Persuasion under Second-Order Uncertainty

Patrick Hummel
Google

John Morgan
University of California, Berkeley

Phillip C. Stocken
Dartmouth College

April 22, 2016

Abstract

We study a general model of persuasion games when there is second-order uncertainty about the sender's knowledge of an uncertain state variable. Unlike situations where such uncertainty is absent, we show that second-order uncertainty eliminates truth-telling as an equilibrium. Instead, equilibrium consists of a convex interval of states where either disclosure or complete non-disclosure occurs, depending on the relative slopes of the ideal action lines of the sender and receiver. When choosing between senders differing in both expertise and preference alignment, we offer conditions where expertise dominates regardless of the degree of preference misalignment. Absent second-order uncertainty, we provide an algorithm for constructing a truthful equilibrium, as well as necessary and sufficient conditions for equilibrium uniqueness.

Keywords: Second-order uncertainty, persuasion, cheap talk, truth-telling, unraveling, full revelation

JEL Codes: C72, D80, D83

1 Introduction

In the spring of 2004, Alberto Gonzales, then White House Counsel to President George W. Bush, was on the cusp of achieving lifelong ambitions. Appointment to the Supreme Court or as Attorney General seemed imminent. When the latter position became vacant at the start of Bush's second term of office, Bush submitted Gonzales for confirmation in the Senate, a process that he breezed through to become the first Latino Attorney General of the U.S.

Gonzales brought his customary energy and lawyer's eye for detail to the post. One of his first actions was the unusual decision to dismiss nine U.S. Attorneys, allegedly for underperformance, but possibly for ideological reasons, as critics charged. It was a typically bold action from Gonzales, but one that would come to haunt him as complaints from the dismissed attorneys first led to press scrutiny and ultimately to a formal Senate investigation.

By the spring of 2007, Gonzales' sole preoccupation was preparing to testify under oath before many of the same senators who confirmed him, a process that would possibly make or break his career. Initially, Gonzales made a number of arguments defending his actions, but by the second day of testimony a clear pattern emerged in his responses. Virtually regardless of the question or questioner, Gonzales replied that he could not recall the answer. Sixty-four times over the course of a four-hour hearing he suffered such memory lapses. In the end, this would cost him his position as Attorney General and, ultimately, his career in the upper echelons of public service.

Senators from Gonzales' own Republican Party were perplexed at this performance. How could Gonzales, a man esteemed for his mental acuity and superb memory, have forgotten so much? One possibility is that Gonzales feared legal repercussions from any admissions he made; however, since this was a Senate hearing rather than a criminal proceeding, the main legal risk was perjury, for which feigning ignorance is no defense.[1] Thus, a strategy of pretending lack of knowledge may have carried more legal risk than one of disclosure. Even today, some ten years later, Gonzales' motives remain murky.[2]

Alberto Gonzalez' situation is, by no means, unique. When evaluating the statements of a sender/witness, the receiving party often faces uncertainty not just about the facts of the situation, i.e. the state variable, but also about the sender's knowledge of those facts. That is, receivers must account for a second layer of uncertainty in evaluating a sender's statements and forming posterior beliefs. Our main contribution is to study how this second layer of uncertainty affects information transmission in persuasion games.

Fundamental results, such as the phenomenon of unraveling, cease to operate when this additional uncertainty is present. To see this, consider a quintessential persuasion situation: A seller produces a good which may be of low or high quality and knows the actual quality. The seller would like to persuade the buyer that the good is high quality. The seller is limited, however, to making statements that cannot be shown to be false. It is well known that persuasion is impossible in these situations—in the unique equilibrium quality is truthfully revealed.

But now add a second layer of uncertainty—suppose that, in ascertaining quality, the seller conducts a test. Most of the time, the test reveals the true quality but

[1] For example, when baseball pitcher Roger Clemens testified in the Senate concerning his use of performance enhancing drugs, he was prosecuted for perjury and other obstruction charges but ultimately acquitted in 2012 (Wilber and Marimow, 2012).

[2] For further details about this case see Eggen and Fletcher (2007), Eggen and Kane (2007), Lewis (2008) and Milbank (2007).

occasionally it produces inconclusive results. Now the seller can, and will, avail himself of an opportunity to make ambiguous statements. When he knows the good is high quality, he says so, but when he knows it to be low quality, he simply remains silent, pooling these outcomes with inconclusive findings. Whereas in the former situation, silence would be viewed as equivalent to low quality, here silence could be taken at face value. Thus, the seller/sender is able to persuade the buyer/receiver that a low quality good is, in fact, of slightly higher (expected) quality.

Such an example is highly stylized, but contains within it the general properties of persuasion in the presence of second-order uncertainty: some information is always withheld in equilibrium, and the nature of disclosure is a combination of full revelation in some states and complete non-disclosure in others. While useful, it tells us little about exactly what information is disclosed under such uncertainty. We offer such a characterization for general preferences, and show how disclosure hinges fundamentally on the relative slopes of the lines describing the ideal actions of the sender and receiver. Our main findings, ordered by importance, are the following:

1. When the sender's knowledge state is private, i.e. receivers face second-order uncertainty, then full revelation is never an equilibrium. Instead, senders choose an interval or intervals of non-disclosure interlaced with full revelation.

2. The exact details of disclosure depend on the relative slopes of the "bliss" (ideal action) lines of the sender and receiver. When the slope of the sender's bliss line is much smaller than the receiver's, full revelation occurs over an interval. When the reverse is true, non-disclosure occurs over an interval.

3. Absent second-order uncertainty, the sender's loyalty is irrelevant to information disclosure. With second-order uncertainty, however, this is no longer true. In

choosing between senders strictly ranked by alignment or loyalty to the receiver, the irrelevance result only applies under a boundary equilibrium, an equilibrium where non-disclosure occurs at the endpoint of the state space; otherwise the receiver strictly prefers to consult a more loyal sender under an interior equilibrium.

4. By way of comparison, we also analyze situations where second-order uncertainty is absent. Our main findings here are: (a) an algorithm for constructing a fully revealing equilibrium; (b) necessary and sufficient conditions for full revelation to be the unique equilibrium; and (c) when multiple equilibria exist, the sender and receiver always disagree as to the better equilibrium.

A key insight of the model is that information disclosure does not depend directly on the amount of disagreement between the two parties, but on the relative slopes of their bliss lines. Indeed, points where both parties agree as to the ideal action are not necessarily associated with full disclosure in their neighborhood.

Since the paper's main novelty is to study persuasion under second-order uncertainty, we describe the related literature mainly in that context, and offer only a cursory review of the vast literature where it is absent.

The nearest antecedents to our work are Dye (1985), Shin (1994a), and their successors.[3] They study situations where second-order uncertainty is present but with preferences akin to the example; that is, where the receiver seeks to match her action to the expected state and where all sender types wish to convince the receiver that the state is at its highest possible level. We derive general conditions on preferences where equilibrium takes the form found in these works—suppressing

[3] See, e.g., Jung and Kwon (1988), Penno (1997), Dye (1998), Pae (2005), Shin (2006), Guttman *et al.* (2014), and Hummel *et al.* (2016).

bad news and revealing good news. We also show how, by changing preferences, equilibrium no longer takes this form—instead extremes on both ends of the state space are suppressed and only moderate states are disclosed fully. More broadly, we add to this literature by generalizing preferences, allowing for the possibility that the sender and receiver agree on the optimal action in some states, and characterizing the set of equilibria.

Shin (1994b), as well as Bhattacharya and Mukherjee (2013), study situations of second-order uncertainty in the presence of multiple experts.[4] In all cases, these experts have "flat" bliss lines, having a most preferred action irrespective of the state. Since we study single sender situations, there is no direct parallel between their results and ours. However, like our study of how the receiver chooses among senders, this literature also indicates that receivers are not necessarily well served by always selecting, or listening to, more aligned senders.

While our main contribution is the study of persuasion games in the presence of second-order uncertainty, we also consider situations where such uncertainty is absent, treating this as a benchmark. Milgrom (1981) and Grossman (1981) spawned a vast literature studying persuasion under this assumption. The main finding is that, under quite general circumstances, full revelation is an equilibrium (Seidmann and Winter, 1997). Seidmann and Winter (1997) also offer sufficient conditions for uniqueness. Giovannoni and Seidmann (2007) further note that, when multiple equilibria are present, the sender might prefer a less informative equilibrium. We generalize both results, identifying necessary and sufficient conditions for uniqueness and showing that, when there are multiple equilibria, the sender and receiver always disagree over their ranking. While initial persuasion models assumed that the sender's mes-

[4] See also Bhattacharya *et al.*(2015) for further extensions of these situations.

sage space was binary, consisting of truthful disclosure or no disclosure whatsoever, considerable work examines the effects of weakening these message space restrictions.[5] Here too the main finding is that full revelation is a robust phenomenon.

The paper proceeds as follows: Section 2 sketches the model. In Section 3, we present findings when second-order uncertainty is absent, leading to finding 4. These findings are included primarily as a benchmark and do not reflect our main contribution. Section 4 characterizes equilibrium with second-order uncertainty and contains findings 1 and 2 above. Section 5 compares information transmission based on the sender's loyalty and shows that finding 3 holds. Finally, Section 6 concludes.

2 The Model

There are two players, a sender S and a receiver R. The sender is possibly knowledgeable about a payoff-relevant state variable, θ. He sends a message m to the receiver who then takes an action $y \in \Re$ that affects the payoffs of both parties. The state variable θ is commonly known to be drawn from an atomless distribution $F(\theta)$ having support $\left[\underline{\theta}, \bar{\theta}\right]$ including, possibly, the entire real line.

Senders can differ in their knowledge state $t \in \{0,1\}$. With probability $p \in (0,1)$, a sender is *knowledgeable*, i.e. $t = 1$, and knows the realized state perfectly.[6] Otherwise,

[5] Okuno-Fujiwara *et al.* (1990) first extended the unraveling result by relaxing the message state to allow partially informative messages. Koessler (2003) further relaxes message space restrictions while showing that full revelation still obtains. Mathis (2008) showed how unraveling extends to situations of partial certifiability of messages. Most recently, Hagenbach *et al.* (2014) offered conditions for full revelation in n-player persuasion games.

[6] The model straightforwardly extends to a situation where the sender is either imperfectly informed about the state or completely uninformed. In this case, reinterpret the state variable as the sender's signal and redefine bliss actions accordingly to use the existing framework.

the sender is *not knowledgeable*, i.e. $t = 0$, and knows nothing more about the state than the receiver.[7] Formally, a sender receives a signal s whose distribution depends on both the knowledge state, t, and the state of nature, θ. Conditional on θ, a knowledgeable sender receives the signal $s = \theta$. A sender who is not knowledgeable receives a signal drawn from an arbitrary distribution, G_0, having the same support as F but independent of its realized value. Since such a signal is uncorrelated with the state, it is uninformative.

We will vary the receiver's information as to the *knowledge state* of the sender. Much of the extant literature implicitly assumes that the knowledge state is public, i.e. it is common knowledge whether the sender is informed. Our main concern, however, is with the case where the knowledge state t is *private information*, i.e., only the sender knows whether he is informed. In this case, the receiver labors under second-order uncertainty—she knows neither the realized state nor the sender's knowledge of the state.

A public knowledge state may arise when the sender must undertake certain observable procedures to learn the state, but the results require the sender's expertise to interpret. On the other hand, a private knowledge state can come about when no explicit procedures are needed to acquire expertise or when such procedures might produce inconclusive results. In the latter situation, even though the receiver knows whether procedures were conducted, her ignorance of the results still leaves her in doubt as to the sender's knowledge state.

A sender, after learning his knowledge state t and receiving signal s, transmits a report m to the receiver consisting of a statement as to the signal received and, where

[7] Throughout, the terms knowledgeable and informed and the terms not knowledgeable and uninformed are used as synonyms.

relevant, the knowledge state. Let σ denote the sender's message as to the signal. We treat this message as hard information in that it cannot be shown to have been false once the state is revealed. Formally, a feasible message σ is a (possibly degenerate) closed subset $I \subseteq [\underline{\theta}, \bar{\theta}]$ that contains the true state θ.[8] Thus, an uninformed sender must send the message $\sigma = [\underline{\theta}, \bar{\theta}]$, which we denote $\sigma = \emptyset$, to be assured of sending a message that contains the true state. Let $\tau \in \{0, 1\}$ denote the reported knowledge state, where $\tau = 0$ indicates that the sender is uninformed and $\tau = 1$ indicates that he is informed. Unlike σ, the message τ is soft information and need not be truthful, capturing the idea that knowledgeability is more subjective than the facts of the case.[9]

After receiving the message $m = (\sigma, \tau)$, the receiver selects an action $y \in \Re$ based on her preferences. Let $U_i(y, \theta)$ denote the payoffs of player $i \in \{R, S\}$ when action y is chosen in state θ. We assume that, for every θ, payoffs are continuous and single-peaked in y, with a unique payoff-maximizing action, $y_i(\theta)$, which we term i's *bliss action.* We assume $y_S(\theta)$ is continuous and weakly increasing in θ, whereas $y_R(\theta)$ is continuous and strictly increasing in θ. When the receiver has no information, $y(\emptyset)$ denotes her bliss action.

Bliss actions alone are not sufficient to define preferences. We assume that preferences satisfy the *distance property,* which requires that payoffs are proportional to the distance between the chosen action and the bliss action. Formally,

Definition 1 *The preferences of agent $i \in \{S, R\}$ satisfy the* ***distance property*** *if and only if $U_i(y, \theta) = u_i(|y - y_i(\theta)|; \theta)$ for some strictly decreasing function u_i.*

[8] Examples of feasible messages include $\{\theta\}$, $[\theta_1, \theta_2]$, $[\theta_1, \theta_2] \cup [\theta_3, \theta_4]$ and $[\underline{\theta}, \bar{\theta}]$ for any $\theta_1, ..., \theta_4$ satisfying $\underline{\theta} \leq \theta_1 < \theta < \theta_2 < \theta_3 < \theta_4 \leq \bar{\theta}$.

[9] The inclusion of the message τ is merely for completeness of the possible messages the sender might convey. Since τ is soft information, no credible signaling will be possible in equilibrium; hence τ effectively does not figure into the analysis.

The distance property implies that, when the action chosen exceeds the agent's bliss action by a given amount, the agent evaluates it in the same fashion as when it falls below the bliss action by the same amount. It does not imply that losses from errors of a given magnitude are the same across states; for instance, in state θ, the sender may suffer a quadratic loss as a function of the distance of the error whereas in state θ' he may suffer a quartic loss.

Disagreement as to the bliss action between the sender and receiver constitutes the key barrier to information transmission. We model this by assuming that $y_S(\theta) \neq y_R(\theta)$ except at finitely many *agreement points*. This disagreement creates an incentive for the sender to try to persuade the receiver to follow his preferred action. The receiver wants to avoid being misled.

We use the following solution concept to characterize the results: The receiver uses Bayes' rule wherever possible in formulating beliefs. We further restrict beliefs such that, if the sender sends the (possibly degenerate) message I then the receiver must believe that the state lies somewhere in I, even if I lies off the equilibrium path. Given her beliefs, the receiver chooses an action maximizing expected payoffs. The sender chooses messages optimally given the receiver's anticipated response.[10]

3 Sender Knowledge Is Public Information

We first study situations where the knowledge state is public. This serves mainly as a benchmark for situations where the knowledge state is private. We offer three main results. First, while it is known that full revelation is generally an equilibrium in these situations, we offer a novel construction of such an equilibrium. Second, full

[10] Bhattacharya and Mukherjee (2013) use a similar solution concept.

revelation is not always the unique equilibrium. Finally, we show that when partially revealing equilibria exist, the sender always prefers these equilibria to full revelation. Thus, unlike the situation of pure cheap talk, where there is a Pareto ranking of equilibria by informativeness, the sender and receiver fundamentally disagree about the equilibrium ranking under persuasion.

