Australian School of Business Research Paper No. 2013 ECON 19

Optimal Information Disclosure: Quantity vs. Quality

Anton Kolotilin

Optimal Information Disclosure: Quantity vs. Quality*

Anton Kolotilin†

This Version: July, 2013

Abstract

A sender chooses ex ante how her information will be disclosed to a privately informed receiver who then takes one of two actions. The sender wishes to maximize the probability that the receiver takes the desired action. The sender faces an ex ante quantity-quality tradeoff: sending positive messages more often (in terms of the sender's information) makes it less likely that the receiver will take the desired action (in terms of the receiver's information). Interestingly, the sender's and receiver's welfare is not monotonic in the precision of the receiver's private information: the sender may find it easier to influence a more informed receiver, and the receiver may suffer from having more precise private information. Necessary and sufficient conditions are derived for full and no information revelation to be optimal.

Key words: information disclosure, persuasion, informed decision maker, two-way communication

JEL Codes: C72, D81, D82, D83

1 Introduction

Decision makers often rely on information obtained from interested parties. Most of the literature on communication assumes that decision makers do not have private information. But

*This paper is a revised version of Chapter 2 of my dissertation at MIT. I thank Robert Gibbons and Muhamet Yildiz for their invaluable guidance and advice. I also thank Ricardo Alonso, Sandeep Baliga, Abhijit Banerjee, Gabriel Carroll, Denis Chetverikov, Glenn Ellison, Richard Holden, Hongyi Li, Uliana Loginova, Niko Matouschek, Parag Pathak, Michael Powell, Andrea Prat, Juuso Toikka, Alexander Wolitzky, Juan Xandri, Luis Zermeño, and the participants at CUHK, HKU, HKUST, NES, MIT, UNSW, and UQ seminars for helpful comments and suggestions.

†UNSW, School of Economics. Email: a.kolotilin@unsw.edu.au.

generically, everyone has private information. The main goal of this paper is to understand economic aspects of optimal information disclosure from a sender to a privately informed receiver.

In my model, the receiver decides whether to act or not to act. The sender's utility depends only on the action taken by the receiver, and she prefers the receiver to act. The receiver's utility depends both on his action and on information. The receiver takes an action that maximizes his expected utility, given his private information and information disclosed by the sender. Before obtaining her private information, the sender can commit to how her private information will be disclosed to the receiver. Formally, the sender can choose any (stochastic) mapping from her information to messages, which I call a *mechanism*. The sender chooses the mechanism that maximizes the ex ante probability that the receiver will act. I impose a single-crossing assumption requiring that receiver's types can be ordered according to their willingness to act.

For example, consider a school that chooses a disclosure policy for a student in order to persuade a potential employer to hire him. The school has a lot of freedom in choosing which part of available information about the student will appear on his transcript. Moreover, the school commits to its disclosure policy before it learns anything about the student. The employer observes the student's transcript but also obtains private information, for example, from conducting an employment interview with the student and competing candidates. In addition, the school uses the same disclosure policy for all students, who apply to different employers. This also contributes to the receiver's private information in terms of my model.

Since the receiver has private information, he acts or does not act depending not only on a message received from the sender but also on his private information. Thus, from the sender's perspective, each message generates a probability distribution over receiver's actions. Therefore, when the sender chooses a mechanism, she faces an important quantity-quality tradeoff of messages that she will later send: sending positive messages more often (in terms of the sender's private information) makes it less likely that the receiver will act upon receiving them (in terms of the receiver's private information). The optimal mechanism balances these two conflicting objectives. For example, when the school chooses lower standards for getting good grades, more students get good-looking transcripts, but employers rationally account for this and each student with a good-looking transcript will find it harder to get a job. This ex ante tradeoff does not appear in cheap talk and verifiable message games where the sender chooses a report at the interim stage when she already has her private information.

Interestingly, under the optimal mechanism, the sender's and receiver's expected utilities are not monotonic in the precision of the receiver's private information. First, as the receiver becomes more informed, his expected utility may decrease despite the fact that he is the only player who takes an action that directly affects his utility. This happens because the optimal mechanism depends on the structure of the receiver's private information, and the sender may prefer to disclose significantly less information if the receiver's information is more precise.[1] Second, it may be easier for the sender to influence a more informed receiver. This happens because the sender may optimally choose to target only the receiver with favorable private information. In this case, it becomes easier for the sender to persuade the receiver with more precise favorable information, so the sender may be able to persuade the receiver with a higher total probability.

The sender's problem of finding an optimal mechanism reduces to a linear program. Using duality theory, I show how to obtain primitive necessary and sufficient conditions for a candidate mechanism to be optimal. This is the main technical contribution of the paper, which can be applied to other models of information disclosure because the sender's expected utility is always linear in probabilities that constitute a mechanism. In reality, schools choose various disclosure policies and duality theory allows us to find primitive conditions on the environment that justify each choice. At the one extreme, schools report all grades and class rank on transcripts. The full revelation mechanism is optimal if and only if the sender prefers to reveal any two of her types than to pool them. At the other extreme, schools release no transcripts. The no revelation mechanism is optimal if and only if the sender prefers to pool any three of her types than to pool two of them and reveal the third one. Under further assumptions, I show that the amount of information that is optimally disclosed is determined by the convexity properties of the distribution of the receiver's type and by the expectation of the sender's type.

In the benchmark model, the receiver is not allowed to communicate with the sender. This assumption fits many real-life examples. In particular, the school gives the same transcripts to students regardless of where they apply for a job and before they get interviewed by employers. However, this assumption is not without loss of generality because the sender can potentially increase the probability that the receiver acts by conditioning a mechanism on receiver's reports. I provide sufficient conditions under which two-way communication does not help the sender and give an example in which it does.

[1] Continuing the school-employer application, Arvey and Campion (1982) summarize research on employment interviews and report low reliability for interview-based assessments, which may actually be beneficial for employers because it motivates schools to design more informative disclosure policies as shown in my model.

The most related literature is the one in which the sender can commit to an information disclosure mechanism. Kamenica and Gentzkow (2011) study a much more general model but focus on the case of an uninformed receiver. They also show that some results generalize to the case of a privately informed receiver. In my companion paper Kolotilin (2013), I derive monotone comparative statics results with respect to the probability distribution of information for the case of an uninformed receiver. In contrast, in this paper I focus on the case in which the receiver does have private information, where both the results and analytical techniques are very different. Similar to this paper, Rayo and Segal (2010) assume that the receiver has a binary action choice, but they allow the sender's utility to depend not only on the action but also on information. To make the analysis tractable, they assume that the receiver's type is uniformly distributed. This assumption would make my model trivial in that the sender's expected utility would be the same under any mechanism, as follows from part 1 of Theorem 1 below. Ostrovsky and Schwarz (2010) study information disclosure in matching markets with private information. The main conceptual difference is that they study equilibrium rather than optimal information disclosure.

