Competition," *Journal of Industrial Economics*, 57(1):167–196.

Lizzeri, A. (1999), "Information Revelation and Certification Intermediaries," *RAND Journal of Economics*, 30(2):214–231.

Luca, M. and J. Smith (2015), "Strategic Disclosure: The Case of Business School Rankings," *Journal of Economic Behavior and Organization*, 112:17–25.

Lyon, T. P. and J. W. Maxwell (2004), "Greenwash: Corporate Environmental Disclosure under Threat of Audit," *Journal of Economics and Management Strategy*, 20(1):3–41.

Mathios, A. D. (2000), "The Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market," *Journal of Law and Economics*, 43(2):651–677.

Mathis, J. (2008), "Full Revelation of Information in Sender-Receiver Games of Persuasion," *Journal of Economic Theory*, 143:571–584.

Mayzlin, D. and J. Shin (2011), "Uninformative Advertising as an Invitation to Search," *Marketing Science*, 30(4):666–685.

Miklós-Thal, J. and J. Zhang (2013), "(De)marketing to Manage Consumer Quality Inferences," *Journal of Marketing Research*, 50(1):55–69.

Milgrom, P. (1981), "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12:380–391.

Milgrom, P. R. and J. Roberts (1986), "Relying on the Information of Interested Parties," *RAND Journal of Economics*, 17:18–32.

— (1994), "Comparing Equilibria," *American Economic Review*, 84:441–459.

Nanda, D. and Y. Zhang (2008), "Selective and Aggregate Disclosure," Mimeo.

Noe, T. H. (2017), "Selection and Stochastic Order," Working paper.

Okuno-Fujiwara, M., A. Postlewaite and K. Suzumura (1990), "Strategic Information Revelation," *Review of Economic Studies*, 57:25–47.

Orzach, R., P. B. Overgaard and Y. Tauman (2002), "Modest Advertising Signals Strength," *RAND Journal of Economics*, 33:340–358.

Orzach, R. and Y. Tauman (2005), "Strategic Dropouts," *Games and Economic Behavior*, 50:79–88.

Powers, N., A. Blackman, T. P. Lyon and U. Narain (2011), "Does Disclosure Reduce Pollution? Evidence from India's Green Rating Project," *Environmental and Resource Economics*.

Ramey, G. (1996), "D1 Signaling Equilibria with Multiple Signals and a Continuum of Types," *Journal of Economic Theory*, 69:508–531.

Sadowski, P. S. (2016), "Overeagerness," *Journal of Economic Behavior & Organization*, 131:114–125.

Seidmann, D. and E. Winter (1997), “Strategic Information with Verifiable Messages,” *Econometrica*, 65:163–169.

Shaked, M. and J. G. Shanthikumar (2007), *Stochastic Orders*, Springer Series in Statistics, Springer-Verlag, New York.

Shin, H. S. (1994), “News Management and the Value of Firms,” *RAND Journal of Economics*, 25:58–71.

— (2003), “Disclosures and Asset Returns,” *Econometrica*, 71:105–134.

Sobel, J. (2009), “Signaling Games,” in R. A. Meyers, ed., “Encyclopedia of Complexity and System Science,” Springer.

Spence, M. (2001), “Signaling in Retrospect and the Informational Structure of Markets,” *American Economic Review*, 92:434–459.

Teoh, S. H. and C. Y. Hwang (1991), “Nondisclosure and Adverse Disclosure as Signals of Firm Value,” *Review of Financial Studies*, 4(2):283–313.

Verrecchia, R. E. (1983), “Discretionary Disclosure,” *Journal of Accounting and Economics*, 5:179–194.

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

Watson, J. (1996), “Information Transmission When the Informed Party is Confused,” *Games and Economic Behavior*, 12(1):143–161.

Xiao, M. (2010), “Is Quality Certification Effective? Evidence from the Childcare Market,” *International Journal of Industrial Organization*, 28:708–721.