When the sender is known not to be knowledgeable, the game is trivial to analyze. Since the sender knows nothing, the receiver takes the optimal action given her prior beliefs. The remainder of the analysis concerns the case where the sender is known to be knowledgeable.

Throughout the analysis, situations where the bliss actions cross, which we term *agreement points*, play an important role.[11] Formally,

Definition 2 *State $\theta' \in \left(\underline{\theta}, \bar{\theta}\right)$ is an* ***agreement point*** *if and only if $y_R(\theta') = y_S(\theta')$ and there is a neighborhood of θ', N, such that $sign\left(y_S(\theta_a) - y_R(\theta_a)\right) \neq sign\left(y_S(\theta_b) - y_R(\theta_b)\right)$ for all $\theta_a, \theta_b \in N$ satisfying $\theta_a < \theta' < \theta_b$.*

It is known that the unraveling argument, first introduced by Grossman (1981) and Milgrom (1981), generalizes substantially to produce a fully revealing equilibrium. The idea underlying the argument is that for any pooling interval some positive measure of sender types are disadvantaged by being pooled. Since these types then have an incentive to deviate and fully reveal, this destroys the possibility of pooling in equilibrium. Such intuition, while powerful, depends on certain key features to operate. The following result shows that bounded state spaces are one such key feature.

[11] An agreement point represents a state where the sender and receiver share an ideal action but where this action does not represent a tangency point between their bliss lines.

Proposition 1 *Full revelation is never an equilibrium when:*

(a) The state space is unbounded from above and $y_R(\theta) > y_S(\theta)$ for θ sufficiently large; or

(b) The state space is unbounded from below and $y_R(\theta) < y_S(\theta)$ for θ sufficiently small.

The unraveling intuition relies essentially on the presence of a worst sender type following any deviation message, a type who could feasibly have sent the message but with whom no other feasible sender type would wish to be confused. For instance, when the sender's bliss line lies strictly above the receiver's, no other feasible sender type would wish to be thought of as the lowest type that could have sent a given deviation message. A receiver who believes that the worst type deviated will choose an action that is lower than full revelation for anyone sending this message; hence, such a deviation is unprofitable for the sender.

When the state space is unbounded, such a worst type no longer exists following certain messages. In the situation described above, for example, the null message always represents a profitable deviation. This message will induce some action y' that is ideal for the receiver in some state θ'. Sender types just below θ', which always exist owing to the unbounded state space, find it profitable to deviate, as the null message induces a slightly higher action than does full revelation. More generally, although it is well known that full revelation is the unique equilibrium when the state space is bounded and bliss lines never intersect, Proposition 1 implies that, for this same preference configuration (and many others), full revelation is never an equilibrium when the state space is unbounded. The key insight is that, unless unraveling eventually produces a distinct worst sender type, persuasion game restrictions on messages do not, in and of themselves, ensure full disclosure.

Proposition 1 establishes the importance of bounded state spaces for the existence of a fully revealing equilibrium. As in the model of Seidmann and Winter (1997), we assume the necessary boundedness conditions hold throughout the remainder of the paper. Since, absent second-order uncertainty, the two models differ little, the commonality in the conditions for existence is not altogether surprising.

What does differ, however, is the method of proof. Most previous existence proofs rely on a combination of contradiction and induction to show that there exists some set of off-equilibrium beliefs and actions such that deviating from full revelation is never optimal.[12] Instead, we offer a constructive proof of the existence of a truth-telling equilibrium, delineating the out of equilibrium beliefs and actions required to support full disclosure.

We now describe the construction. Following any singleton message, $\sigma = \theta$, the receiver believes the state equals θ with probability one and hence plays the action $y_R(\theta)$. By contrast, for any message $\sigma = I$, where I is a non-degenerate subset, the receiver forms beliefs and chooses actions according to the Bifurcation Algorithm.

Bifurcation Algorithm

The point of the algorithm is to construct an action following any message $\sigma = I$ such that all sender types $\theta \in I$ would prefer truth-telling than $y(I)$. The algorithm first identifies the sender type who disagrees most vigorously with the receiver (i.e. the type where the gap between bliss points is largest) in $C(I)$, the convex hull of I. It then identifies a set of receiver bliss actions, associated with states in $C(I)$, that are worse than truth-telling for this sender type. Next, we study the subinterval of

[12] The exception is Hagenbach, *et al.* (2014) who delineate out of equilibrium beliefs supporting full revelation for a broad class of games including persuasion. Our algorithm, however, differs from theirs.

$C(I)$ consisting of the threatened actions described above and again find the sender type with the greatest disagreement. Once more, we construct threats to induce truth-telling for this type. This process continues until ultimately converging on a feasible threatened action that induces all sender types in $C(I)$ to tell the truth.

Formally, suppose the convex hull of the sender's message, $C(I) = [\theta_0, \theta_{n+1}]$, contains exactly $n \geq 0$ agreement points $\{\theta_1, \theta_2, ..., \theta_n\}$ where $\theta_i < \theta_{i+1}$. Partition the state space into intervals $I_k = [\theta_{k-1}, \theta_k]$ for $k = 1, 2, ..., n+1$. Define the **gap** of interval I_k to be the largest value $|y_R(\theta) - y_S(\theta)|$ for $\theta \in I_k$. Let $\theta_k^* \in \arg\max_{\theta \in I_k} |y_R(\theta) - y_S(\theta)|$.

The algorithm successively bifurcates this set of intervals into ever smaller subsets up to the point where no further bifurcation is possible. Specifically:

Step 1: Let $\bar{I}_k$ be an interval containing the largest gap, i.e. the largest value of $|y_R(\theta) - y_S(\theta)|$.

- Branch U: If $y_S(\theta) < y_R(\theta)$ for $\theta \in \bar{I}_k$, we will only consider intervals $J = \{I_{k+1}, I_{k+2}, ..., I_n, I_{n+1}\}$.

- Branch D: If $y_S(\theta) > y_R(\theta)$ for $\theta \in \bar{I}_k$, we will only consider intervals $J = \{I_1, I_2, ..., I_{k-2}, I_{k-1}\}$.

For J appropriately defined (depending on the branch), let $\bar{I}_{k'}$ be the interval containing the largest value of $|y_R(\theta) - y_S(\theta)|$ conditional on $\theta \in J$.

Step 2: Consider the intervals J for Branch U. We bifurcate this interval as follows:

- Branch UU: If $y_S(\theta) < y_R(\theta)$ for $\theta \in \bar{I}_{k'}$, we will only consider intervals $J = \{I_{k'+1}, I_{k'+2}, ..., I_n, I_{n+1}\}$.

- Branch UD: If $y_S(\theta) > y_R(\theta)$ for $\theta \in \bar{I}_{k'}$, we will only consider intervals $J = \{I_{k+1}, I_{k+2}, ..., I_{k'-2}, I_{k'-1}\}$.

For Branch D, let branches DU and DD be analogously defined.

Steps $3, 4, ..., m$: Repeat this process performing an analogous bifurcation procedure. Continue until a bifurcation leads to the empty set.

Definition 3 *The sequence of* ***bifurcation intervals*** *is* $\bar{I} = \left\{\bar{I}_k, \bar{I}_{k'}, \bar{I}_{k''}, ...\bar{I}_{k^{(m)'}}\right\}.$

We now specify the receiver's choice of action following message σ. There are three possibilities:

Case (I): In every instance where a bifurcation occurred, $y_S(\theta) < y_R(\theta)$, i.e., the terminal branch was UUU...U.

Case (II): In every instance where a bifurcation occurred, $y_S(\theta) > y_R(\theta)$, i.e., the terminal branch was DDD...D.

Case (III): There exists at least one instance of a bifurcation where $y_S(\theta) > y_R(\theta)$ and at least one instance where $y_S(\theta) < y_R(\theta)$, i.e., the terminal branch contains at least one U and one D.

Given these possibilities, the receiver's beliefs and actions are as follows:

In Case (I), let the receiver hold beliefs that place probability one on state θ_{n+1} and choose $y(\sigma) = y_R(\theta_{n+1})$.

In Case (II), let the receiver hold beliefs that place probability one on state θ_0 and choose $y(\sigma) = y_R(\theta_0)$.

In Case (III), let the receiver choose an action $y(\sigma) = y_R(\theta_K)$ where $\theta_K = \theta_{k^{(m)'}}$ if $y_S(\theta) < y_R(\theta)$ for $\theta \in \bar{I}_{k^{(m)'}}$ and $\theta_K = \theta_{k^{(m)'}-1}$ if $y_S(\theta) > y_R(\theta)$ for $\theta \in \bar{I}_{k^{(m)'}}$. The receiver holds beliefs that place probability π on state θ_0 and $(1-\pi)$ on state θ_{n+1} where π is chosen so that $y_R(\theta_K)$ is optimal.

The following proposition shows that the sender can never profitably deviate by sending messages that are non-degenerate intervals.

Proposition 2 *If the knowledge state is public and sender preferences are bounded, then full revelation is an equilibrium. Specifically, any message $\sigma = I$, where I is a non-degenerate subset, is not a profitable deviation from full revelation when $y(I)$ is determined by the Bifurcation Algorithm.*

Uniqueness and Equilibrium Selection

Although having an algorithm for constructing a fully revealing equilibrium is useful, Proposition 2 plods familiar ground—it is known that under general persuasion games, full revelation is an equilibrium. Less well understood are conditions for uniqueness and equilibrium selection in the face of multiplicity. The latter is particularly important given the emphasis on fully revealing equilibria (e.g. Hagenbach, *et al.,* 2014). We show, however, that fully revealing equilibria need not be unique and, when multiple equilibria are present, the sender and receiver disagree as to which should be played.

In this section, we offer two results. First, we identify *necessary and sufficient* conditions for the full revelation to be the unique equilibrium. Second, we show that, when multiple equilibria exist, senders and receivers fundamentally disagree as to which should be played. In particular, senders always strictly prefer a *less* informative to a fully revealing equilibrium.

By way of background, Seidmann and Winter (1997), among others, offer sufficient conditions for full revelation to be the unique equilibrium. The degree to which these conditions might be weakened remains an open question; however, we can repurpose the Bifurcation Algorithm to strengthen results on uniqueness. The required condition, it turns out, permits an if and only if answer to the uniqueness question.

The basic idea of the proof is to operate the Bifurcation Algorithm in reverse—if the receiver chooses a non-extreme action following a pooling message $\sigma = I$, under what preferences will the sender prefer this non-extreme action to full revelation? The key condition is "conservatism"—a sender is conservative (relative to the receiver) in the neighborhood of an agreement point if his bliss line is much less responsive to changes in the state, in a sense to be made precise below. For instance, suppose that the sender's bliss line is relatively flat and cuts the receiver's bliss line from above at an agreement point. Now, the gap between the sender's bliss line and the agreement action is smaller than the gap between bliss lines away from the agreement point. Hence, so long as pooling near the agreement point induces an action close to the agreement action, all senders will prefer this to full revelation. To ensure that full revelation is the unique equilibrium, such circumstances must be ruled out, i.e. the sender must not be conservative. Formally,

Condition 1 *A sender is* ***conservative*** *in state θ if there exists an agreement point, θ', such that $U_S(y_R(\theta'), \theta) > U_S(y_R(\theta), \theta)$. If, for every agreement point θ' and (almost) all states θ, $U_S(y_R(\theta'), \theta) < U_S(y_R(\theta), \theta)$, a sender is* ***not conservative.***

Note that conservatism only occurs in reference to agreement points. When the sender and receiver always disagree as to the ideal action, full revelation is indeed the *unique* equilibrium. Ironically, it is the presence of points where conflict is absent—points where full revelation can hardly be in doubt, that create the possibility of equilibrium information withholding. Specifically,

Proposition 3 *If the sender is not conservative, then full revelation is the unique equilibrium. However, if the sender is conservative for some positive measure of states θ, then a partial pooling equilibrium exists.*

We now turn to equilibrium selection when the sender is conservative. It might seem obvious that, since full revelation is an equilibrium, both parties ought to prefer its selection. Indeed, such an informativeness selection criterion is common in the cheap talk literature. Crawford and Sobel (1982) first showed that, under pure cheap talk, such a selection could be justified by an *ex ante* Pareto ranking—both sender and receiver *ex ante* prefer the more informative equilibrium in their setting.

Persuasion games, however, do not have this structure. Instead, senders strictly prefer a less informative equilibrium when one is present whereas receivers prefer full revelation. To see why, note that, in such a partially revealing equilibrium, the sender could, if desired, induce the receiver to play the actions associated with full revelation merely by revealing truthfully. But, since partial revelation is an equilibrium, the sender's payoff from doing so must be lower than that obtained by withholding information. Therefore, senders always prefer less informative equilibria to full revelation.

The same holds more generally when comparing equilibria with differing levels of informativeness. Specifically, if we compare an equilibrium in which some set of states is revealed with one in which those states are not revealed, the same argument implies that the sender will prefer the latter equilibrium to the former.[13]

Before proceeding to the formal result, it is useful to define a strict informativeness ordering over equilibria which consist of either full revelation or no revelation in each state. We say that an equilibrium is *strictly more informative* than another if, in every situation of the latter where there is full revelation, the sender fully reveals in the former. Furthermore, there exist some positive measure of states in the latter

[13]The limiting case of this argument implies that full pooling is the sender's most preferred disclosure regime. However full pooling is generically not an equilibrium in this setting.

where the sender offers the null message and fully reveals in the former. With this ordering in mind, we have shown:

Proposition 4 *If there are multiple equilibria strictly ordered by informativeness, the sender strictly prefers the least informative equilibrium whereas the receiver prefers the most informative equilibrium.*

An important corollary of the proposition concerns fully revealing equilibria:

Corollary 1 *The sender strictly prefers any equilibrium with partial pooling to one producing full revelation. The receiver prefers the opposite.*

Proposition 4 does not imply that focusing on the receiver's preferred equilibrium is unreasonable, merely that it cannot be justified on Pareto grounds. The appropriate equilibrium selection in these circumstances remains an open question.

Before proceeding to our main results, context is important. The situation where the sender's knowledge state is public has been much more widely studied than settings where the knowledge state is private. The main findings here show the robustness of full revelation. We modestly add to this collection of results by strengthening conditions for uniqueness and noting difficulties in equilibrium selection favoring full revelation. Nonetheless, it should be emphasized that these are mere benchmarks for our main concerns—situations where the receiver has doubts about whether the sender is informed, which we turn to next.

4 Sender Knowledge Is Private Information

We now turn to the heart of the paper, the study of situations where the knowledge state is private. As mentioned above, such situations arise whenever there is room for doubt as to the “expertise” of the sender. This might be due to the possibility of forgetfulness on the sender’s part or to situations where the signals about the state are possibly uninformative, as when a test or experiment designed to measure the state delivers an ambiguous or inconclusive outcome. Expert witness testimony represents a canonical case of such situations.