A few papers study cheap talk with a privately informed receiver. In the cheap talk version of my model, the unique equilibrium outcome involves no information revelation because the sender's utility depends only on the receiver's action and the information structure satisfies the single-crossing assumption. If either of these two assumptions fails, a fully revealing equilibrium may exist (Seidmann (1990) and Watson (1996)). Chen (2009), de Barreda (2012), and Lai (2013) study cheap talk with an informed receiver under the standard Crawford and Sobel (1982) assumptions. They all show that the receiver's expected utility may be not monotonic in the precision of his private information. The mechanics of these results, however, is different from that of my non-monotone comparative statics results. In stark contrast to optimal information disclosure games (Kolotilin (2013)), in cheap talk games with an uninformed receiver, the sender's and receiver's expected utilities are non-monotonic in the precision of the sender's information (Green and Stokey (2007) and Ivanov (2010)) and in the precision of public information (Chen (2012)).

The rest of the paper is organized as follows. Section 2 develops a general model. Section 3 presents two examples that illustrate the quantity-quality tradeoff of the sender and non-monotone comparative statics. Section 4 analyzes the model under a fairly general information structure of the sender and receiver. This section partially characterizes the optimal mechanism and derives primitive necessary and sufficient conditions for optimality of the full revelation and no revelation mechanisms. Section 5 extends the model to allow two-way information disclosure between the sender and receiver. Section 6 concludes. The appendix contains formal proofs.

2 Model

Consider a communication game between a female sender and a male receiver. The receiver takes a binary action: *to act* ($a = 0$) or *not to act* ($a = 0$). The sender's utility depends only on a, but the receiver's utility depends both on a and on (r, s) where components r and s denote the receiver's and sender's types, respectively. That is, the sender's utility is a, and the receiver's utility is $au(r, s)$ where u is a continuously differentiable function. Before s is realized, the sender can commit to a mechanism that sends a message m to the receiver as a (stochastic) function of her type s; specifically, the sender chooses the conditional distribution $\phi(m|s)$ of m given s. With a slight abuse of notation, the joint distribution of (m, s) is denoted by $\phi(m, s)$.

Assume that the set of messages is the continuum, the set R of receiver's types is $[\underline{r}, \overline{r}]$, and the set S of sender's types is $[\underline{s}, \overline{s}]$. The information (r, s) has some joint distribution. Unless stated otherwise, assume that for this distribution, the marginal distribution $F(s)$ of s and the conditional distribution $G(r|s)$ of r given s admit strictly positive continuously differentiable densities $f(s)$ and $g(r|s)$.

The timing of the communication game is as follows:

1. The sender publicly chooses a mechanism $\phi(m|s)$.

2. A triple (m, r, s) is drawn according to ϕ, F, and G.

3. The receiver observes (m, r) and takes an action a.

4. Utilities of the sender and receiver are realized.

The solution concept used is Perfect Bayesian Equilibrium (PBE). I view PBEs as identical if they have the same equilibrium mapping from information (s, r) to the receiver's action a. At the third stage, the receiver forms a belief about s and acts if and only if the conditional expectation $\mathbb{E}_\phi[u(r, s)|m, r]$ of u given (m, r) is at least 0. At the first stage, the sender chooses an *optimal mechanism* that maximizes her expected utility, the probability that the receiver acts. The main assumption, formally imposed later, is the single-crossing assumption: each message m induces types $r \geq r^*(m)$ to act, for some function r^*.

Hereafter, use the following definitions and conventions. All notions are in the weak sense, unless stated otherwise. For example, increasing means non-decreasing and higher means not lower. Two mechanisms are *equivalent* if they result in the same probability that the receiver acts. One mechanism *dominates* another mechanism if the former results in a higher probability that the receiver acts than the latter. *The full revelation mechanism* (denoted by ϕ_{full}) is a mechanism that sends a different message for each s. *The no revelation mechanism* (denoted by ϕ_{no}) is a mechanism that sends the same message regardless of s. The *survival function* $\overline{H}$ of a random variable with distribution H is defined as $\overline{H} \equiv 1 - H$.

3 Examples

In this section, I discuss two complimentary examples. In the first example, the sender's and receiver's types are binary and the receiver's type is a noisy signal about the state, known by the sender. In the second example, the sender's and receiver's types are continuous and the receiver's type is independent of the sender's type. For these examples, I derive the optimal mechanism and illustrate the sender's quantity-quality tradeoff. Further, I show that the sender's and receiver's expected utilities are non-monotonic in information. Finally, I discuss what determines how much of information is optimally disclosed.

3.1 Binary Example

In this example, the sender is perfectly informed, but the receiver is partially informed. That is, the sender knows the receiver's utility from acting, but the receiver only gets a signal about his utility. Specifically, the receiver's utility from acting is equal to the sender's type s that takes two values: $s = 1$ with probability $1/5$ and $s = -1$ with probability $4/5$. The receiver's type (equivalently signal) r also takes two values $r = 1$ and $r = -1$ according to the following conditional probabilities:

$$\Pr(r = 1|s = 1) = \Pr(r = -1|s = -1) = p.$$

The parameter p captures the precision of the receiver's private signal. In the school-employer application, p may correspond to the quality of an employment interview. Without loss of generality, assume that $p \in [1/2, 1]$. For a given mechanism, the receiver $r = 1$ assigns a higher probability that s is 1, than the receiver $r = -1$. Moreover, the difference in their assessments of the probability that s is 1 increases with p. Thus, p can be alternatively viewed as the measure of *polarization* between the *optimistic receiver* ($r = 1$) and the *pessimistic receiver* ($r = -1$).

A message m under a mechanism ϕ generates a posterior probability $\Pr_\phi(s|m)$ of s given m for each value s. The probability $\Pr_\phi(s = 1|m, r)$ that s is 1 given m and r can be calculated using Bayes' rule. The receiver acts if $\Pr_\phi(s = 1|m, r) \geq 1/2$. It is straightforward to calculate that upon receiving m, the optimistic receiver acts if $\Pr_\phi(s = 1|m) \geq 1 - p$, and the pessimistic receiver acts if $\Pr_\phi(s = 1|m) \geq p$. Clearly, if m induces the pessimistic receiver to act, it also induces the optimistic receiver to act. Thus, by the revelation principle, we can restrict attention to mechanisms with three messages: (i) m_0 that induces the receiver not to act regardless of his signal ($\Pr_\phi(s = 1|m_0) \in [0, 1 - p)$), (ii) m_1 that induces only the optimistic receiver to act ($\Pr_\phi(s = 1|m_1) \in [1 - p, p)$), and (iii) m_2 that induces the receiver to act regardless of his signal ($\Pr_\phi(s = 1|m_2) \in [p, 1]$). Because the sender's expected utility is equal to the probability that the receiver acts, she would strictly prefer to send m_2 over m_1 and m_1 over m_0 if there were no constraints on how often she can send various messages.