Equilibrium disclosure in this setting is the polar opposite of that when the knowledge state is public. There it is well known that full revelation is an equilibrium. When the knowledge state is private, however, this is never the case or, to be more precise, for *generic* parameters of the model, all equilibria entail some degree of information loss. By generic, we mean all cases save for the knife-edge situation where there is an agreement point whose action corresponds exactly to the optimal action the receiver would undertake given her prior beliefs.

With a private knowledge state, non-disclosure is always on the equilibrium path since uninformed senders have no recourse but to send the null message. This, in turn, limits the ability of the receiver to strategically respond to such a message so as to create incentives for the informed types to disclose. To see why this wrecks full revelation, notice that the null message produces an action that differs from the fully revealing action. Since the sender and receiver disagree as to the ideal action, it then follows that, by deviating to the null message, certain sender types can profitably shift the action in a favorable direction. Formally,

Theorem 1 *If the knowledge state is private, then full revelation is generically not an equilibrium.*

Dye (1985) offered a version of this result for the special case where the receiver chooses an action equal to the expected state, and the sender always prefers the highest possible action.[14] Analogous to Seidmann and Winter (1997), who extended the full revelation findings of Milgrom (1981) and Grossman (1981) to arbitrary preferences when the knowledge state is public, Theorem 1 extends the non-existence result to a broad class of preferences when the knowledge state is private. Since one can no longer focus on full revelation, it remains to determine the nature of equilibrium and the degree of information loss in these settings. We do this next.

4.1 Convex Disclosure Equilibrium

Initially, we restrict attention to a type of equilibrium we label a convex disclosure equilibrium. This is an equilibrium where the disclosure region is convex and where complete non-disclosure occurs elsewhere. We will later show that, for a certain class of preferences, the restriction to this type of equilibrium is without loss of generality—all equilibria are of this form. The precise condition on preferences where this is the case is when preferences satisfy what we term the *gradual slope ordering property,* which holds when the slope of the sender's bliss line is less than half that of the receiver's bliss line. Formally,

Definition 4 *The bliss lines $y_S(\theta)$ and $y_R(\theta)$ satisfy the* gradual slope ordering property *if for all $\theta' > \theta$, we have $y_S(\theta') - y_S(\theta) < \frac{1}{2}(y_R(\theta') - y_R(\theta))$.*

Much of the applied literature (e.g. Dye, 1985) studies a special case of this form of preferences, assuming that all sender types prefer the highest possible action while receivers prefer to match the action to the state. This literature typically finds

[14] Dye (1985) also allows for the possibility that the shareholders commit to a disclosure policy, but assumes that this option is not exercised. See Hummel, *et al.* (2016) for an analysis of disclosure with commitment.

that "good news" is revealed whereas "bad news" is suppressed. In this section, we contribute by showing how this good news-bad news feature of equilibria generalizes and what conditions are required for equilibria of this form.

Critical to the analysis is the receiver's response following the null message $\sigma = \emptyset$, which now occurs in equilibrium. Let $y\left(\emptyset; D = [\theta_1, \theta_2]\right)$ denote the equilibrium action following the null message when the sender discloses over the interval $[\theta_1, \theta_2]$. Absent agreement points, a sharp result is available:

Proposition 5 *If there are no agreement points and the gradual slope ordering property holds, then there is a unique equilibrium. In this equilibrium, full disclosure occurs over some interval* $[\theta_1, \theta_2]$ *and non-disclosure results otherwise. Moreover, in this equilibrium either* $\theta_1 = \underline{\theta}$ *or* $\theta_2 = \bar{\theta}$*, but not both.*

The proof follows as a consequence of three propositions. The first shows that a convex disclosure equilibrium exists under the gradual slope ordering property regardless of whether there is an agreement point. The second shows that when agreement points are absent, there is a unique convex disclosure equilibrium. The third shows that, in these circumstances, all equilibria are convex disclosure equilibria. We now establish the first result:

Proposition 6 *If the gradual slope ordering property holds, then there exists a convex disclosure equilibrium.*

Proposition 6 is illustrated in Figure 1. While existing work studies how the threshold for revelation varies with the chance that the sender is informed in a model where sender bliss lines are flat, little is known about how differing sender preferences affect disclosure. An implication of Proposition 6 is that, when agreement points

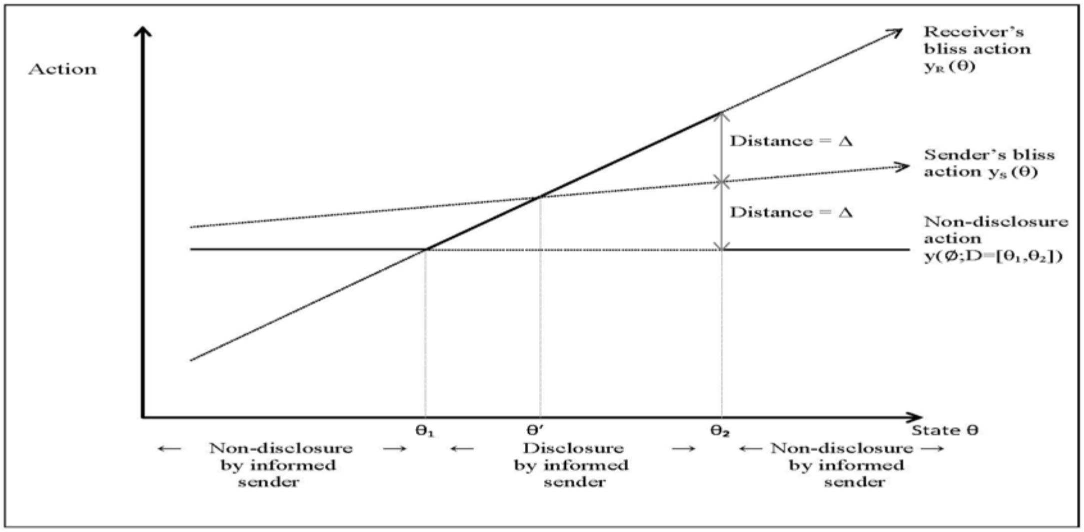

Figure 1: Convex Disclosure Equilibrium: *If the gradual slope ordering property holds, then a convex disclosure equilibrium exists.*

are absent and the gradual slope ordering property holds, equilibrium information disclosure is *independent* of the particulars of sender preferences. We return to this point in more detail in Section 5, but the basic idea is the following: disclosure begins at the point where the receiver's bliss line equals the action taken following the null message. As neither expression depends on sender preferences, equilibrium is undisturbed when these preferences change.

The notion that the two parties might agree in certain states represents a novelty in the modeling of preferences absent from most extant work causing equilibria to differ in fundamental ways from earlier characterizations.[15] Rather than dividing the state space into good news, which is disclosed, and bad news, which is withheld, when preferences exhibit some agreement, a second cutoff can arise. Each cutoff satisfies a

[15] Note though that Bhattacharya and Mukherjee (2013) also allow for agreement points in their analysis but restrict attention to flat sender bliss lines.

similar indifference condition to the more usual case, but together these cutoffs imply the suppression of "extreme" news, consisting of both extremely high and extremely low states, rather than the simple good versus bad news dichotomy. This leaves the receiver in the perplexing situation that, when faced with non-disclosure, the state might be extremely high, extremely low, or simply unknown to all. Thus, unlike the more standard situations where agreement points are absent, a receiver's action following non-disclosure is fraught with considerable risk.

Having identified circumstances where convex disclosure equilibria exist, the next proposition shows that, so long as agreement points are absent, the gradual slope ordering property also guarantees that there is a unique equilibrium cutoff supporting a convex disclosure equilibrium. Formally,

Proposition 7 *If the gradual slope ordering property holds and there are no agreement points, then there is a unique convex disclosure equilibrium.*

The Severity of Information Loss

The contrast between Theorem 1 and Proposition 2 suggests that second-order uncertainty leads to entirely different disclosure regimes. When agreement points are absent and the gradual slope ordering property holds, we can trace how equilibrium disclosure varies in the two regimes. Specifically, the parameter p reflects the degree of second-order uncertainty. In the limit as $p \to 1$, where second-order uncertainty vanishes, we show that the sequence of convex disclosure equilibria converges to full revelation. Formally,

Proposition 8 *If the gradual slope ordering property holds and there are no agreement points, then the unique convex disclosure equilibrium converges to full revelation as* $p \to 1$.

Agreement Points

The presence of agreement points may create additional complexity in that, in some cases, the equilibrium configuration must satisfy two indifference conditions (corresponding to the endpoints of the interior disclosure interval) rather than one. This, in turn, leads to the possibility of equilibrium multiplicity.

To see why, notice that the receiver's response to non-disclosure is non-linear in the parameters θ_1 and θ_2 of the disclosure region. Nonlinearities arise both from the Bayes' rule calculation weighing the likelihood that the sender is uninformed and from the calculation of the conditional expectation of θ under strategic non-disclosure. In general, such non-linear two equation systems produce multiple solutions. Indeed, even in the canonical case where bliss lines are linear and the state is uniformly distributed, multiple equilibria can arise:

Example 1 *Suppose the state is uniformly distributed on* $[-50, 50]$ *and the probability the sender is informed is* $p = \frac{1}{2}$. *The receiver's bliss line is* $y_R(\theta) = \theta$ *while the sender's is* $y_S(\theta) = \frac{9}{20}\theta + \frac{3}{16}$*; thus, the bliss lines satisfy the gradual slope ordering property and there is an agreement point at* $\theta \approx 0.34$. *Finally, suppose the receiver's payoffs are quadratic in the distance from the bliss action. This specification yields exactly two convex disclosure equilibria, one in which the sender discloses if and only if* $\theta \in [-2.58, 29.54]$ *and one in which the sender discloses if and only if* $\theta \in [-0.05, 4.20]$.

The example demonstrates that the presence of agreement points destroys the possibility of equilibrium uniqueness. It also offers two other lessons. First, much like cheap talk games, persuasion games can also produce multiple equilibria with very different informational characteristics. Second, it shows that the presence of agreement points need not prevent considerable information loss.

Thus far, we have limited the search for multiple equilibria to other convex disclosure equilibria. Our next proposition shows that restricting the search in this way captures all possible equilibria, provided there are no agreement points.

Proposition 9 *If the gradual slope ordering property holds and there are no agreement points, then every equilibrium is a convex disclosure equilibrium.*

The intuition for why disclosure intervals must be convex can be most readily seen when the sender's bliss line lies above the receiver's. The key to the result is the structure of how the conflict between the sender and receiver evolves with the state, i.e. the changes in the two bliss lines. Recall that disclosure is preferred if and only if the gap between the two bliss points is smaller than the gap between the sender's bliss action and the non-disclosure action, i.e., the "non-disclosure gap". If, at any two points, the former gap is smaller, then it remains smaller everywhere in between. This is easily seen when the non-disclosure action lies below the sender's bliss action at both points. Since, in these circumstances, the receiver's bliss action must lie between the non-disclosure action and the sender's bliss action, disclosure remains optimal everywhere in between. When the non-disclosure action lies above the sender's bliss action at both points, incentives to disclose depend on the speed with which the bliss line gap closes compared to the non-disclosure gap. The gradual slope ordering property ensures that the bliss line gap closes faster, ensuring disclosure remains optimal in between.

Non-Convex Disclosure

If we restrict the message space, as in Dye (1985), so that the sender can either fully disclose or send the null message, then Proposition 9 extends to cases where an agreement point is present—all equilibria entail a convex disclosure interval. But relaxing this restriction on messages permits the possibility of non-convex disclosure equilibria as the following result shows.

Proposition 10 *If the gradual slope ordering property holds and there is an agreement point θ', then there exists an equilibrium characterized by cutoffs $\theta_1 < \theta_1' < \theta_2' < \theta_2$, where $\theta' \in (\theta_1', \theta_2')$, such that partial disclosure occurs when $\theta \in [\theta_1', \theta_2']$, full disclosure occurs when $\theta \in [\theta_1, \theta_2] / [\theta_1', \theta_2']$, and non-disclosure occurs when $\theta \notin (\theta_1, \theta_2)$.*

Notice that, when an agreement point is present, the gradual slope ordering property also implies that the sender is conservative in the neighborhood of the agreement point. As was the case when the knowledge state was public, conservatism gives rise to multiple equilibria, and fundamentally alters some features of disclosure in persuasion settings. Ironically, the effect of an agreement point in this situation is to *reduce* disclosure by the sender. Indeed, using arguments identical to those in the public knowledge setting for comparing equilibria, it follows that the sender strictly prefers an equilibrium with non-convex disclosure to the convex disclosure equilibrium.

To conclude, despite the potential complexity of messaging and response strategies when the sender is privately informed, there exist equilibria with a simple and intuitive form—the sender hides bad news (from his perspective) and discloses good news. Moreover, when agreement points are absent, the equilibrium is unique. Adding agreement points and enriching the message space, however, can admit radically different types of equilibria.

4.2 Convex Non-Disclosure Equilibrium

We now study equilibria where the non-disclosure region is a convex set. Let $y(\emptyset; ND = [\theta_1, \theta_2])$ denote the equilibrium action following the null message when the *non*-disclosure interval is $[\theta_1, \theta_2]$. A sufficient condition for such equilibria to exist is that the slope of the sender's bliss line is *more* than half that of the receiver's bliss line; we call this the *steep slope ordering property.* Formally,

Definition 5 *The bliss lines $y_S(\theta)$ and $y_R(\theta)$ satisfy the* steep slope ordering property *if for all $\theta' > \theta$, we have $y_S(\theta') - y_S(\theta) > \frac{1}{2}(y_R(\theta') - y_R(\theta))$.*

One might wonder why a cutoff of $\frac{1}{2}$ is used in defining whether the bliss lines satisfy the gradual slope ordering property or the steep slope ordering property. The reason is that if the slope of the sender's bliss line is more than half that of the receiver's, then the difference between the sender's bliss action and the receiver's bliss action changes less rapidly as a function of the state than the difference between the sender's bliss action and the action taken upon non-disclosure. But if the slope of the sender's bliss line is less than half that of the receiver's then the opposite holds. Hence $\frac{1}{2}$ represents the critical value dividing the two cases.

Unlike the gradual slope ordering property, under the steep slope ordering property there may be multiple agreement points. Among other things, this property implies that the sender is not conservative at any agreement point. Thus, our earlier arguments ruling out partially informative messages when the knowledge state is public apply here. Hence, without loss of generality, we restrict attention to full disclosure versus no disclosure.

When preferences satisfy the steep slope ordering property, we first establish that a convex non-disclosure equilibrium always exists and then that *all* equilibria are convex non-disclosure equilibria. Therefore, restricting equilibrium search to this class is of no consequence. One might then be tempted to conjecture that, like convex disclosure equilibria, the absence of agreement points leads to uniqueness. Sadly, this is not the case—regardless of the presence or absence of agreement points, multiple equilibria can arise.

To establish equilibrium existence, the following definition and lemma are helpful.

Definition 6 *Let $\theta_\emptyset$ be the state θ solving $y_R(\theta) = y(\emptyset)$, where $y(\emptyset)$ is the receiver's optimal action given her prior beliefs.*

Lemma 1 *Define θ_L to be the largest agreement point $\theta' < \theta_\emptyset$, if such a point exists, and $\theta_L = \underline{\theta}$ otherwise. Likewise, define θ_H to be the smallest agreement point $\theta'' > \theta_\emptyset$, if such a point exists, and $\theta_H = \bar{\theta}$ otherwise. When the steep slope ordering property holds and $y_S(\theta) > y_R(\theta)$ for all $\theta \in (\theta_L, \theta_H)$, there exists a non-disclosure equilibrium $[\theta_1, \theta_2] \subseteq [\theta_L, \theta_H]$ solving*

$$|y(\emptyset; ND = [\theta_1, \theta_2]) - y_S(\theta_1)| \leq |y_S(\theta_1) - y_R(\theta_1)| \quad (1)$$

$$y(\emptyset; ND = [\theta_1, \theta_2]) = y_R(\theta_2) \quad (2)$$

where (1) holds with equality if $\theta_1 > \theta_L$.