The prior distribution of s, however, imposes a constraint on how often the sender can send various messages:

$$\sum_{i=0}^{2} \Pr{}_\phi(s = 1|m_i) \Pr{}_\phi(m_i) = \Pr(s = 1) = \frac{1}{5}, \tag{1}$$

where $\Pr{}_\phi(m_i)$ denotes the probability that m_i is sent under a mechanism ϕ. Constraint (1) implies that to maximize the probability of the messages m_2 and m_1, the sender should choose a mechanism that satisfies: $\Pr_\phi(s = 1|m_0) = 0$, $\Pr_\phi(s = 1|m_1) = 1-p$, and $\Pr{}_\phi(s = 1|m_2) = p$.[2] That is, m_0 gives the most possible evidence against acting; m_1 gives the minimal possible evidence to make the optimistic receiver act; and m_2 gives the minimal possible evidence to make the pessimistic receiver act. These observations imply that the sender's expected utility simplifies to:[3]

$$2p(1-p)\Pr(m_1) + \Pr(m_2), \tag{2}$$

and constraint (1) simplifies to:

$$(1-p)\Pr(m_1) + p\Pr(m_2) = \frac{1}{5}. \tag{3}$$

[2] Formally, the optimal mechanism is derived in the supplemental appendix for a setting that nests this example.

[3] Equation (2) is obtained using the fact that m_2 induces the receiver to act with probability 1, m_1 induces the receiver to act with probability $p\Pr_\phi(s = 1|m_1) + (1-p)\Pr_\phi(s = -1|m_1)$, and m_0 induces the receiver to act with probability 0.

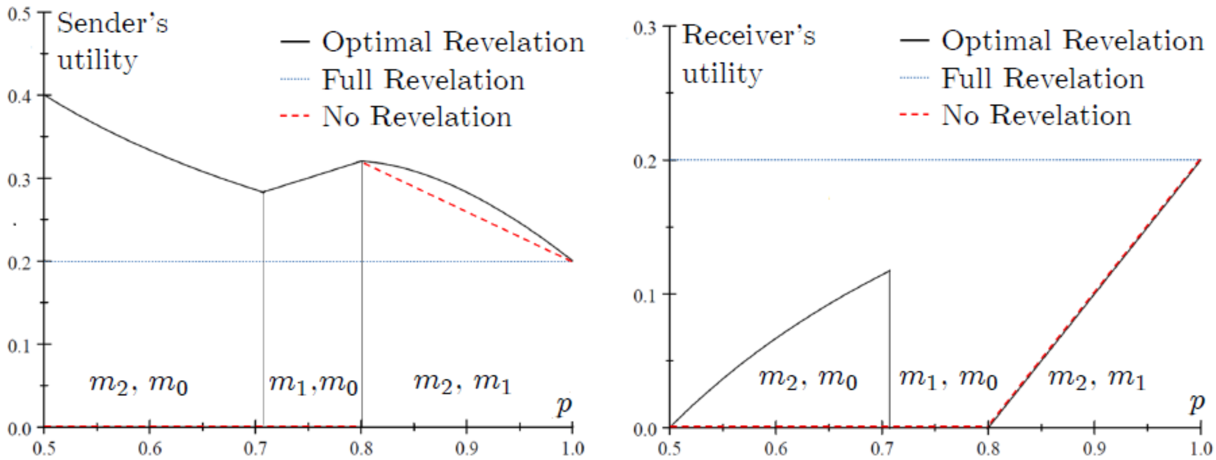

Figure 1: The sender's and receiver's expected utilities as a function of the precision of the receiver's private information.

The sender's problem of finding the optimal mechanism can be viewed as a problem of maximizing the linear utility function (2) over probabilities Pr (m_0), Pr (m_1), and Pr (m_2) subject to the budget constraint (3). That is, the marginal utilities of the messages m_0, m_1, and m_2 are 0, $2p(1-p)$, and 1; and the prices of these messages are 0, $1-p$, and p. Thus, the sender faces a quantity-quality tradeoff: to send m_1 with a high probability and persuade only the optimistic receiver or to send m_2 with a small probability and persuade both the pessimistic and optimistic receivers. This tradeoff is resolved by a choice of a mechanism that sends messages with the highest *marginal utility-price ratio*. Before discussing the optimal mechanism in a greater detail, I highlight non-monotone comparative statics.

Figure 1 shows the sender's and receiver's expected utilities under the optimal mechanism. Naive intuition may suggest that (i) the sender's expected utility should decrease with p because it is harder to influence a better informed receiver and (ii) the receiver's expected utility should increase with p because a better informed receiver takes a more appropriate action. This naive intuition, however, does not take into account that the optimal mechanism changes with p, and the sender may choose to disclose significantly less information if the receiver is more informed. This effect may overturn the results. In fact, the sender's expected utility strictly increases with p for $p \in \left(1/\sqrt{2}, 4/5\right)$, and the receiver's expected utility jumps down to zero as p exceeds $1/\sqrt{2}$.[4] Thus, a more informative employment interview may help

[4]Consistent with the naive intuition, the sender's expected utility decreases and the receiver's expected utility increases with the precision of the receiver's information if this precision is either low or high. Indeed, the sender is best off and the receiver is worst off when the receiver is uniformed. Moreover, the sender is

the school in influencing the employer's decision and may hurt the employer in making him hire worse students on average.

I stress that these non-monotone comparative statics results with respect to the precision of information arise only when the receiver is privately informed. If the receiver's signal was public, then the sender's and receiver's expected utilities would be monotonic both in the precision of the sender's private information and in the precision of public information (Kolotilin (2013)).

Figure 1 also sheds light on the extent to which information disclosure can affect the receiver's action and on the informativeness of the optimal mechanism. As the left panel shows, for a wide range of p, the probability that the receiver acts is considerably higher under the optimal mechanism than under the two benchmark mechanisms: the full revelation and no revelation mechanisms. As the right panel shows, from the receiver's perspective, the optimal mechanism is maximally uninformative if $p = \frac{1}{2}$ or $p \in \left[1/\sqrt{2}, 1\right]$, and its informativeness gradually increases with p for $p \in \left(1/2, 1/\sqrt{2}\right)$. I now explain the three forms that the optimal mechanism can take as p increases from $1/2$ to 1.

First, if the receiver's signal is imprecise in that p is close to $1/2$, then it is almost as cheap to persuade the pessimistic receiver to act as it is to persuade the optimistic receiver to act, because the prices p and $1-p$ are close. Thus, the sender prefers to target the pessimistic receiver, so the optimal mechanism sends the messages m_2 and m_0. As p increases, it becomes harder to persuade the pessimistic receiver to act and, thus, sending m_2 becomes more expensive. As a result, the sender's expected utility decreases with p. Since the optimal mechanism gives no rent to the pessimistic receiver, the optimistic receiver gets a strictly positive rent, which increases with p.