We now establish the existence result.

Proposition 11 *If the steep slope ordering property holds, then there exists a convex non-disclosure equilibrium.*

Proposition 11 is illustrated in Figure 2. There are many plausible situations where the steep slope ordering property arises. Perhaps the most obvious is the canonical cheap talk specification wherein a decision maker consults a biased expert who may report freely her views as to the state. The most explored case of this model occurs when the bliss lines of the decision maker and expert are parallel. Comparing the situation of cheap talk with persuasion, we see that they share in common the

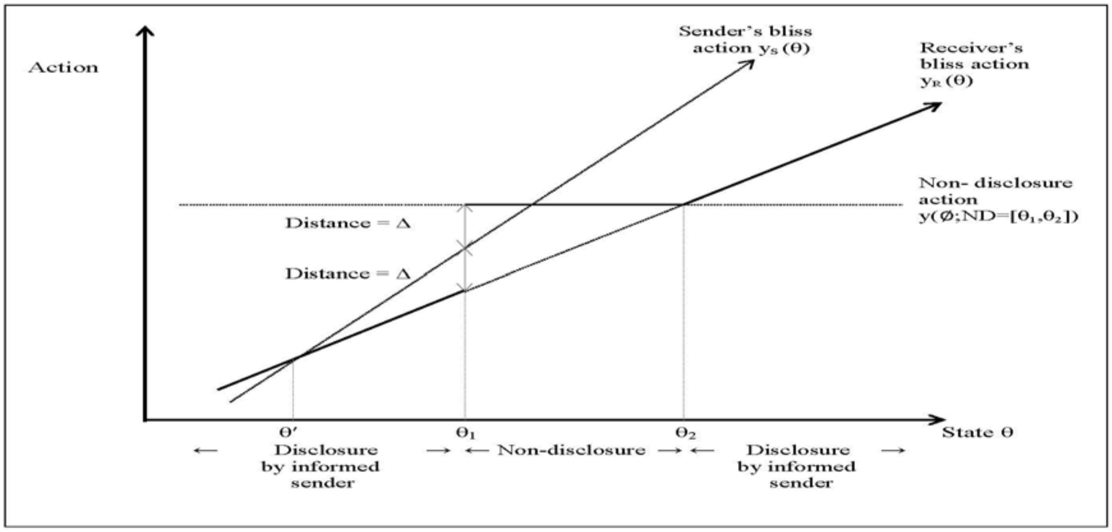

Figure 2: Convex Non-Disclosure Equilibrium: *If the steep slope ordering property holds, then a convex non-disclosure equilibrium exists.*

inevitable loss of information, but that this loss occurs via a convex non-disclosure interval under persuasion rather than partitional equilibria under cheap talk.

Perhaps more interesting are situations where the sender's bias switches direction; that is, where there is an agreement point. For instance, suppose θ is drawn from a distribution with mean $\mu > 0$. The receiver wishes to match the state while the sender prefers an action equal to $\alpha\theta$, where $\alpha > 1$. Such a setting might occur when a policy maker consults an expert to deliver a fact-laden report, and the expert is more ideologically polarized than the policy maker. Proposition 11 reveals that, even though incentives grow arbitrarily misaligned as the state becomes extreme in either direction, the expert discloses for extreme values of θ. By contrast, in states over a subset of $[0, \mu]$ where there is little conflict, no disclosure occurs.

Why does the steep slope ordering property change the nature of equilibrium disclosure so starkly compared to the gradual slope ordering property? The differing workings of these properties are most easily seen for extreme states. Under the gradual slope ordering property, the receiver prefers more extreme actions than the sender, at least near one of the endpoints. Hence, the sender resorts to non-disclosure, which produces a moderate action in response. By contrast, under the steep slope ordering property, preferences are not as misaligned at the extremes. Indeed, the sender may even prefer more extreme actions than the receiver. Accordingly, the option of moderation holds no interest and disclosure is best. The opposite is true for moderate states, where non-disclosure itself can represent the extreme choice. As a consequence, it offers a useful counter for senders that are relatively responsive to the state, but not for those that are relatively unresponsive.

We now characterize the qualitative properties of *all* equilibria when the steep slope ordering property holds.

Proposition 12 *If the steep slope ordering property holds, then every equilibrium is a convex non-disclosure equilibrium.*

Examples of Multiple Convex Non-Disclosure Equilibria

Unlike the situation of convex disclosure equilibria, where equilibrium was unique in the absence of agreement points, no such property obtains for convex non-disclosure equilibria. To illustrate this, suppose the state is uniformly distributed on $[-50, 50]$, the receiver suffers quadratic losses in the difference between her action and the state, the sender is informed with probability $p = \frac{3}{4}$, and has a bliss line:

$$y_S(\theta) = \begin{cases} \theta + 1 & \text{if} \quad \theta \geq -2.06 \\ 0.528\,41\theta + 0.028\,525 & \text{if} \quad \theta < -2.06 \end{cases} \tag{3}$$

Notice that the slope of $y_S(\theta)$ is greater than half the slope of $y_R(\theta)$ and therefore the steep slope ordering property holds. Moreover, there are no agreement points.

It may be readily verified that non-disclosure over the interval $[-2.06, -0.060]$ and disclosure elsewhere comprises a convex non-disclosure equilibrium. This construction depends only on preferences in the upper region, $\theta \geq -2.06$. But there is another non-disclosure equilibrium that has an upper bound $\theta_2 = -0.061$. Since a necessary condition in any equilibrium is that $y(\emptyset; ND = [\theta_1, \theta_2]) = y_R(\theta_2)$, the lower bound of the non-disclosure interval must be $\theta_1 = -2.0776$. Equilibrium also requires that the sender is indifferent between disclosing and revealing when $\theta = \theta_1$. To ensure this relation, we identified sender preferences to satisfy this condition while maintaining continuity. This accounts for the part of the sender's bliss line where $\theta < -2.06$.

The choice of $\theta_2 = -0.061$ was arbitrary. For any θ_2 slightly lower than -0.061, a similar construction of sender preferences below -2.06 produces a second equilibrium without disturbing the equilibrium characterized by non-disclosure over the interval $[-2.06, -0.060]$. In short, the example represents neither a knife-edge nor a pathological case.

Although the example does not feature any agreement points, adding agreement points in no way eliminates multiple equilibria. For instance, if we modify preferences for sufficiently high values of θ such that one or more agreement points arise (while maintaining the steep slope ordering property), the equilibria we identified remain undisturbed because adding agreement points does not alter the incentives to disclose.

5 Choosing Among Senders

While the model thus far envisages an exogenous assignment of a sender to a receiver, in many settings the receiver can choose between senders. For instance, when the sender represents a consultant and the receiver a CEO, the receiver will typically interview several senders before deciding on one. We study a receiver's choice of a sender when senders are equally likely to be informed (have the same value of p) but differ in their preferences. Taking a mechanism design perspective, we resolve situations where there are multiple equilibria by choosing the equilibrium most favored by the principal/receiver. To capture preference differences, suppose the receiver can consult with exactly one of two senders, labeled S_1 and S_2, and ordered by their alignment with the receiver.

Definition 7 *Sender S_1 is* ***more aligned*** *with the receiver than sender S_2 if and only if $y_{S_1}(\theta)$ falls strictly between $y_R(\theta)$ and $y_{S_2}(\theta)$ for all non-agreement points θ.*

While intuitive, alignment is a strong ordering property—it requires that, for all states where the sender and receiver disagree, the gap between bliss points of the more aligned sender and the receiver is smaller than that of the less aligned sender and the receiver. The strength of this condition is that the ordering between the two senders is insensitive to the state distribution.

Given this ordering, it seems apparent that the receiver will always confer with the more aligned sender, even if required to pay a small cost to do so. Results under cheap talk games reinforce this conclusion. Indeed, a key finding from Crawford and Sobel (1982) is that the receiver should confer with the least biased agent possible. Under persuasion, however, this is not necessarily the case. In some cases, the intuitive result applies, yet in surprisingly many situations the receiver would not pay any amount, even if arbitrarily small, to consult with a more aligned rather than a less aligned sender.

5.1 Sender Knowledge Is Public Information

When the sender's knowledge state is public, the comparison between senders is straightforward. Since full revelation is an equilibrium regardless of preferences, the receiver gains nothing by consulting with the more aligned sender.

Proposition 13 *If the knowledge state is public and sender preferences are bounded, then the receiver is indifferent over all senders, regardless of the degree of alignment.*

5.2 Sender Knowledge Is Private Information

When the knowledge state is private, full revelation is no longer possible and, as we established in Section 4, the qualitative features of equilibrium vary intricately with whether sender preferences satisfy the gradual slope ordering property or the steep slope ordering property. Throughout we assume one of these two properties holds for both senders. The question of whom to consult depends on the properties of the sender preferences. Accordingly, we divide the analysis into two parts.

5.2.1 Gradual Slope Ordering Property

Under the gradual slope ordering property, the best equilibrium for the receiver is a convex disclosure equilibrium. We divide the analysis by equilibrium types, boundary or interior. A boundary convex disclosure equilibrium is one in which one of the endpoints of the disclosure interval is also an endpoint of the state space. An interior convex disclosure equilibrium is one in which the disclosure interval lies in the strict interior of the state space. An analogous definition holds for boundary and interior convex non-disclosure equilibria.

Boundary Equilibrium

When the equilibrium under the less aligned sender is a boundary equilibrium, the analysis is stark. The receiver derives no additional benefit from consulting a more aligned sender, just as when the knowledge state was public information.

Proposition 14 *If the knowledge state is private, the gradual slope ordering property holds, and there is a boundary equilibrium when consulting the less aligned sender S_2, then the receiver cannot do any better by consulting a more aligned sender S_1.*

We prove this proposition in two parts. The first shows that, given a boundary equilibrium under the less aligned sender S_2, there are no better boundary equilibria under the more aligned sender S_1, regardless of how closely aligned this sender is with the receiver. The second part shows that the receiver strictly prefers a boundary equilibrium to any interior equilibrium under a given sender. Together, these two findings imply Proposition 14.

While the presence of a boundary equilibrium is sufficient, it is not necessary for the observation that the receiver gains nothing by consulting a more aligned sender. Suppose that we weaken the alignment property to a weak inequality; that is, the gap between bliss points for a more aligned sender is at least as small as that of the less aligned sender, but need not be strictly smaller for all states. With this weakening, this finding can arise even with interior equilibria.

Proposition 15 *If the knowledge state is private, the gradual slope ordering property holds, there is an agreement point θ', $y(\emptyset) > (<)\, y_R(\theta')$, and $y_{S_1}(\theta) = y_{S_2}(\theta)$ for all $\theta \geq (\leq)\, \theta'$, then the receiver is indifferent between the senders.*

Proposition 15 demonstrates that it is not the presence of boundary equilibria *per se*, that produces neutrality; rather it is the fact that the sender's preferences only matter in a subset of the state space. Such a situation cannot arise under alignment as we have defined it, where a strict ordering holds for all states. Proposition 15 illustrates why that condition cannot be easily weakened.

Now suppose that, instead of choosing among senders by alignment, the receiver's choice concerns sender expertise, reflected by differences in p. In this situation, we can show:

Proposition 16 *If the senders share the same bliss line $y_S(\theta)$, there are no agreement points, the gradual slope ordering property holds, and S_1 has more expertise than S_2 in that $p_1 > p_2$, then the receiver strictly prefers to consult sender S_1.*

An implication of this result is that a receiver obtains more value from consulting a more informed sender than a sender with more aligned preferences. Consulting a more informed sender improves the receiver's payoff whereas consulting a more aligned sender does not.

Interior Equilibrium

We now study interior equilibria when senders are ordered by alignment. In contrast to our earlier findings, neutrality no longer holds. Instead, the receiver strictly prefers to consult the more aligned sender. To establish this result, we require that the receiver's preferences be well-behaved in the following sense:

Definition 8 *The receiver's indifference condition is said to be* ***regular*** *if, for any θ_2, the value of θ_1 solving $y_R(\theta_1) = y(\emptyset; D = [\theta_1, \theta_2])$ has the property that $\frac{d\theta_1}{d\theta_2} < 0$ where the disclosure interval is $[\theta_1, \theta_2]$.*

The regularity condition says that, if disclosure occurs over slightly higher states (i.e., θ_2 goes up), the receiver's optimal action under non-disclosure falls and hence the point at which the receiver is indifferent between the disclosure and non-disclosure action also falls. It ensures the receiver's optimal action under non-disclosure behaves in a sensible fashion. Commonly used expressions of receiver preferences satisfy the regularity condition. For instance, regularity holds whenever $y(\emptyset; D = [\theta_1, \theta_2]) = E[\theta|\sigma = \emptyset]$, which occurs, among other cases, when the receiver has quadratic preferences.[16]

Proposition 17 *If the knowledge state is private, the gradual slope ordering property holds, there is an agreement point, and the receiver's preferences are regular, then for any interior equilibrium under the less aligned sender, S_2, there exists a strictly more informative equilibrium under the more aligned sender, S_1.*

5.2.2 Steep Slope Ordering Property

We again divide the analysis by whether the equilibria are boundary or interior.

Boundary Equilibrium

Whereas in the previous section, the presence of a boundary equilibrium under S_2 implied that such an equilibrium arose when consulting S_1, this is no longer the case for convex non-disclosure equilibria. Instead, the opposite implication holds. The other key difference is that boundary equilibria need not be preferred to interior equilibria for convex non-disclosure equilibria. Despite these differences, there are settings where the receiver derives no incremental benefit from consulting the more aligned sender.

[16] Proof that $y(\emptyset; D = [\theta_1, \theta_2]) = E[\theta|\sigma = \emptyset]$ implies the regularity of receiver preferences is available on request.

Proposition 18 *If the unique equilibrium under the more aligned sender, S_1, is a boundary equilibrium, then the same strategies are also an equilibrium when consulting the less aligned sender, S_2.*

The following example demonstrates that the opposite implication, i.e. the presence of a unique boundary equilibrium under S_2, does not imply that there is no value to consulting a more aligned sender.

Example 2 *Suppose the state is uniformly distributed on the unit interval, the receiver's preferences satisfy $U_R(y,\theta) = -(y-\theta)^2$, and the sender's bliss line satisfies $y_S(\theta) = \theta + b$ for some $b > 0$. When $b \geq \frac{1}{2} - \frac{1}{2p} + \frac{\sqrt{1-p}}{2p}$, the unique equilibrium is characterized by non-disclosure in $\left[0, 1 - \frac{1}{p} + \frac{\sqrt{1-p}}{p}\right]$. When $b < \frac{1}{2} - \frac{1}{2p} + \frac{\sqrt{1-p}}{2p}$, the unique equilibrium is characterized by non-disclosure in $\left[\frac{1}{2} - 2b^2\frac{p}{1-p} - 2b, \frac{1}{2} - 2b^2\frac{p}{1-p}\right]$.*

In this example, lower values of b indicate more aligned senders. If $b_2 > \frac{1}{2} - \frac{1}{2p} + \frac{\sqrt{1-p}}{2p} > b_1$, then there is a unique boundary equilibrium under S_2 and there is a unique, and more informative, interior non-disclosure equilibrium under S_1. Thus, the receiver benefits by consulting the more aligned sender.