Second, as p exceeds $1/\sqrt{2}$ (but falls behind $4/5$), the polarization between the optimistic and pessimistic receivers becomes so high that it becomes much more expensive to persuade the pessimistic receiver to act than to persuade the optimistic receiver to act. Thus, the sender prefers to target the optimistic receiver, so the optimal mechanism sends the messages m_1 and m_0. In other words, the sender switches from the more expensive and more persuasive message m_2 to the less expensive and less persuasive message m_1. As p increases, the price $1-p$ of sending m_1 decreases and it becomes easier to persuade the optimistic receiver to act. As a result, the sender's expected utility increases with p. The receiver's expected utility jumps down to 0 as p exceeds $1/\sqrt{2}$, and it stays at 0 because the optimal mechanism makes the receiver indifferent to act whenever he acts.

worst off and the receiver is best off when the receiver is perfectly informed.

Third, as p exceeds $4/5$, the receiver's signal becomes so precise that the sender can persuade the optimistic receiver to act by disclosing no information. Thus, the sender prefers to target the optimistic receiver with certainty and the pessimistic receiver with some probability, so the optimal mechanism sends m_1 and m_2. As p increases further, the sender can persuade the pessimistic receiver to act more often, so the optimal mechanism sends m_2 with a higher probability. But the probability of the receiver being optimistic decreases, so m_1 induces the receiver to act with a lower probability. In this example, the latter effect dominates the former, so the sender's expected utility decreases with p. The receiver's expected utility increases with p because the optimal mechanism gives no rent to the receiver, so a better informed receiver takes a more appropriate action.

The sender's quantity-quality tradeoff illustrated here carries on to a general version of the model. If the sender's signal is binary, this tradeoff is resolved by the choice of messages with the highest marginal utility-price ratio (Section 4 and the supplemental appendix), otherwise the tradeoff becomes more intricate because the budget constraint becomes multidimensional (Sections 3.2 and 4).

3.2 Continuous Example

In this example, the receiver's utility is additive in sender's and receiver's types that are independent of each other. More formally, $u(r, s) = s - r$ where s and r are independently distributed with distributions F and G. The supports are such that the receiver $\underline{r}$ always acts ($\underline{r} < \underline{s}$) and the receiver $\overline{r}$ never acts ($\overline{r} > \overline{s}$). For example, s may correspond to the student's ability privately known by the school, and r to the opportunity cost from hiring privately known by the employer. For simplicity, each message m is identified with the receiver's type who is indifferent to act, so that m induces the receiver to act if and only if $r \leq m$.[5]

Proposition 1 simplifies the sender's problem of finding an optimal mechanism to a problem of finding an optimal distribution of messages.

Proposition 1 *Let H denote the marginal distribution of m under the optimal mechanism.*

[5]This example is more general than it may seem. In particular, it includes the case where $u(r, s) = b(r) c(s) + d(r)$ for some functions b, c, and d where b is positive and all functions satisfy certain regularity conditions. Indeed, the receiver acts whenever $-d(r)/b(r) \leq \mathbb{E}_{\phi}[c(s)|m]$, so redefining the receiver's type as $-d(r)/b(r)$ and the sender's type as $c(s)$ gives the required result.

Then H

$$\begin{array}{c} \textit{maximizes } \int_{\underline{r}}^{\overline{r}} G(m)\, dH(m) \\ \textit{subject to } F \textit{ is a mean-preserving spread of } H. \end{array} \tag{4}$$

The objective function in (4) is simply the probability that the receiver acts under a mechanism ϕ. If F is a mean-preserving spread of H, then F is more informative about the underlying (hypothetical) state than H (Blackwell (1953)). Since the sender has full commitment, she can garble her information to achieve any less informative distribution H than her prior F. If she then fully reveals this garbled information to the receiver, then the distribution of m will be H. Conversely, because the sender cannot make her information more precise in any sense, F must be a mean-preserving spread of H for any feasible mechanism.

Proposition 1 suggests that the sender faces a similar tradeoff to that of the binary example. The sender's marginal utility from sending m is $G(m)$. But besides requiring the expectation of m and s to be equal, the budget constraint now also requires the distribution of m to be less variable than the prior distribution. Proposition 2 shows that the shape of the optimal mechanism is determined by the curvature of G and the expectation of s.

Proposition 2 *In this example:*

1. *All mechanisms are equivalent if and only if G is linear on S.*

2. *ϕ_{full} is optimal if and only if G is convex on S.*

3. *ϕ_{no} is optimal if and only if the concave closure $\mathbf{G}$ of G on S is equal to G at $r_{no} \equiv \mathbb{E}_F[s]$ in that*
$$G(r_{no}) \geq \frac{r_2 - r_{no}}{r_2 - r_1} G(r_1) + \frac{r_{no} - r_1}{r_2 - r_1} G(r_2)$$
for all $r_1, r_2 \in S$ such that $r_1 < r_{no} < r_2$.[6]

4. *If G is convex on $[\underline{s}, s_i]$, concave on $[s_i, \overline{s}]$, and $G(r_{no}) < \mathbf{G}(r_{no})$, then the optimal mechanism reveals s for $s < s_c$ and sends the same message $r_c \equiv \mathbb{E}[s | s \geq s_c]$ for $s \geq s_c$ where $s_c < s_i$ is uniquely determined by*
$$g(r_c) = \frac{G(r_c) - G(s_c)}{r_c - s_c}.$$

[6] Intuitively, a concave closure of a function (defined on a convex set) is the smallest concave function that is everywhere greater than the original function.

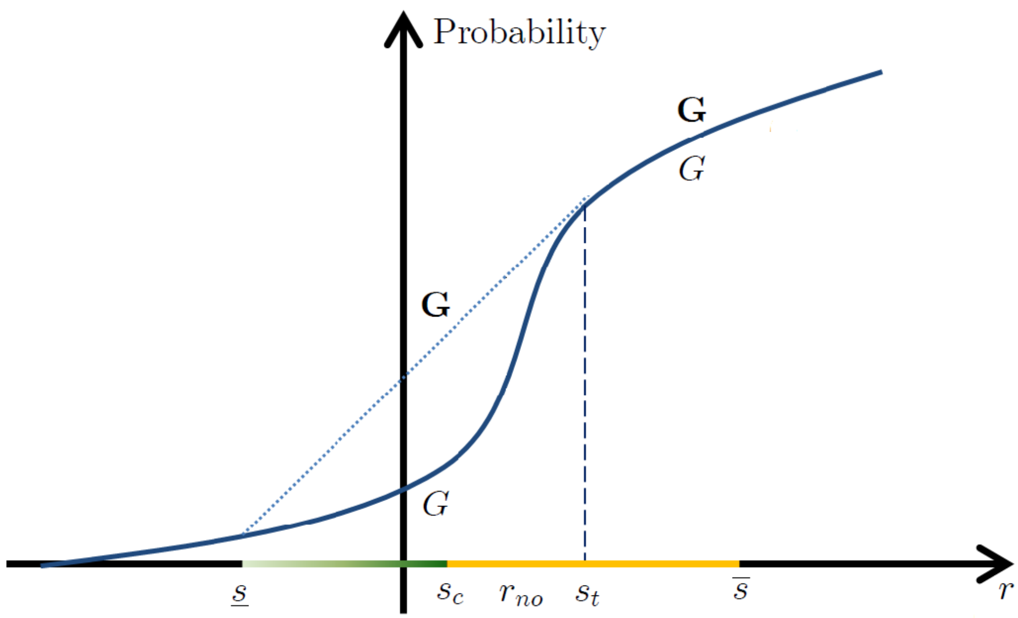

Figure 2: The distribution G and its concave closure $\mathbf{G}$ for $r \in S$.

The first three parts of Proposition 2 are straightforward because the optimal mechanism is the solution to problem (4). First, if G is linear, then the sender is risk neutral, so all mechanisms are equivalent. Second, if G is convex, then the sender is risk loving, so the full revelation mechanism is optimal. Third, if G is concave, then the sender is risk averse, so the no revelation mechanism is optimal.[7]

The last part of Proposition 2 derives the optimal mechanism under a natural assumption that the distribution G has an "S" shape as shown in Figure 2.[8] Assume that $\mathbb{E}_F[s] < s_t$ (equivalently $G(r_{no}) < \mathbf{G}(r_{no})$), otherwise ϕ_{no} is optimal by part 3. If F were to put strictly positive probabilities only on $\underline{s}$ and $\overline{s}$, then the optimal mechanism would send two messages $\underline{s}$ and s_t and the receiver would act with probability $\mathbf{G}(r_{no})$. This mechanism, however, is

[7] The mathematical structure of this example is similar to Ostrovsky and Schwarz (2010) who analyze information disclosure in matching markets. In particular, we can reinterpret this continuous example as if a student with ability s receives a transcript m according to a distribution $\phi(m|s)$ and then he is matched to an employer of quality $G(m)$. The main technical difference is that in Ostrovsky and Schwarz (2010) the function G is endogenously determined by information disclosure mechanisms of schools. The first three parts of Proposition 2 can alternatively be derived using tools developed in Kamenica and Gentzkow (2011). Moreover, these three parts are similar to the results obtained in Section VIII B of Rayo and Segal (2010). Proposition 1 and Part 4 of Proposition 2, however, are new to the literature to the best of my knowledge.

[8] It is straightforward (though notationally heavy) to characterize the optimal mechanism if G has more than one inflection point at which the curvature changes sign.

not feasible when F admits a density because s is equal to $\underline{s}$ with probability 0. Thus, the optimal mechanism reveals s for $s < s_c$ and sends the same message for all $s \geq s_c$ where the cutoff s_c is such that the sender is indifferent between revealing s_c or pooling it with $s \geq s_c$. In the school-employer application, if the average student ability is high, then the school should reveal no information about its students, otherwise it should fully separate bad students but should pool good and very good students.

Note that the optimal mechanism may be very sensitive to primitives of the model. For example, if G is almost uniform but strictly convex, then ϕ_{full} is uniquely optimal. However, if G is almost uniform but concave, then ϕ_{no} is uniquely optimal. This observation gives an explanation for why many similar-looking schools may choose very different disclosure policies regarding what information (if any) to report on transcripts (grading scale, class rank, distinctions).

I conclude by discussing comparative statics in this example. By Proposition 1, as F becomes more informative in the mean-preserving spread sense, the set of feasible mechanisms expands, so the sender's expected utility increases.[9] That is, an additional information about a student can only help a school if it can commit to a disclosure policy in advance. Moreover, Proposition 1 implies that as the sender's and receiver's priors become more favorable for acting (F increases and G decreases in the first-order stochastic dominance sense), the sender's expected utility increases. These monotone comparative statics results are similar in spirit to the results in Kolotilin (2013).

However, similarly to the binary example, the sender's and receiver's expected utilities are not monotonic in the precision of the receiver's private information. In particular, the sender's expected utility may decrease as the receiver's private information G becomes more precise in the mean-preserving spread sense. To see this, consider F that puts probability one on some s and note that the sender's expected utility $G(s)$ changes ambiguously. Moreover, the receiver's expected utility may decrease with the precision of his private information. To see this, suppose that G_1 is almost uniform but convex, and G_2 is concave and slightly more informative than G_1. By Proposition 2, ϕ_{full} is optimal under G_1 and ϕ_{no} is optimal under G_2. Thus, from the receiver's perspective a small gain from having more precise private information under G_2 is outweighed by a large loss from getting less precise information from the sender.

[9] In the two extreme cases, if F were to put probability one on some s, then the only feasible H would put probability one on $m = s$, but if F were to put strictly positive probabilities only on $\underline{s}$ and $\overline{s}$, then any H supported on S with $\mathbb{E}_H[s] = \mathbb{E}_F[s]$ would be feasible.

4 General Case

This section generalizes the examples of Section 3. The key assumption that is maintained throughout this section is that the receiver with a higher type is always more willing to act.[10] Section 4.1 develops the necessary machinery for characterization of an optimal mechanism. Section 4.2 characterizes necessary and sufficient conditions for optimality of the two most important mechanisms: the full revelation and no revelation mechanisms.

4.1 Characterization of Optimal Mechanism

If the sender's type is binary, then, similarly to the binary example, an optimal mechanism maximizes a linear utility function subject to a linear budget constraint. However, if the sender's type is not binary, then the budget constraint becomes multidimensional and it becomes hard to solve for an optimal mechanism. Nevertheless, an optimal mechanism always solves a linear program; so duality theory applies. Duality theory gives a relatively simple solution to the reverse problem of finding necessary and sufficient conditions on the primitives of the model that ensure that a candidate mechanism is optimal. For example, we can derive conditions of an environment under which an actual disclosure policy chosen by a school is optimal.[11]

Section 4 maintains the following *single-crossing assumption*: $v_H(r) \equiv \int_S \widetilde{u}(r,s)\, dH(s)$ crosses the horizontal axis once and from below for all distributions H on S where $\widetilde{u}(r,s) \equiv u(r,s)\, g(r|s)$; moreover, $r^*(s)$ is strictly decreasing in s where $r^*(s)$ is the unique r that solves $u(r,s) = 0$. The single-crossing assumption allows us to restrict attention to mechanisms ϕ in which a message m induces the receiver to act if and only if $r \geq m$.

The essence of the single-crossing assumption is that $v_H(r)$ crosses the horizontal axis at most once and in the same direction, the remaining requirements are just technical conditions.[12] The continuous example satisfies the single-crossing assumption, and the binary

[10] In the binary example, the optimistic receiver is more willing to act. In the continuous example, the receiver with a lower opportunity cost (a higher type $-r$) is more willing to act.