Interior Equilibrium

We now study situations where there is an interior convex non-disclosure equilibrium when consulting the more aligned sender S_1. As we will show, this situation is analogous to that under the gradual slope ordering property, but no longer requires the regularity condition. Such a weakening occurs even though comparing equilibria between senders is less straightforward than under the opposite slope ordering. Whereas interior convex disclosure equilibria were always subsets or supersets of one another, and hence directly comparable, this is not the case with non-disclosure equilibria. Instead, changes in sender preferences provoke shifts in the location of the non-disclosure interval as well as adjustments in the length of the interval. In principle, this makes comparison difficult or even ambiguous; however, a key feature of the equilibrium conditions enables us to resolve the comparison unambiguously. In particular, when the sender's bliss actions lie above those of the receiver in the relevant non-disclosure interval, then the receiver's action at the right-hand endpoint of the interval coincides with her bliss action. This implies that small changes to this endpoint have only second-order effects on the receiver's payoffs while shifts in the left-hand endpoint have first-order effects.

We study non-disclosure intervals of the form $[\theta_1, \theta_2(\theta_1)]$ where $\theta_2(\theta_1)$ solves $y(\emptyset; ND = [\theta_1, \theta_2(\theta_1)]) = y_R(\theta_2(\theta_1))$; that is, the right-hand endpoint satisfies the equilibrium condition (provided $y_S(\theta) > y_R(\theta)$), but the left-hand endpoint is treated as a parameter. This class nests all interior equilibrium intervals. We show in the next lemma that as long as the non-disclosure interval shifts to the right, the receiver is better off.

Lemma 2 *If the knowledge state is private and the steep slope ordering property holds, then for any non-disclosure interval* $[\theta_1, \theta_2(\theta_1)]$*, the receiver's utility is increasing in* θ_1*.*

We now prove that, when consulting with the more aligned sender, a non-disclosure equilibrium that lies to the right of that under a less aligned sender always exists.

Proposition 19 *If the knowledge state is private, the steep slope ordering property holds, and the receiver's preferred equilibrium under the less aligned sender,* S_2*, is interior, then the receiver strictly prefers to consult with the more aligned sender,* S_1*.*

To summarize, when considering boundary equilibria, sender loyalty is irrelevant. By contrast, under an interior equilibrium, the receiver strictly prefers a more aligned sender. The latter also contrasts with the situation where sender knowledge is public, in which case alignment is irrelevant as well.

6 Conclusion

Since the early 1980s, persuasion games have been widely studied. General models have been offered in situations of information transmission where the receiver knows exactly how knowledgeable the sender is concerning the issue of interest. Yet in many situations the sender's knowledge is unclear. Here, the receiver is not only uncertain about the realized state, but also about the degree to which the sender knows the realized state. Our main contribution is to offer a general model of persuasion games exhibiting this type of second-order uncertainty. We show that the presence of such uncertainty qualitatively changes conclusions about the degree to which hard information facilitates information transfer.

Most notably, such uncertainty destroys the possibility that truthful revelation is an equilibrium when the preferences of the sender and receiver conflict. Instead, equilibrium consists of non-trivial intervals of disclosure and non-disclosure that depend not just on the degree of conflict, but on the relative sensitivity of the sender and receiver bliss actions to changing information. When the sender is relatively insensitive to information, disclosure occurs over a interval and non-disclosure otherwise. When the sender's bliss action is relatively sensitive to the state, equilibria take the reverse form, with non-disclosure occurring over a interval.

Information loss stems from two sources: the degree of preference misalignment and the likelihood that the sender is informed about the realized state. When the sender is highly likely to be well-informed, information loss is minor, even when preferences are badly misaligned. When the receiver is maximally uncertain about the sender's expertise, however, information loss can be severe—even when preferences are relatively well-aligned.

Equilibria also differ in another key respect: whether they are interior or boundary. This distinction proves crucial in examining situations where the receiver must choose whom to consult. When sender alignment can be strictly ranked in that one sender's bliss line is strictly closer to the receiver's than another sender for all states, the choice of sender would seem to be obvious—the receiver should choose the more aligned sender. But when boundary equilibria arise, this proves mistaken. In these situations, the amount of disclosure is independent of preference alignment, but not of sender expertise. Thus, the receiver should choose the more expert sender, regardless of preference alignment. Moreover, the requirements for boundary equilibria are quite mild: So long as the sender and receiver never agree on the ideal action and the sender's bliss actions are relatively insensitive to new information, then *all* equilibria are boundary. Only under an interior equilibrium does closer alignment provide more information.

To conclude then, the practical implication of our analysis is that the knowledgeability of a sender, his likelihood of being informed as to the state, matters much more than preference alignment, especially in situations where sender bliss actions are relatively insensitive to the state. Thus, decision makers should focus on knowledge over loyalty when interviewing prospective senders. More broadly, settings with second-order uncertainty behave quite differently from the standard analysis, where the informedness of the sender is commonly known. Full revelation no longer prevails, and equilibrium structure changes dramatically depending on the relative *slopes* of the bliss lines.

References

[1] Bhattacharya, S., M. Goltsman, and A. Mukherjee. 2015. On the optimality of diverse expert panels in persuasion games. University of London working paper.

[2] Bhattacharya, S., and A. Mukherjee. 2013. Strategic information revelation when experts compete to influence. RAND Journal of Economics 44 (3): 522–544.

[3] Crawford, V. P., and J. Sobel. 1982. Strategic information transformation. Econometrica 50 (6): 1431-1451.

[4] Dye, R. A. 1985. Disclosure of nonproprietary information. Journal of Accounting Research 23 (1): 123-145.

[5] Dye, R. A. 1998. Investor sophistication and voluntary disclosures. Review of Accounting Studies. 3 (3): 261-287.

[6] Eggen, D., and M. A. Fletcher. 2007. Embattled Gonzales resigns. Washington Post (Tuesday, August 28, 2007): http://www.washingtonpost.com/wp-dyn/content/article/2007/08/27/AR2007082700372.html.

[7] Eggen, D., and P. Kane. 2007. Senators chastise Gonzales at hearing. Washington Post (Friday, April 20, 2007): http://www.washingtonpost.com/wp-dyn/content/article/2007/04/19/AR2007041902935.html.

[8] Giovannoni , F., and D. J. Seidmann. 2007. Secrecy, two-sided bias and the value of evidence. Games and Economic Behavior 59 (2): 296–315.

[9] Grossman, S. 1981. The role of warranties and private disclosure about product quality. Journal of Law and Economics 24: 461-483.

[10] Guttman, I., I. Kremer, and A. Skrzypacz. 2014. Not only what but also when: a theory of dynamic voluntary disclosure. American Economic Review 104 (8): 2400-2420.

[11] Hagenbach, J., F. Koessler, and E. Perez-Richeti. 2014. Certifiable pre-play communication: full disclosure. Econometrica 82 (3): 1093–1131

[12] Hummel, P., J. Morgan, and P. C. Stocken. 2016. Optimal firm (non-)disclosure. University of California, Berkeley working paper.

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

[14] Koessler, F. 2003. Persuasion games with higher-order uncertainty. Journal of Economic Theory 110 (2): 393–399.

[15] Lewis, N. A. 2008. In searching for new job, Gonzales sees no takers. New York Times (April 13, 2008): http://www.nytimes.com/2008/04/13/washington/13gonzales.html?_r=0.

[16] Mathis, J. 2008. Full revelation of information in Sender–Receiver games of persuasion. Journal of Economic Theory 143(1): 571–584.

[17] Milbank, D. 2007. Maybe Gonzales won't recall his painful day on the hill. Washington Post (Friday, April 20, 2007): http://www.washingtonpost.com/wp-dyn/content/article/2007/04/19/AR2007041902571.html.

[18] Milgrom, P. 1981. Good news and bad news: representation theorems and applications. Bell Journal of Economics 12 (2): 380-391.

[19] Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura. 1990. Strategic information revelation. Review of Economic Studies 57 (1): 25-47.

[20] Pae, S. 2005. Selective disclosures in the presence of uncertainty about information endowment. Journal of Accounting and Economics 39 (3): 383–409.

[21] Penno, M. C. 1997. Information quality and voluntary disclosure. The Accounting Review 72 (2): 275-284.

[22] Seidmann, D., and E. Winter. 1997. Strategic information transmission with verifiable messages. Econometrica 65 (1): 163–170.

[23] Shin, H. S. 1994a. News management and the value of firms. RAND Journal of Economics 25 (1): 58-71.

[24] Shin, H. S. 1994b. The burden of proof in a game of persuasion. Journal of Economic Theory 64 (1): 253–264.

[25] Shin, H. S. 2006. Disclosure risk and price drift. Journal of Accounting Research 44 (2): 351–379.

[26] Wilber, D. Q. and A. E. Marimow. 2012. Roger Clemens acquitted of all charges. Washington Post (Monday, 18 June 2012): https://www.washingtonpost.com/local/crime/roger-clemens-trial-verdict-reached/2012/06/18/gJQAQxvzlV_story.html.

Appendix

Proof of Proposition 1: Suppose not. Consider case (a) in which $y_R(\theta) > y_S(\theta)$ for all $\theta \geq \theta^*$ for some θ^*. Suppose the sender sends a message $\sigma = [\theta^*, \infty)$ and the receiver's response is $y(\sigma) = y_R\left(\hat{\theta}\right)$ for some $\hat{\theta} \geq \theta^*$. All sender types with $\theta \in \left(\hat{\theta}, \hat{\theta} + \varepsilon\right)$ for some small $\varepsilon > 0$ prefer to send the message σ rather than fully reveal yielding a contradiction. Case (b) is analogous. ■

Proof of Proposition 2: Consider the following three cases:

Case I (terminal branch was UUU...U). Since $y(\sigma) = y_R(\theta_{n+1})$, for all $\theta \in [\theta_0, \theta_{n+1}]$ where $y_S(\theta) < y_R(\theta)$, we have $y_S(\theta) < y_R(\theta) < y_R(\theta_{n+1})$. Hence, these types have no incentive to deviate. Now consider $\theta \in I_k$ where $y_S(\theta) > y_R(\theta)$. Since $I_k \notin \bar{I}$ and $\bar{I}_{k^{(m)\prime}} = I_{n+1}$, there exists some interval I_l for $l > k$ whose gap is at least as large as the gap in I_k. Thus, for $\theta \in I_k$, $|y_R(\theta) - y_S(\theta)| \leq |y_R(\theta_l^*) - y_S(\theta_l^*)| \leq |y_R(\theta_{n+1}) - y_S(\theta_l^*)| < |y_R(\theta_{n+1}) - y_S(\theta)|$. The first inequality follows because the gap in I_l is at least as large as that in I_k. The second inequality follows because $y_R(\theta_{n+1}) \geq y_R(\theta_l^*)$ and because, for any interval in the set of bifurcation intervals, $y_S(\theta) < y_R(\theta)$. The third inequality follows because $\theta_l^* > \theta$ for all $\theta \in I_k$. Therefore, deviating to σ is not profitable.

Case II (terminal branch was DDD...D). This case is analogous to Case I.

Case III (terminal branch contains at least one U and one D). Recall that $\theta_K \in \left\{\theta_{k^{(m)\prime}-1}, \theta_{k^{(m)\prime}}\right\}$. We prove that full revelation is an equilibrium when, for $\theta \in \bar{I}_{k^{(m)\prime}}$, $y_S(\theta) < y_R(\theta)$, and hence $y(\sigma) = y_R(\theta_{k^{(m)\prime}})$. The case where $y_S(\theta) > y_R(\theta)$ for $\theta \in \bar{I}_{k^{(m)\prime}}$ is analogous.

Case IIIA: Suppose $\theta < \theta_K = \theta_{k^{(m)\prime}}$. If $y_S(\theta) < y_R(\theta)$, then full revelation is incentive compatible since $y_S(\theta) < y_R(\theta) < y_R(\theta_{k^{(m)\prime}}) = y(\sigma)$. It remains to show incentive compatibility for $\theta \in I_k$ where $y_S(\theta) > y_R(\theta)$. Notice that there exists some interval I_l with $k < l \leq K$ with a larger gap than I_k and with the property that, for $\theta \in I_l$, $y_S(\theta) < y_R(\theta)$. The reason is that at some prior bifurcation we selected an interval I_l with $k < l \leq K$ such that for $\theta \in I_l$, $y_S(\theta) < y_R(\theta)$. When I_l was selected I_k was available but not chosen implying that the gap in I_k is less than that in I_l. For $\theta \in I_k$, observe that $|y_R(\theta) - y_S(\theta)| \leq |y_R(\theta_l^*) - y_S(\theta_l^*)| \leq |y_R(\theta_K) - y_S(\theta_l^*)| < |y_R(\theta_K) - y_S(\theta)|$. The first inequality follows because the gap in I_l is at least as large as that in I_k. The second inequality follows because $y_R(\theta_K) \geq y_R(\theta_l^*)$ and $y_S(\theta_l^*) < y_R(\theta_l^*)$. And the third inequality follows because $\theta < \theta_l^*$ for all $\theta \in I_k$. Therefore, deviating to σ is not profitable.

Case IIIB: Suppose $\theta > \theta_K = \theta_{k(m)'}$. The argument establishing that deviating is not profitable is analogous to that of Case IIIA.

Since this exhausts all of the possibilities, the proof is complete. ■

Proof of Proposition 3: We first show that if the sender is not conservative, then full revelation is the unique equilibrium. When there are no agreement points, the not conservative condition is satisfied vacuously. Moreover, it follows from Seidmann and Winter (1997), Theorem 3 (part a) that the equilibrium is unique. When there are one or more agreement points, Proposition 2 showed that full revelation is an equilibrium. We now show that, when the sender is not conservative, no equilibrium exists in which partial pooling of information occurs.

Suppose, contrary to the proposition, that there exists an equilibrium without full revelation. Fix a message $\sigma = I$, where I consists of a non-degenerate interval $[\theta_0, \theta_1]$. Define a set of positive measure $\Theta(\sigma)$ such that, for all $\theta \in \Theta(\sigma)$, the message σ is sent in equilibrium. Let $C(\Theta(\sigma))$ be the convex hull of $\Theta(\sigma)$. Let θ' be the largest agreement point in $C(\Theta(\sigma))$ such that the set $\{\theta : \theta > \theta' \text{ and } \theta \in \Theta(\sigma)\}$ has positive measure if such an agreement point exists, and let $\theta' = \inf C(\Theta(\sigma))$ otherwise.

Case 1: Suppose that for (almost) all $\theta > \theta'$ in $C(\Theta(\sigma))$, $y_R(\theta) < y_S(\theta)$. Then, since the receiver believes that all types are contained in $C(\Theta(\sigma))$ following the equilibrium message σ, it then follows that $y(\sigma) = y_R(\theta'')$ for some value of θ'' strictly in the interior of $C(\Theta(\sigma))$ and furthermore, there exist a positive measure of values of $\theta \in \Theta(\sigma)$ such that $\theta > \max\{\theta', \theta''\}$. These sender types can deviate by revealing truthfully, thus inducing an action $y_R(\theta) > y(\sigma)$, which is strictly profitable since $y_S(\theta) > y_R(\theta) > y(\sigma)$.

Case 2: Suppose that for (almost) all $\theta > \theta'$ in $C(\Theta(\sigma))$, $y_R(\theta) > y_S(\theta)$. As in the previous case, the putative equilibrium action $y(\sigma)$ lies strictly in the interior of $C(\Theta(\sigma))$. If $y(\sigma) \leq y_R(\theta')$, then for a positive measure of types θ where $\theta \in \Theta(\sigma)$ and $\theta > \theta'$, we have $U_S(y_R(\theta), \theta) > U_S(y_R(\theta'), \theta) \geq U_S(y(\sigma), \theta)$. Hence, they prefer full revelation to $y_R(\theta')$.

If $y(\sigma) > y_R(\theta')$, then either there exists a positive measure of sender types θ where $\theta \in \Theta(\sigma)$ and $\theta < \theta'$ or there exists a positive measure of sender types θ where $\theta \in \Theta(\sigma)$ and $\theta \in \left(\theta', y_R^{-1}(y(\sigma))\right)$. In the former case, $y_S(\theta) < y_R(\theta') < y(\sigma)$ for any such types θ, and hence, $U_S(y_R(\theta), \theta) > U_S(y_R(\theta'), \theta) > U_S(y(\sigma), \theta)$. In the latter case, $y_S(\theta) < y_R(\theta) < y(\sigma)$ for any such sender type θ, and hence $U_S(y_R(\theta), \theta) > U_S(y(\sigma), \theta)$. Thus, in either case, a positive measure of senders can profitably deviate by revealing truthfully. Since this exhausts all possibilities, the result follows.

We now show that if at some agreement point a positive measure of sender types are conservative, then there exists a partial pooling equilibrium. We do this by construction.

First, suppose that a positive measure of conservative senders lie on both sides of agreement point θ'. Let $S_L = \{\theta : \theta < \theta'$ and sender is conservative$\}$ and $S_H = \{\theta : \theta > \theta'$ and sender is conservative$\}$.

We will construct an interval I containing a positive measure of conservative senders on both sides of θ' such that, when the message $\sigma = I$ is sent, in equilibrium, the action $y(\sigma) = y_R(\theta')$ is chosen. Formally, choose an interval I such that $I \cap S_L$ and $I \cap S_H$ are of positive measure. Suppose that in equilibrium, for all $\theta \in (I \cap S_L) \cup (I \cap S_H)$ senders send the message $\sigma = I$. All other senders fully reveal.

We now show that one can choose I such that, given equilibrium posterior beliefs, the action $y(\sigma) = y_R(\theta')$ maximizes the receiver's payoffs. Clearly I may be constructed to contain a positive measure of conservative senders on both sides of θ'. To show that it produces the action $y_R(\theta')$ in equilibrium, notice that, if $I \cap S_L$ is sufficiently small, then the receiver will optimally choose $y(\sigma) > y_R(\theta')$, while if $I \cap S_H$ is sufficiently small, then the receiver will optimally choose $y(\sigma) < y_R(\theta')$. By continuity of the receiver's best response, there exists I containing a positive measure of conservative senders on both sides of θ' such that $y(\sigma) = y_R(\theta')$. Thus, such a construction is feasible.

To see that this construction is incentive compatible, suppose that the receiver uses the Bifurcation Algorithm to respond to out-of-equilibrium messages. For non-conservative senders, this ensures that deviation is unprofitable. For conservative senders where $\theta \notin I$, the message σ is not feasible and, by construction, they prefer full revelation to any feasible deviation. Finally, for conservative senders such that $\theta \in I$, by construction, full revelation is weakly preferred to any deviation and, by the definition of conservatism, pooling and obtaining $y_R(\theta')$ is strictly preferred to full revelation. Therefore, choosing $\sigma = I$ is incentive compatible. Finally, the receiver is acting optimally given beliefs on and off the equilibrium path. Thus, we have constructed an equilibrium.

Next, suppose that a positive measure of senders are conservative on only one side of agreement point θ'. Suppose that a positive measure of sender types below θ' are conservative while a zero measure of senders above θ' are conservative; that is, S_L is of positive measure and S_H of zero measure. Consider a point $\theta'' > \theta'$ but close to it. Notice that, for θ in the interval $(\theta'', \theta'' + \varepsilon]$ for some small $\varepsilon > 0$, we have the ordering $y_S(\theta) < y_R(\theta'') < y_R(\theta)$ since $y_R(\theta)$ is strictly increasing. Therefore, in this interval $U_S(y_R(\theta'), \theta) \leq U_S(y_R(\theta), \theta) < U_S(y_R(\theta''), \theta)$. Moreover, since θ'' is close to θ', then, by continuity, for a positive measure $\theta < \theta'$ where the sender is conservative with respect to θ', the sender is also conservative with respect to θ''. We can then use an analogous construction to the case where senders are conservative on both sides of an agreement point to establish a partial pooling equilibrium.

The case where there are a positive measure of conservative senders above θ' and not below is analogous. ■

Proof of Theorem 1: Suppose to the contrary that full revelation is an equilibrium. We will derive a contradiction by constructing a profitable deviation. Clearly there exists a state θ' where $y_R(\theta') = y(\emptyset)$. Since, generically $y_S(\theta') \neq y_R(\theta')$ and $y_R(\theta)$ is continuous and strictly increasing, there is a positive measure of sender types near θ' who will prefer to report that they are uninformed and induce action $y(\emptyset)$, than informed and induce action $y_R(\theta)$. Therefore, full revelation is not an equilibrium. ■

Proof of Proposition 6: The gradual slope ordering property implies that there is no more than one agreement point. We can assume that the sender either fully discloses or engages in complete non-disclosure. If the sender selects a message $\sigma = I$, where I is a non-degenerate interval, then this message immediately reveals the sender's information state. Thus, arguments analogous to those supporting Proposition 2 imply that the receiver can respond in such a way that the sender prefers full disclosure to partial disclosure. As a result, the remainder of the proof restricts attention to incentive compatibility of full versus no disclosure.

First, assume there are no agreement points. Suppose $y_S(\theta) > y_R(\theta)$ for all θ; the opposite case is analogous. Define θ^* as follows: An informed sender sends the message $\sigma = \emptyset$ for $\theta < \theta^*$ and sends the message $\sigma = \theta$ for $\theta \geq \theta^*$. Following the message $\sigma = \emptyset$, the receiver's action $y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right)$ maximizes her expected payoff conditional on the message coming from a sender who is uninformed with probability $\frac{1-p}{1-p+pF(\theta^*)}$, and from a sender who is informed and where the state is $\theta < \theta^*$ with the remaining probability. The value of θ^* is defined to satisfy $y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right) = y_R(\theta^*)$.

To establish that such a θ^* exists, notice that when $\theta^* \to \underline{\theta}$ or $\theta^* \to \bar{\theta}$, the action $y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right)$ reflects the optimal action conditional on the sender being uninformed and, therefore, $\lim_{\theta^* \to \underline{\theta}} y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right) > \lim_{\theta^* \to \underline{\theta}} y_R(\theta^*)$ and $\lim_{\theta^* \to \bar{\theta}} y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right) < \lim_{\theta^* \to \bar{\theta}} y_R(\theta^*)$. Since $y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right)$ is continuous in θ^*, it follows that a value of θ^* satisfying $y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right) = y_R(\theta^*)$ exists.

Next, we show the sender can do no better than to send the message $\sigma = \emptyset$ for all $\theta < \theta^*$ and $\sigma = \theta$ for all $\theta \geq \theta^*$. For $\theta \geq \theta^*$, notice that $y_S(\theta) > y_R(\theta) \geq y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right)$; therefore disclosure is preferred to non-disclosure by an informed sender in this state.

For $\theta < \theta^*$, when not disclosing, a sender earns $U_S(|y_R(\theta^*) - y_S(\theta)|)$ and when disclosing, a sender earns $U_S(|y_R(\theta) - y_S(\theta)|)$. We claim that for all $\theta < \theta^*$, $|y_R(\theta) - y_S(\theta)| > |y_R(\theta^*) - y_S(\theta)|$.

Case 1: $y_R(\theta^*) < y_S(\theta)$. Then $|y_R(\theta) - y_S(\theta)| > |y_R(\theta^*) - y_S(\theta)|$ holds if and only if $y_S(\theta) - y_R(\theta) > y_S(\theta) - y_R(\theta^*)$ or $y_R(\theta) < y_R(\theta^*)$, and since $\theta < \theta^*$, this condition holds.

Case 2: $y_R(\theta^*) > y_S(\theta)$. Then $|y_R(\theta) - y_S(\theta)| > |y_R(\theta^*) - y_S(\theta)|$ holds if and only if $y_S(\theta) - y_R(\theta) > y_R(\theta^*) - y_S(\theta)$. To establish this inequality, observe that $y_S(\theta) - y_R(\theta) > y_S(\theta) - y_S(\theta^*) + y_R(\theta^*) - y_R(\theta) > y_S(\theta) - y_S(\theta^*) + 2(y_S(\theta^*) - y_S(\theta)) = y_S(\theta^*) - y_S(\theta) > y_R(\theta^*) - y_S(\theta)$ where the first inequality follows because $y_S(\theta) > y_R(\theta)$ while the second inequality follows from the gradual slope ordering property. This establishes that non-disclosure is preferred to disclosure, and completes the proof for the case where there are no agreement points.

Next, assume there is a single agreement point occurring in state θ'. Suppose that $y(\emptyset) < y_R(\theta')$ (the situation where $y(\emptyset) > y_R(\theta')$ follows an analogous line of proof). We will show that there is an interval $[\theta_1, \theta_2]$ where disclosure occurs. In the remaining states, an informed sender chooses not to disclose.

To construct $[\theta_1, \theta_2]$, we require (1) $\theta_1 < \theta' < \theta_2$, (2) $y_R(\theta_1) = y(\emptyset; D = [\theta_1, \theta_2])$, and (3) $|y_R(\theta_2) - y_S(\theta_2)| \leq |y(\emptyset; D = [\theta_1(\theta_2), \theta_2]) - y_S(\theta_2)|$ with equality if $\theta_2 < \bar{\theta}$. To see that such a construction is possible, fix $\theta_2 > \theta'$ and find a value $\theta_1(\theta_2)$ solving condition (2). Notice that, for θ_1 sufficiently small, $y_R(\theta_1) < y(\emptyset; D = [\theta_1, \theta_2])$ while for θ_1 close to θ', $y_R(\theta_1) > y(\emptyset; D = [\theta_1, \theta_2])$. Therefore a solution $\theta_1(\theta_2)$ exists. Similarly, by varying θ_2, one can show that there exists a value of $\theta_2 > \theta'$ satisfying condition (3). Therefore, such a construction is feasible.

When $\theta < \theta_1$, we claim the sender prefers non-disclosure. To establish this claim, we show $|y(\emptyset; D = [\theta_1, \theta_2]) - y_S(\theta)| \leq |y_R(\theta) - y_S(\theta)|$. The combination of $\theta < \theta'$ and the gradual slope ordering property implies that $y_S(\theta) > y_R(\theta)$. Thus, $|y_R(\theta) - y_S(\theta)| = y_S(\theta) - y_R(\theta) = y_S(\theta) - y_R(\theta_1) + y_R(\theta_1) - y_R(\theta) > y_S(\theta) - y_S(\theta_1) + y_R(\theta_1) - y_R(\theta) > y_S(\theta) - y_S(\theta_1) + 2(y_S(\theta_1) - y_S(\theta)) = y_S(\theta_1) - y_S(\theta) > y_R(\theta_1) - y_S(\theta) = y(\emptyset; D = [\theta_1, \theta_2]) - y_S(\theta)$, where the first and third inequalities follow because $y_S(\theta) > y_R(\theta)$ for $\theta < \theta'$ while the second inequality follows from the slope property.

Next, when $\theta > \theta_2$, we also claim the sender prefers non-disclosure. To establish this claim, we show $|y(\emptyset; D = [\theta_1, \theta_2]) - y_S(\theta)| \leq |y_R(\theta) - y_S(\theta)|$. In this case $y_S(\theta) < y_R(\theta)$ for all $\theta > \theta'$. Therefore, $|y_R(\theta) - y_S(\theta)| = y_R(\theta) - y_S(\theta) = y_R(\theta_2) - y_S(\theta_2) + (y_R(\theta) - y_R(\theta_2)) - (y_S(\theta) - y_S(\theta_2)) \geq y_R(\theta_2) - y_S(\theta_2) + 2(y_S(\theta) - y_S(\theta_2)) - (y_S(\theta) - y_S(\theta_2)) = y_R(\theta_2) - y_S(\theta_2) + y_S(\theta) - y_S(\theta_2) = y_S(\theta_2) - y(\emptyset; D = [\theta_1(\theta_2), \theta_2]) + y_S(\theta) - y_S(\theta_2) = y_S(\theta) - y(\emptyset; D = [\theta_1(\theta_2), \theta_2])$, where the weak inequality follows from the slope property and the penultimate equality follows from condition (3).

Finally, for $\theta \in (\theta_1, \theta_2)$, we claim the sender prefers to reveal. To establish this claim, we show $|y_R(\theta) - y_S(\theta)| \leq |y(\emptyset; D = [\theta_1, \theta_2]) - y_S(\theta)|$. We consider two cases: $\theta < \theta'$ and $\theta > \theta'$. When $\theta < \theta'$, $y_S(\theta) > y_R(\theta) > y_R(\theta_1) = y(\emptyset; [\theta_1, \theta_2])$. Hence, the required inequality holds. Alternatively, when $\theta > \theta'$, we know $y_R(\theta) - y_S(\theta) = y_R(\theta_2) - y_S(\theta_2) + (y_R(\theta) - y_R(\theta_2)) - (y_S(\theta) - y_S(\theta_2)) \leq y_R(\theta_2) - y_S(\theta_2) + 2(y_S(\theta) - y_S(\theta_2)) - (y_S(\theta) - y_S(\theta_2)) = y_R(\theta_2) - y_S(\theta_2) + y_S(\theta) - y_S(\theta_2) = y_S(\theta_2) - y(\emptyset; D = [\theta_1(\theta_2), \theta_2]) + y_S(\theta) - y_S(\theta_2) = y_S(\theta) - y(\emptyset; D = [\theta_1(\theta_2), \theta_2])$, where the inequality follows from the gradual slope ordering property and because $\theta < \theta_2$. ■

Proof of Proposition 7: Recall that an equilibrium consists of a value of θ^* that solves $y_R(\theta^*) = y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right)$. We will show that, at any such solution, it must be the case that $\frac{dy_R(\theta^*)}{d\theta^*} > \frac{d}{d\theta^*} y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right)$. Recall that $y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right)$ is the argument y which maximizes

$$\frac{pF(\theta^*)}{1-p+pF(\theta^*)} \frac{1}{F(\theta^*)} \int_{\underline{\theta}}^{\theta^*} U_R(y,\theta) f(\theta)\, d\theta + \frac{1-p}{1-p+pF(\theta^*)} \int_{\underline{\theta}}^{\bar{\theta}} U_R(y,\theta) f(\theta)\, d\theta$$

Our assumptions imply $y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right]\right)$ satisfies the first-order condition, $\Psi(y, \theta^*) \equiv$

$$\frac{p}{1-p+pF(\theta^*)} \int_{\underline{\theta}}^{\theta^*} \frac{\partial U_R(y,\theta)}{\partial y} f(\theta)\, d\theta + \frac{1-p}{1-p+pF(\theta^*)} \int_{\underline{\theta}}^{\bar{\theta}} \frac{\partial U_R(y,\theta)}{\partial y} f(\theta)\, d\theta = 0$$

where $\partial\Psi(y,\theta^*)/\partial y < 0$ since y is a maximum.