[11] Since all mechanisms are equivalent in the continuous example from Section 3.2 if G is linear, we know that any disclosure policy is optimal under some conditions.

[12] Indeed, extending $u(r,s)$ to $\widetilde{R} \supset R$ for all s and making $g(r|s)$ infinitesimally small for all s and $r \notin R$ yields that $v_H(r)$ crosses the horizontal axis exactly once on $\widetilde{R}$, not just at most once. Reordering R yields that $v_H(r)$ crosses the horizontal axis from below, not just in the same direction. Considering H that puts probability one on s yields that $u(r,s)$ crosses the horizontal axis once for all s, so $r^*(s)$ is well defined. Finally, reordering S yields that $r^*(s)$ is decreasing.

example satisfies this weak version of the single-crossing assumption.[13] To illustrate broad applicability of this assumption, Proposition 3 provides an alternate representation and primitive sufficient conditions for the weak version of the single-crossing assumption.

Proposition 3 *Let all assumptions imposed in Section 2 hold.*

1. *The function $v_H(r)$ crosses the horizontal axis at most once and from below for all distributions H if and only if for any $r_2 \geq r_1$ there exists a constant $b \geq 0$ such that $\widetilde{u}(r_2, s) \geq b\widetilde{u}(r_1, s)$ for all s.*

2. *If $u(r, s)$ is increasing in both r and s, and the density $g(r|s)$ has the monotone likelihood ratio property in that $g(r_2|s_2)\, g(r_1|s_1) - g(r_2|s_1)\, g(r_1|s_2) \geq 0$ for all $s_2 \geq s_1$ and $r_2 \geq r_1$, then $v_H(r)$ crosses the horizontal axis at most once and from below for all distributions H.*

Before turning to the general problem where both the sender's and receiver's types are continuous, it is instructive to consider the case where the receiver's type is continuous but the sender's type is binary in that $G(r|s)$ admits a density $g(r|s)$ but F is supported on $\underline{s}$ and $\overline{s}$. For all $r \in R^* \equiv [r^*(\overline{s}), r^*(\underline{s})]$, denote $p(r)$ as the probability of $\overline{s}$ at which the receiver r is indifferent to act. In the optimal mechanism, the distribution H of messages

$$\text{maximizes } \int_{R^*} \Pr(r \geq m|m)\, dH(m)$$

$$\text{subject to } \int_{R^*} p(m)\, dH(m) = \Pr(\overline{s})\,.$$ [14]

The objective function is the probability that the receiver acts and the constraint is the feasibility constraint that requires that posterior probabilities $\Pr(\overline{s}|m)$ average out to the prior probability $\Pr(\overline{s})$. Again, the objective function can be interpreted as a linear utility function and the constraint as a Bayesian budget constraint. As a result, the sender faces the same quantity-quality tradeoff as in the binary example of Section 3.1: sending a lower message m is more expensive (the price $p(m)$ is higher), but it has a greater impact on the receiver (the marginal utility $\Pr(r \geq m|m)$ is higher). To resolve this tradeoff, the

[13] The reader interested in an example that does not satisfy even the weak version of the single-crossing assumption is referred to the supplemental appendix.

[14] Explicitly, $p(r) = \widetilde{u}(r, \underline{s}) / (\widetilde{u}(r, \underline{s}) - \widetilde{u}(r, \overline{s}))$ and $\Pr(r \geq m|m) = p(m)\, \overline{G}(m|\overline{s}) + (1 - p(m))\, \overline{G}(m|\underline{s})$.

optimal mechanism sends at most two messages with the highest marginal utility-price ratio $\Pr(r \geq m|m)/p(m)$.[15]

In general (if both the sender's and receiver's types are continuous), the optimal mechanism is a distribution ϕ that

$$\text{maximizes} \int_{R\times S} \overline{G}(r|s)\, d\phi(r,s) \tag{5}$$

$$\text{subject to} \int_{R\times\widetilde{S}} d\phi(r,s) = \int_{\widetilde{S}} f(s)\, ds \text{ for any measurable set } \widetilde{S}\subset S, \tag{6}$$

$$\int_{\widetilde{R}\times S} \widetilde{u}(r,s)\, d\phi(r,s) = 0 \text{ for any measurable set } \widetilde{R}\subset R. \tag{7}$$

The objective function is the probability that the receiver acts under a mechanism ϕ. The first constraint (6) is the requirement that the marginal distribution of s for ϕ is F. Intuitively, (6) is a multidimensional Bayesian budget constraint. The second constraint (7) is the requirement that a message r makes the receiver r indifferent to act. Intuitively, (7) determines multidimensional prices of various messages.

The problem (5) is called the *primal problem.* This primal problem is an infinite dimensional linear program, because the objective function and both constraints are linear in a probability distribution ϕ.[16] The *dual problem* is to find bounded functions η and ν that

$$\text{minimize} \int_S \eta(s) f(s)\, ds \tag{8}$$

$$\text{subject to } \eta(s) + \widetilde{u}(r,s)\,\nu(r) \geq \overline{G}(r|s) \text{ for all } (r,s)\in R\times S. \tag{9}$$

Intuitively, the variables $\eta(s)$ and $\nu(r)$ are multipliers for constraints (6) and (7).

Say that ϕ is *feasible* for (5) if it is a distribution that satisfies (6) and (7). Similarly, say that η and ν are *feasible* for (8) if they are bounded functions that satisfy (9). Feasible ϕ and (η,ν) that solve their respective problems (5) and (8) are called *optimal solutions.*

The reader should not be concerned about how the dual problem is derived; what is important is the linkage between the primal and dual problems stated in Lemmas 1 and 2.

[15] The optimal mechanism is a solution to a linear program, so it is an extreme point of the constraint set. If s is binary, then the constraint is one dimensional, so the optimal mechanism sends at most two messages.

[16] The primal problem would be a finite dimensional linear program if the sets R and S were finite. I impose enough smoothness on functions u, G, and F to guarnatee that standard results for finite dimensional linear programs extend to an infinite dimensional case. If R and S are finite, neither full revelation mechanism nor no revelation mechanism is generically optimal, because G and F are step functions. For this reason, I assume that the sets S and R are intervals in which case both full revelation and no revelation mechanisms can be generically optimal.