Using the Implicit Function Theorem and that $\frac{\partial U_R(y,\theta)}{\partial y}|_{\theta=\theta^*} = 0$, we have

$$\begin{aligned}
& \frac{dy\left(\emptyset; D=\left[\theta^*,\bar{\theta}\right]\right)}{d\theta^*} \\
= \; & -\frac{\partial\Psi(y,\theta^*)}{\partial\theta^*} \Big/ \frac{\partial\Psi(y,\theta^*)}{\partial y} \\
= \; & \frac{pf(\theta^*)}{1-p+pF(\theta^*)} \left[\frac{\frac{p}{1-p+pF(\theta^*)}\int_{\underline{\theta}}^{\theta^*}\frac{\partial U_R(y,\theta)}{\partial y}f(\theta)\,d\theta + \frac{(1-p)}{1-p+pF(\theta^*)}\int_{\underline{\theta}}^{\bar{\theta}}\frac{\partial U_R(y,\theta)}{\partial y}f(\theta)\,d\theta}{\partial\Psi(y,\theta^*)/\partial y}\right] \\
= \; & \frac{pf(\theta^*)}{1-p+pF(\theta^*)}\left[\frac{\Psi(y,\theta^*)}{\partial\Psi(y,\theta^*)/\partial y}\right] = 0
\end{aligned}$$

Since $\frac{dy_R(\theta^*)}{d\theta^*} > 0$, it then follows that $\frac{dy_R(\theta^*)}{d\theta^*} > \frac{d}{d\theta^*}y\left(\emptyset; D=\left[\theta^*,\bar{\theta}\right]\right)$ at any intersection point. Hence, there is a unique solution, θ^*. ■

Proof of Proposition 8: To see this, suppose, without loss of generality, that the sender's bliss line lies above the receiver's bliss line. In that case, the equilibrium equation, defining the cutoff θ^* where information revelation takes place, is given by $y\left(\emptyset; D=\left[\theta^*,\bar{\theta}\right]\right) = y_R(\theta^*)$. We claim that $\lim_{p\to1}\theta^*(p) = \underline{\theta}$. To see this, suppose to the contrary that $\limsup_{p\to1}\theta^*(p) = \theta' > \underline{\theta}$. Then $\limsup_{p\to1} y\left(\emptyset; D=\left[\theta^*(p),\bar{\theta}\right]\right) = \lim_{p\to1} y\left(\emptyset; D=\left[\theta',\bar{\theta}\right]\right) = y_R(\hat{\theta})$, where $\hat{\theta} < \theta'$. But this is a contradiction since, in equilibrium $\lim_{p\to1} y\left(\emptyset; D=\left[\theta',\bar{\theta}\right]\right) = y_R(\theta')$. Thus, the sequence of convex disclosure equilibria converges to full revelation in the limit. ■

Proof of Proposition 9: First, we rule out partial disclosure in any equilibrium. To see this, suppose to the contrary that, for some set of states having positive measure, the message $\sigma = I$, where I consists of an interval not including the entire state space, is sent in equilibrium. For states where this message is sent in equilibrium, the situation is identical to one in which the knowledge state is public. As a consequence, our previous arguments for that case imply that some positive measure of sender types can profitably deviate, contradicting the notion that I is sent in equilibrium.

Next, we consider the situation where disclosure regions are non-convex as a result of non-disclosure. Then there exist states θ' and θ'' where $\theta' < \theta''$ such that disclosure occurs in equilibrium in each of these states, but, for some $t \in (0,1)$, non-disclosure occurs in state $\theta''' = t\theta' + (1-t)\theta''$. Disclosure in states θ' and θ'' implies that $|y(\emptyset; D) - y_S(\theta')| \geq |y_R(\theta') - y_S(\theta')|$ and $|y(\emptyset; D) - y_S(\theta'')| \geq |y_R(\theta'') - y_S(\theta'')|$, where D denotes the set of states in which disclosure occurs. To show that non-disclosure will not occur in state θ''' we show that $|y(\emptyset; D) - y_S(\theta''')| < |y_R(\theta''') - y_S(\theta''')|$ cannot occur. We prove this for two separate cases:

Case 1: Suppose $y_S(\theta) > y_R(\theta)$ for all $\theta \in (\theta', \theta'')$. Then $|y(\emptyset; D) - y_S(\theta''')| < |y_R(\theta''') - y_S(\theta''')|$ can only hold if $y(\emptyset; D) > y_R(\theta''')$, and hence, $y(\emptyset; D) - y_S(\theta''') < y_S(\theta''') - y_R(\theta''')$, or equivalently, $2y_S(\theta''') - y_R(\theta''') > y(\emptyset; D)$. By the gradual slope ordering property, this implies $2y_S(\theta') - y_R(\theta') > y(\emptyset; D)$, which may be rewritten as $y_S(\theta') - y_R(\theta') > y(\emptyset; D) - y_S(\theta')$. But this contradicts our previous finding that $|y(\emptyset; D) - y_S(\theta')| \geq |y_R(\theta') - y_S(\theta')|$ (regardless of whether $y(\emptyset; D) > y_S(\theta')$) because if $y(\emptyset; D) < y_S(\theta')$, then we have $y_S(\theta') > y(\emptyset; D) > y_R(\theta')$. Thus, $|y(\emptyset; D) - y_S(\theta''')| < |y_R(\theta''') - y_S(\theta''')|$ cannot hold in this case.

Case 2: Suppose $y_S(\theta) < y_R(\theta)$ for all $\theta \in (\theta', \theta'')$. The proof establishing that $|y(\emptyset; D) - y_S(\theta''')| < |y_R(\theta''') - y_S(\theta''')|$ cannot hold is analogous to Case 1. ■

Proof of Proposition 10: The proof is by construction. Let θ' be an agreement point. Recall that there exists a convex disclosure equilibrium with disclosure interval $[\theta_1, \theta_2]$ satisfying $\theta_1 < \theta'$ and $\theta_2 > \theta'$. Consider an interval $I' = [\theta_1', \theta_2']$, where $\theta' \in [\theta_1', \theta_2'] \subset [\theta_1, \theta_2]$, and the sender sends message $\sigma = I'$ in equilibrium with the resulting action $y(I') = y_R(\theta')$. To see that such a construction is feasible, notice that, by continuity of the receiver's bliss line, there exists a continuum of pairs (θ_1', θ_2') that induce $y(I') = y_R(\theta')$. Moreover, these pairs can be made arbitrarily close to θ' and hence $[\theta_1', \theta_2'] \subset [\theta_1, \theta_2]$. Finally, since the non-disclosure region remained unchanged by this amendment, the equilibrium conditions for (θ_1, θ_2) are undisturbed.

It remains to show that this strategy is incentive compatible. By our previous arguments, we know that disclosure is preferred to non-disclosure in the region $[\theta_1, \theta_2]$ and vice-versa. Since the gradual slope ordering property implies the sender is conservative in the neighborhood of θ', it follows that there exists an interval sufficiently close to θ' where the sender prefers the action $y_R(\theta')$ to the disclosure action. Thus sending the message I' in the interval $[\theta_1', \theta_2']$ is preferred to full disclosure. ■

Proof of Lemma 1: We first show that there exists some θ_1 and θ_2 satisfying conditions (1) and (2). To see this, fix $\theta_1 < \theta_\emptyset$. Since $y_R(\theta_\emptyset) > y(\emptyset; ND = [\theta_1, \theta_\emptyset])$ and $y_R(\theta_1) < y(\emptyset; ND = [\theta_1, \theta_1])$, it follows from the Intermediate Value Theorem that there exists some $\theta_2 \in (\theta_1, \theta_\emptyset)$ satisfying $y(\emptyset; ND = [\theta_1, \theta_2]) = y_R(\theta_2)$. Let $\theta_2(\theta_1)$ denote this value of θ_2. For values of θ_1 close to $\theta_\emptyset$, we have $y_S(\theta_1) > y(\emptyset; ND = [\theta_1, \theta_2(\theta_1)]) > y_R(\theta_1)$, and thus $|y(\emptyset; ND = [\theta_1, \theta_2(\theta_1)]) - y_S(\theta_1)| < |y_S(\theta_1) - y_R(\theta_1)|$. But this implies that when $\lim_{\theta_1 \to \theta_L} |y(\emptyset; ND = [\theta_1, \theta_2]) - y_S(\theta_1)| > \lim_{\theta_1 \to \theta_L} |y_S(\theta_1) - y_R(\theta_1)|$, there exists some $\theta_1 \in (\theta_L, \theta_\emptyset)$ such that $|y(\emptyset; ND = [\theta_1, \theta_2(\theta_1)]) - y_S(\theta_1)| = |y_S(\theta_1) - y_R(\theta_1)|$. Thus, there exists some θ_1 and θ_2 satisfying conditions (1) and (2).

It remains to show that for such θ_1 and θ_2, it is incentive compatible for the sender not to disclose if and only if $\theta \in [\theta_1, \theta_2]$. First, consider $\theta > \theta_2$. If $y_S(\theta) > y_R(\theta)$, then $y_S(\theta) > y_R(\theta) > y(\emptyset; ND = [\theta_1, \theta_2])$. It follows immediately that disclosure is strictly preferred to non-disclosure. Conversely, if θ is such that $y_S(\theta) \le y_R(\theta)$, define θ'' to be the largest agreement point where $\theta'' < \theta$. (Since $y_S(\theta) > y_R(\theta)$ in the region $[\theta_L, \theta_H]$, then such an agreement point θ'' must exist for it to be the case that

$y_S(\theta) \leq y_R(\theta)$.) For θ such that $\theta > \theta''$, we have $y_R(\theta) - y_S(\theta) = y_R(\theta) - y_R(\theta'') - (y_S(\theta) - y_S(\theta'')) < y_S(\theta) - y_S(\theta'') < y_S(\theta) - y(\emptyset; ND = [\theta_1, \theta_2])$, where the first equality follows from $y_S(\theta'') = y_R(\theta'')$, the first inequality follows from the steep slope ordering property, and the second inequality follows because $y(\emptyset; ND = [\theta_1, \theta_2]) = y_R(\theta_2) < y_R(\theta'') = y_S(\theta'') \leq y_S(\theta)$. Therefore, the sender prefers disclosure in this region. Thus, for all $\theta > \theta_2$, disclosure is preferred.

Next, consider $\theta \in (\theta_1, \theta_2)$. We claim that $|y(\emptyset; ND = [\theta_1, \theta_2]) - y_S(\theta)| < |y_S(\theta) - y_R(\theta)|$. For θ close to θ_2, $y_S(\theta) \geq y(\emptyset; ND = [\theta_1, \theta_2]) > y_R(\theta)$ and hence non-disclosure is strictly preferred to disclosure. For θ close to θ_1, $y(\emptyset; ND = [\theta_1, \theta_2]) > y_S(\theta) > y_R(\theta)$. It follows that $y_S(\theta) - y_R(\theta) = y_S(\theta_1) - y_R(\theta_1) + \{(y_S(\theta) - y_S(\theta_1)) - (y_R(\theta) - y_R(\theta_1))\} > y_S(\theta_1) - y_R(\theta_1) - \{y_S(\theta) - y_S(\theta_1)\} \geq y(\emptyset; ND = [\theta_1, \theta_2]) - y_S(\theta_1) - \{y_S(\theta) - y_S(\theta_1)\} = y(\emptyset; ND = [\theta_1, \theta_2]) - y_S(\theta)$, where the inequality follows from the steep slope ordering property, and the next substitution follows from the equilibrium properties of θ_1 and θ_2. Since this exhausts the space of possibilities for $\theta \in (\theta_1, \theta_2)$, we have shown that the sender prefers non-disclosure to disclosure in this region.

For $\theta < \theta_1$, if $y_S(\theta) > y_R(\theta)$, a similar argument shows $y_S(\theta) - y_R(\theta) = y_S(\theta_1) - y_R(\theta_1) + \{(y_S(\theta) - y_S(\theta_1)) - (y_R(\theta) - y_R(\theta_1))\} < y_S(\theta_1) - y_R(\theta_1) - \{y_S(\theta) - y_S(\theta_1)\} = y(\emptyset; ND = [\theta_1, \theta_2]) - y_S(\theta_1) - \{y_S(\theta) - y_S(\theta_1)\} = y(\emptyset; ND = [\theta_1, \theta_2]) - y_S(\theta)$, so disclosure is preferred to non-disclosure in this region. Conversely, if $y_S(\theta) \leq y_R(\theta)$, then $y_S(\theta) \leq y_R(\theta) < y(\emptyset; ND = [\theta_1, \theta_2])$, and thus disclosure is preferred. ■

Proof of Proposition 11: When there are no agreement points and $y_S(\theta) > y_R(\theta)$, existence follows from Lemma 1. The case where $y_S(\theta) < y_R(\theta)$ is analogous.

When there is one agreement point, θ', where for $\theta > \theta'$, $y_S(\theta) > y_R(\theta)$ and

$y(\emptyset) > y_R(\theta')$, then, setting $\theta_L = \theta'$ and $\theta_H = \bar{\theta}$, we can invoke Lemma 1 to show existence. The case where $y(\emptyset) < y_R(\theta')$ is analogous. Conversely, when for $\theta < \theta'$, $y_S(\theta) > y_R(\theta)$ and $y(\emptyset) < y_R(\theta')$, then, setting $\theta_L = \underline{\theta}$ and $\theta_H = \theta'$, we can invoke Lemma 1. The case where $y(\emptyset) > y_R(\theta')$ is analogous.

Where there are multiple agreement points, define θ' and θ'' to be adjacent agreement points relative to $y(\emptyset)$ as set out in Lemma 1. When $y_S(\theta) > y_R(\theta)$ in (θ', θ''), the result follows immediately. An analogous argument shows existence when $y_S(\theta) < y_R(\theta)$ in (θ', θ''). ■

Proof of Proposition 12: Suppose to the contrary that non-disclosure regions are non-convex. Then there exist states θ' and θ'' where $\theta' < \theta''$ such that non-disclosure occurs in equilibrium in each of these states, but, for some $t \in (0,1)$, disclosure occurs in state $\theta''' = t\theta' + (1-t)\theta''$. Non-disclosure in state θ' and θ'' implies that $|y(\emptyset; ND) - y_S(\theta')| \leq |y_R(\theta') - y_S(\theta')|$ and $|y(\emptyset; ND) - y_S(\theta'')| \leq |y_R(\theta'') - y_S(\theta'')|$, where ND denotes the set of states in which there is non-disclosure. To show that non-disclosure occurs in θ''', we show that $|y(\emptyset; ND) - y_S(\theta''')| > |y_R(\theta''') - y_S(\theta''')|$ cannot occur. We prove this for three separate cases.

Case 1: Suppose that for all $\theta \in (\theta', \theta'')$, $y_S(\theta) > y_R(\theta)$. This implies that $y_R(\theta') \leq y(\emptyset; ND)$ and $y_R(\theta'') \leq y(\emptyset; ND)$, since if either of these inequalities were reversed, we would have $y_S(\theta) > y_R(\theta) > y(\emptyset; ND)$, implying the sender would prefer to disclose. It follows that $y_R(\theta''') < y(\emptyset; ND)$ since $y_R(\theta''') < y_R(\theta'')$.