Weak duality gives an easy way to check that candidate feasible solutions ϕ and (η, ν) are optimal:

Lemma 1 *If ϕ is feasible for (5), and (η, ν) is feasible for (8), then*

$$\int_S \eta(s) f(s) \, ds \geq \int_{R \times S} \overline{G}(r|s) \, d\phi(r,s). \tag{10}$$

Moreover, if inequality (10) holds with equality, then ϕ and (η, ν) are optimal solutions, and

$$\int_{R \times S} \left(\eta(s) + \widetilde{u}(r,s) \nu(r) - \overline{G}(r|s) \right) d\phi(r,s) = 0. \tag{11}$$

Strong duality establishes the existence of optimal solutions:

Lemma 2 *There exists an optimal mechanism ϕ, an optimal solution to the primal problem (5). There exists an optimal solution to the dual problem (8). Moreover, inequality (10) holds with equality for these optimal ϕ and (η, ν).*

In the next section, using duality theory, I derive necessary and sufficient conditions for the full revelation mechanism ϕ_{full} to be optimal and for the no revelation mechanism ϕ_{no} to be optimal. There are at least two reasons that make mechanisms ϕ_{no} and ϕ_{full} prominent, besides their widespread use. First, if the sender did not have commitment power, then ϕ_{no} would be the unique equilibrium outcome under unverifiable information of the sender in the sense of Crawford and Sobel (1982), and ϕ_{full} would be the unique equilibrium outcome under verifiable information of the sender in the sense of Milgrom (1981).[17] The second reason is that these two mechanisms are extremal:

Proposition 4 *Let the single-crossing assumption hold.*

1. *The receiver's expected utility under ϕ_{no} is strictly lower than under any other mechanism.*

2. *The receiver's expected utility under ϕ_{full} is strictly higher than under any other mechanism.*

[17] Under unverifiable communication, if the sender sent two different messages r_1 and r_2 in equilibrium, then she would strongly prefer to send $\min\{r_1, r_2\}$ regardless of s, which leads to a contradiction. Under verifiable communication, if the sender sent the same message r for two or more different s in equilibrium, then there would exist $\widetilde{s}$ such that the sender $\widetilde{s}$ sent r but $u(r, \widetilde{s}) > 0$, which leads to a contradiction because the sender $\widetilde{s}$ would strongly prefer to reveal $\widetilde{s}$ instead.

A more informed receiver is better at maximizing his expected utility by taking a more appropriate action, so a weak version of Proposition 4 is immediate. The single-crossing assumption together with the smoothness assumption guarantees that the strict version of Proposition 4 holds. Note that the strict version does not hold in the binary example: the optimal mechanism is different from ϕ_{no}, yet the receiver's expected utility under the optimal mechanism is the same as under ϕ_{no} for $p \geq 1/\sqrt{2}$.

4.2 Optimality of Specific Mechanisms

By definition, a mechanism ϕ is optimal if and only if it dominates all feasible mechanisms. This observation gives trivial necessary and sufficient conditions for optimality of ϕ. However, to check these conditions, one needs to compare ϕ with all feasible mechanisms, which requires a lot of comparisons. It turns out that for the optimality of ϕ, it is necessary and sufficient to check that only certain deviations from ϕ do not increase the probability that the receiver acts.

I now define the deviations that we need to check for optimality of ϕ_{full} and ϕ_{no}. Note that ϕ_{full} sends the message $r^*(s)$ for each $s \in S$, and ϕ_{no} sends the same message r_{no} for each $s \in S$, where r_{no} is the unique r that solves $\int_S \widetilde{u}(r,s) f(s)\, ds = 0$.

For any s_1, s_2, and r such that $r \in (r^*(s_2), r^*(s_1))$, consider the prior distribution of s that puts probabilities $\widetilde{u}(r,s_2) / (\widetilde{u}(r,s_2) - \widetilde{u}(r,s_1))$ on s_1 and $\widetilde{u}(r,s_1) / (\widetilde{u}(r,s_1) - \widetilde{u}(r,s_2))$ on s_2. The sender *prefers to reveal s_1 and s_2 than to pool them at r* if for this prior distribution, the full revelation mechanism, which sends $r^*(s_1)$ and $r^*(s_2)$ for s_1 and s_2, dominates the no revelation mechanism, which sends the same message r for s_1 and s_2. Mathematically, this requirement is given by:

$$\frac{\overline{G}(r^*(s_2)|s_2) - \overline{G}(r|s_2)}{\widetilde{u}(r,s_2)} \geq \frac{\overline{G}(r^*(s_1)|s_1) - \overline{G}(r|s_1)}{\widetilde{u}(r,s_1)}. \tag{12}$$

Similarly, the sender *prefers to pool s_1 and s_2 at r than to reveal them* if inequality (12) is reversed to "$\leq$". Finally, the sender *is indifferent to reveal s_1 and s_2 or to pool them at r* if (12) holds with equality.

Say that *s_1, s_2, s_3, r are feasible* if there exists the prior distribution that puts positive probabilities p_1, p_2, p_3 only on s_1, s_2, s_3 such that $\sum_{i=1}^{3} p_i \widetilde{u}(r_{no}, s_i) = 0$, and $\sum_{i=1}^{2} p_i \widetilde{u}(r, s_i) = 0$. The sender *prefers to pool s_1, s_2, s_3 at r_{no} than to pool s_1, s_2 at r and to reveal s_3* if for the above prior distribution, the no revelation mechanism, which sends r_{no} for s_1, s_2, s_3, dominates the mechanism that sends r for s_1, s_2 and $r^*(s_3)$ for s_3. Mathematically, this requirement is given by:

$$\begin{aligned}&\tfrac{1}{\widetilde{u}(r,s_2)}\left(\overline{G}\left(r_{no}|s_2\right)-\overline{G}\left(r|s_2\right)+\tfrac{\widetilde{u}(r_{no},s_2)}{\widetilde{u}(r_{no},s_3)}\left(\overline{G}\left(r^*\left(s_3\right)|s_3\right)-\overline{G}\left(r_{no}|s_3\right)\right)\right)\\ \geq\ &\tfrac{1}{\widetilde{u}(r,s_1)}\left(\overline{G}\left(r_{no}|s_1\right)-\overline{G}\left(r|s_1\right)+\tfrac{\widetilde{u}(r_{no},s_1)}{\widetilde{u}(r_{no},s_3)}\left(\overline{G}\left(r^*\left(s_3\right)|s_3\right)-\overline{G}\left(r_{no}|s_3\right)\right)\right).\end{aligned} \tag{13}$$

I now present the main result of this section, which states that it is necessary and sufficient to consider only the pairwise and triplewise deviations defined above for optimality of ϕ_{full} and ϕ_{no}, respectively.