Non-disclosure at θ' implies that $y(\emptyset; ND) - y_S(\theta') \leq y_S(\theta') - y_R(\theta')$ or, equivalently $2y_S(\theta') - y_R(\theta') \geq y(\emptyset; ND)$. And disclosure at θ''' implies $y(\emptyset; ND) - y_S(\theta''') \geq y_S(\theta''') - y_R(\theta''')$ or, equivalently $2y_S(\theta''') - y_R(\theta''') \leq y(\emptyset; ND)$. However, by the steep slope ordering property, $2y_S(\theta''') - y_R(\theta''') > 2y_S(\theta') - y_R(\theta') \geq y(\emptyset; ND)$. Thus, $|y(\emptyset; ND) - y_S(\theta''')| > |y_R(\theta''') - y_S(\theta''')|$ cannot hold.

Case 2: Suppose that, for all $\theta \in (\theta', \theta'')$, $y_S(\theta) < y_R(\theta)$. The proof of this case is analogous to that of Case 1.

Case 3: Suppose that some $\theta''' \in (\theta', \theta'')$ is an agreement point (possibly one of many). We will show that in equilibrium there cannot exist non-disclosure intervals $[\theta'_L, \theta''_L]$ and $[\theta'_H, \theta''_H]$ such that $\theta''_L < \theta''' < \theta'_H$. Suppose to the contrary that such intervals exist. There are four cases to consider.

Case 3(a): Suppose that, for all $\theta \in [\theta'_L, \theta''_L] \cup [\theta'_H, \theta''_H]$, $y_S(\theta) > y_R(\theta)$. Then it must be that $|y(\emptyset; ND) - y_S(\theta'_L)| \leq |y_R(\theta'_L) - y_S(\theta'_L)|$ and $|y(\emptyset; ND) - y_S(\theta''_H)| \leq |y_R(\theta''_H) - y_S(\theta''_H)|$.

When $y_S(\theta''') > y(\emptyset; ND)$, it follows that, since $y_S(\theta''') = y_R(\theta''') > y(\emptyset; ND)$, then for $\theta \in [\theta'_H, \theta''_H]$, we have that $y_S(\theta) > y_R(\theta) > y(\emptyset; ND)$ and hence disclosure is strictly preferred in the interval $[\theta'_H, \theta''_H]$, which is a contradiction.

Conversely, when $y_S(\theta''') \leq y(\emptyset; ND)$, then $y(\emptyset; ND) - y_S(\theta''') > y_S(\theta''') - y_R(\theta''')$ or, equivalently, $2y_S(\theta''') - y_R(\theta''') < y(\emptyset; ND)$. From the steep slope ordering property, it follows that $2y_S(\theta) - y_R(\theta) < y(\emptyset; ND)$ for $\theta \in [\theta'_L, \theta''_L]$. Hence $|y_R(\theta) - y_S(\theta)| < |y(\emptyset; ND) - y_S(\theta)|$ and disclosure is strictly preferred in states $\theta \in [\theta'_L, \theta''_L]$, which is a contradiction.

Case 3(b): Suppose that $y_S(\theta) < y_R(\theta)$ for all $\theta \in [\theta'_L, \theta''_L] \cup [\theta'_H, \theta''_H]$. A proof analogous to Case 3(a) establishes a contradiction.

Case 3(c): Suppose that $y_S(\theta) < y_R(\theta)$ for $\theta \in [\theta'_L, \theta''_L]$ while $y_S(\theta) > y_R(\theta)$ for $\theta \in [\theta'_H, \theta''_H]$. When $y_S(\theta''') > y(\emptyset; ND)$, it then follows immediately that, since $y_S(\theta''') = y_R(\theta''') > y(\emptyset; ND)$, then $y_S(\theta) > y_R(\theta) > y(\emptyset; ND)$ for $\theta \in [\theta'_H, \theta''_H]$, and therefore, disclosure is strictly preferred in the interval $[\theta'_H, \theta''_H]$, which is a contradiction. In contrast, when $y_S(\theta''') \leq y(\emptyset; ND)$, then, since $y_S(\theta''') = y_R(\theta''') \leq y(\emptyset; ND)$, it follows that $y_S(\theta) < y_R(\theta) < y(\emptyset; ND)$ for $\theta \in [\theta'_L, \theta''_L]$. Consequently, disclosure is strictly preferred in the interval $[\theta'_L, \theta''_L]$, which is a contradiction.

Case 3(d): Suppose that $y_S(\theta) > y_R(\theta)$ for $\theta \in [\theta'_L, \theta''_L]$ while $y_S(\theta) < y_R(\theta)$ for $\theta \in [\theta'_H, \theta''_H]$. The proof is analogous to the proof where in Case 3(c).

Since this exhausts all of the possibilities, the proof is complete. ■

Proof of Proposition 14: The proof follows from the following two lemmas.

Lemma 3 *If the knowledge state is private, the gradual slope ordering property holds, and there is a boundary equilibrium under S_2, then any boundary equilibrium under S_1 is identical.*

Proof. First, suppose $\left[\theta_1, \bar{\theta}\right]$ is the equilibrium disclosure interval under S_2. Such an equilibrium satisfies $y_R(\theta_1) = y\left(\emptyset; D = \left[\theta_1, \bar{\theta}\right]\right)$ and $y_R\left(\bar{\theta}\right) - y_{S_2}\left(\bar{\theta}\right) \leq y_{S_2}\left(\bar{\theta}\right) - y\left(\emptyset; D = \left[\theta_1, \bar{\theta}\right]\right)$. Since S_1 is more aligned than S_2, $y_R\left(\bar{\theta}\right) - y_{S_1}\left(\bar{\theta}\right) < y_{S_1}\left(\bar{\theta}\right) - y\left(\emptyset; D = \left[\theta_1, \bar{\theta}\right]\right)$ when S_1 discloses on the same interval. The equilibrium equation $y_R(\theta_1) = y\left(\emptyset; D = \left[\theta_1, \bar{\theta}\right]\right)$ is independent of the sender preferences, and so also holds under S_1. Therefore, if $\left[\theta_1, \bar{\theta}\right]$ is a boundary equilibrium under S_2, it is also a boundary equilibrium under S_1. We know from the proof of Proposition 7 that the solution to the equation $y_R(\theta_1) = y\left(\emptyset; D = \left[\theta_1, \bar{\theta}\right]\right)$ is unique, so $\left[\theta_1, \bar{\theta}\right]$ is the unique boundary equilibrium under S_1. ■

Next, we show that the receiver strictly prefers a boundary equilibrium under S_2 to any interior equilibrium.

Lemma 4 *If the knowledge state is private, the gradual slope ordering property holds, and a boundary equilibrium exists, then it is the most informative equilibrium.*

Proof. A boundary equilibrium consists of disclosure in the interval $\left[\theta_1, \bar{\theta}\right]$ where $y_R(\theta_1) = y\left(\emptyset; D = [\theta_1, \bar{\theta}]\right)$. Suppose, contrary to the lemma, that an interior equilibrium $[\theta'_1, \theta'_2]$ exists and $[\theta'_1, \theta'_2] \nsubseteq [\theta_1, \bar{\theta}]$. First, by construction, $\theta'_2 < \bar{\theta}$. If $\theta'_1 \leq \theta_1$, then this implies that $y(\emptyset; D = [\theta_1', \theta_2']) > y\left(\emptyset; D = [\theta_1, \bar{\theta}]\right)$ and $y_R(\theta_1') \leq y_R(\theta_1)$. But, in equilibrium $y_R(\theta_1') = y(\emptyset; D = [\theta_1', \theta_2'])$, so this is a contradiction. Hence, it must be the case that, in any interior equilibrium, $\theta_1' > \theta_1$. But this implies that $[\theta_1', \theta_2'] \subseteq [\theta_1, \bar{\theta}]$, which is a contradiction. ■

This claim completes the proof. ■

Proof of Proposition 15: We will establish the result when $y(\emptyset) > y_R(\theta')$ and $y_{S_1}(\theta) = y_{S_2}(\theta)$ for $\theta \geq \theta'$. The opposite case is analogous. We claim that, for any pair of senders S_1 and S_2 satisfying the above preferences, every equilibrium under S_1 is an equilibrium under S_2. To see this, fix a convex disclosure equilibrium under S_1, $[\theta_1^*, \theta_2^*]$, and recall that such an equilibrium solves $y_R(\theta_1^*) = y(\emptyset; D = [\theta_1^*, \theta_2^*])$ and $y_R(\theta_2^*) - y_{Sj}(\theta_2^*) = y_{Sj}(\theta_2^*) - y(\emptyset; D = [\theta_1^*, \theta_2^*])$. For a fixed θ_2^*, the solution θ_1^* given by the first equation is independent of sender preferences. And given θ_1^*, any solution $\theta_2^* > \theta'$ found for S_1 is also a solution for S_2, because sender preferences are identical in this region. Finally, the proof of Proposition 6 shows that all convex disclosure equilibria have the property that $\theta_2^* > \theta'$. ■

Proof of Proposition 16: Suppose $y_S(\theta) > y_R(\theta)$ for all θ; the opposite case is analogous. When consulting S_2, the resulting equilibrium is characterized by θ^* solving $y_R(\theta^*) = y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right], p_2\right)$, where $y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right], p_2\right)$ denotes the action taken following the null message when disclosure occurs on the interval $\left[\theta^*, \bar{\theta}\right]$ and the sender's expertise is p_2. If, instead, the receiver consulted S_1, then $y_R(\theta^*) > y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right], p_1\right)$ since $y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right], p_1\right) < y\left(\emptyset; D = \left[\theta^*, \bar{\theta}\right], p_2\right)$, where this inequality follows because the receiver now considers the null message as more likely to emanate from an informed sender. Moreover, since $y_R(\underline{\theta}) < y\left(\emptyset; D = \left[\underline{\theta}, \bar{\theta}\right], p_1\right)$, it follows that, for some $\theta_1^* \in (\underline{\theta}, \theta^*)$, $y_R(\theta_1^*) = y\left(\emptyset; D = \left[\theta_1^*, \bar{\theta}\right], p_1\right)$. This implies strictly more disclosure occurs when the receiver consults S_1 than S_2 , so the receiver strictly prefers to consult S_1. ■

Proof of Proposition 17: We will construct a more informative equilibrium under S_1. Let $[\theta_1, \theta_2]$ be an interior equilibrium under S_2. Then θ_1 and θ_2 satisfy $y_R(\theta_1) = y(\emptyset; D = [\theta_1, \theta_2])$ and $y_R(\theta_2) - y_{S_2}(\theta_2) = y_{S_2}(\theta_2) - y(\emptyset; D = [\theta_1, \theta_2])$. Since S_1 is more aligned than S_2, it follows that $y_R(\theta_2) - y_{S_1}(\theta_2) < y_{S_1}(\theta_2) - y(\emptyset; D = [\theta_1, \theta_2])$. Now, let θ_2 vary and let $\theta_1(\theta_2)$ solve $y_R(\theta_1) = y(\emptyset; D = [\theta_1, \theta_2])$ implicitly. By regularity, we know that $\theta_1(\theta_2)$ is strictly decreasing in θ_2. There are then two cases to consider:

Case 1: Suppose $y_R\left(\bar{\theta}\right) - y_{S_1}\left(\bar{\theta}\right) \leq y_{S_1}\left(\bar{\theta}\right) - y\left(\emptyset; D = \left[\theta_1\left(\bar{\theta}\right), \bar{\theta}\right]\right)$. Then, the boundary equilibrium $\left[\theta_1\left(\bar{\theta}\right), \bar{\theta}\right]$ is an equilibrium under S_1 and is strictly more informative than the equilibrium under S_2.

Case 2: Suppose $y_R\left(\bar{\theta}\right) - y_{S_1}\left(\bar{\theta}\right) > y_{S_1}\left(\bar{\theta}\right) - y\left(\emptyset; D = \left[\theta_1\left(\bar{\theta}\right), \bar{\theta}\right]\right)$. By the Intermediate Value Theorem, there exists $\theta_2' \in \left(\theta_2, \bar{\theta}\right)$ such that $y_R(\theta_2') - y_{S_1}(\theta_2') = y_{S_1}(\theta_2') - y(\emptyset; D = [\theta_1(\theta_2'), \theta_2'])$. Hence, $[\theta_1(\theta_2'), \theta_2']$ is an equilibrium. Further, since $\theta_2' > \theta_2$ and $\theta_1(\theta_2') < \theta_1$ by regularity, this equilibrium is more informative than that under S_2. ■

Proof of Proposition 18: The receiver's preferences solely determine the convex non-disclosure region in a boundary equilibrium. Suppose $y_S(\theta) > y_R(\theta)$ for θ in the neighborhood of $\underline{\theta}$; the opposite case is analogous. Then non-disclosure occurs in the region $[\underline{\theta}, \theta_2]$, where θ_2 satisfies $y_R(\theta_2) = y(\emptyset; ND = [\underline{\theta}, \theta_2])$. Since this is a boundary equilibrium under S_1, it follows that $y(\emptyset; ND = [\underline{\theta}, \theta_2]) - y_{S_1}(\underline{\theta}) \leq y_{S_1}(\underline{\theta}) - y_R(\underline{\theta})$. Since S_2 is less aligned than S_1, $y(\emptyset; ND = [\underline{\theta}, \theta_2]) - y_{S_2}(\underline{\theta}) < y_{S_2}(\underline{\theta}) - y_R(\underline{\theta})$. Hence $[\underline{\theta}, \theta_2]$ is a boundary equilibrium under S_2. ■

Proof of Lemma 2: Let $U_R([\theta_1, \theta_2])$ denote the receiver's expected utility when the non-disclosure interval is $[\theta_1, \theta_2]$. Since $y(\emptyset; ND = [\theta_1, \theta_2]) = y_R(\theta_2)$, it follows that when $\theta_2 = \theta_2(\theta_1)$, we have $\partial U_R([\theta_1, \theta_2]) / \partial\theta_2 = 0$. However, increasing θ_1 makes the receiver better off because of increased disclosure. Since $\partial U_R([\theta_1, \theta_2]) / \partial\theta_1 > 0$ while $\partial U_R([\theta_1, \theta_2]) / \partial\theta_2 = 0$ the result follows. ■

Proof of Proposition 19: Suppose without loss of generality that $y_{S_i}(\theta_\emptyset) > y_R(\theta_\emptyset)$ for $i \in \{1, 2\}$. Since S_1 is more aligned than S_2, we have $y(\emptyset; ND = [\theta_1, \theta_2]) - y_{S_1}(\theta_1) > y_{S_1}(\theta_1) - y_R(\theta_1)$ at the original equilibrium under S_2. Also note that $\theta_2(\theta_\emptyset) = \theta_\emptyset$ since $y(\emptyset; ND = [\theta_\emptyset, \theta_\emptyset]) = y_R(\theta_\emptyset)$. Hence $y(\emptyset; ND = [\theta_\emptyset, \theta_2(\theta_\emptyset)]) - y_{S_1}(\theta_\emptyset) < 0 < y_{S_1}(\theta_\emptyset) - y_R(\theta_\emptyset)$. Thus there exists some θ_1' in $(\theta_1, \theta_\emptyset)$ for which $y(\emptyset; ND = [\theta_1', \theta_2(\theta_1')]) - y_{S_1}(\theta_1') = y_{S_1}(\theta_1') - y_R(\theta_1')$, meaning $[\theta_1', \theta_2(\theta_1')]$ constitutes a non-disclosure equilibrium under S_1. The result then follows from Lemma 2. ■