Theorem 1 *Let the single-crossing assumption hold. Then:*

1. *All mechanisms are equivalent if and only if the sender is indifferent to reveal s_1 and s_2 than to pool them at r for all $s_1, s_2 \in S$ and $r \in R$ such that $r \in (r^*(s_2), r^*(s_1))$, so that there exists a strictly positive function $b(r)$ such that*

$$\widetilde{u}(r,s) = b(r)\left(\overline{G}\left(r^*(s)|s\right) - \overline{G}(r|s)\right) \text{ for all } r \in \left(r^*(\overline{s}), r^*(\underline{s})\right). \tag{14}$$

2. *ϕ_{full} is optimal if and only if the sender prefers to reveal s_1 and s_2 than to pool them at r for all $s_1, s_2 \in S$ and $r \in R$ such that $r \in (r^*(s_2), r^*(s_1))$, so that (12) holds.*

3. *ϕ_{no} is optimal if and only if the sender prefers to pool s_1, s_2, s_3 at r_{no} than to pool s_1, s_2 at r and to reveal s_3 for all feasible s_1, s_2, s_3, r, so that (13) holds.*[18]

The "only if" parts of Theorem 1 are straightforward because, for optimality of a candidate mechanism, we need to check all deviations from a candidate optimal mechanism, including those described in the theorem.

The "if" parts are derived using weak duality (Lemma 1). For a given candidate optimal mechanism ϕ, the complementarity condition (11) implies that (9) holds with equality ($\eta(s) = \overline{G}(r|s) - \widetilde{u}(r,s)\nu(r)$) at each (r,s) in the support of ϕ for a candidate optimal solution (η,ν) to the dual problem (8). Then we can find primitive conditions on u, G, and F such that the constraint (9) of the dual problem (8) is satisfied for all $(r,s) \in R \times S$ and some $\nu(r)$ (this step is known as Fourier-Motzkin elimination of $\nu(r)$). Weak duality implies that these conditions are sufficient for ϕ to be optimal. For candidates ϕ_{full} and ϕ_{no}, these algebraic conditions correspond to the pairwise deviations (12) and the triplewise deviations (13), respectively.

[18] As apparent from the proof, this part can be restated with the triplewise deviations only for $s_3 \to s_{no}$ where s_{no} is the unique s that solves $u(r_{no}, s) = 0$.

The intuition for "if" parts of Theorem 1 relies on Lemma 3. Consider a message r of a mechanism ϕ. This message r generates a lottery $\phi(s|r)$ that makes the receiver r indifferent to act. Lemma 3 shows that this lottery can be decomposed into simpler lotteries indexed by e in such a way that (i) the support of each lottery e contains at most two elements, and (ii) each lottery e makes the receiver r indifferent to act.[19]

Lemma 3 *Let the single-crossing assumption hold. For each mechanism $\phi(r,s)$, there exists a mechanism $\varphi(m,s)$ with two dimensional messages $m=(r,e)\in R\times[0,1]$ such that for each m, the support of $\varphi(.|m)$ contains at most two elements of S and $\int_S \widetilde{u}(r,s)\,d\varphi(s|m)=0$.*

We now discuss each "if" part of Theorem 1 in turn. Suppose that the sender is indifferent to reveal s_1 and s_2 or to pool them at r for all feasible s_1, s_2, and r. By Lemma 3 we can focus on mechanisms in which each message is sent only by some two types s_1 and s_2. Consider such a mechanism. Since the sender is indifferent to reveal s_1 and s_2 or to pool them, this mechanism is equivalent to the mechanism that differs only in that it reveals s_1 and s_2. Sequentially modifying the mechanism for each message, we get that any mechanism is equivalent to ϕ_{full}; so part 1 follows. Note that all mechanisms are equivalent in the knife-edge case when $\widetilde{u}(r,s)$ has representation (14).

Similar to this paper, Rayo and Segal (2010) assume that actions are binary and the sender has full commitment. In contrast to this paper, they allow the sender's utility to depend on both the action and the state. However, to get tractable results, they assume that the utility of the receiver from acting is $u(r,s)=r+s$, where r is uniformly distributed on $[-1,0]$, and the support of s is contained in the interval $[0,1]$. Under this assumption, all mechanisms are equivalent in my model, as part 1 of Theorem 1 shows.[20]

We now turn to part 2 of Theorem 1. Again we can focus on mechanisms in which each message is sent only by two types s_1 and s_2. Consider such a mechanism. Since the sender prefers to reveal s_1 and s_2 than to pool them, this mechanism is dominated by the mechanism that differs only in that it reveals s_1 and s_2. Sequentially modifying the mechanism for each message, we get that ϕ_{full} dominates all mechanisms; so part 2 follows.

Finally, I provide the intuition for a weaker version of part 3 of Theorem 1 with quadraplewise rather than triplewise deviations. Namely, if the sender prefers to pool s_1, s_1', s_2, s_2' at r_{no} than to pool s_1, s_1' at r_1 and to pool s_2, s_2' at r_2 for all feasible s_1, s_1', s_2, s_2', r_1, r_2, then

[19] In a recent paper, Golosov et al. (2013) use a similar result to construct a fully revealing equilibrium in a dynamic cheap talk game.

[20] Note that the continuous example of Section 3 has the same functional form of the receiver's utility (after redefining r as $-r$), but it does not assume that r is uniformly distributed.

ϕ_{no} is optimal. Again we can focus on mechanisms such that any message $r_1 \leq r_{no}$ is sent only by two types s_1 and s_1' and any message $r_2 \geq r_{no}$ is sent only by two types s_2 and s_2'. Any such mechanism is dominated by the mechanism that differs only in that it sends the message r_{no} instead of r_1 and r_2. Sequentially applying this argument for pairs of messages, we get that ϕ_{no} dominates all mechanisms; so the weaker version of part 3 follows.

5 Extension: Two-Way Communication

This section explores robustness of the benchmark model of Section 2 to introduction of communication from the receiver to sender. If the sender has full commitment, the revelation principle applies (Myerson (1982)).[21] Thus, it is without loss of generality to consider the following timing: 1. The sender publicly chooses a mechanism, a conditional distribution $\gamma(m|n,s)$ of a message m given the sender's type s and the receiver's report n. 2. The receiver's type r is drawn according to G. 3. The receiver privately observes r and makes a report $n \in N$. 4. A pair (m,s) is drawn according to γ and F. 5. The receiver gets a message m and takes an action a. 6. Utilities of the sender and receiver are realized.

Further, it is without loss of generality to focus on incentive compatible direct mechanisms in which: (i) the set of receiver's reports N coincides with the set R; (ii) a mechanism sends m_1 with probability $\gamma(m_1|n,s)$ and m_0 with probability $1-\gamma(m_1|n,s)$; (iii) the receiver r prefers to report $n=r$; and (iv) the receiver prefers to act if he receives m_1 and not to act if he receives m_0 for all r and $n=r$.

If the single crossing assumption holds and the sender's type is binary, as in the motivating example[22], then it is without loss of generality to focus on benchmark mechanisms where the receiver is not allowed to make reports:

Proposition 5 *Let the sender's type be binary in that F is supported on $\underline{s}$ and $\overline{s}$, and let the single-crossing assumption hold. Then the set of mappings from (s,r) to the receiver's action a that can be supported by a mechanism is the same under $\phi(m|s)$ and under $\gamma(m|n,s)$.*

However, if the single crossing assumption does not hold, the benchmark mechanisms can be improved upon as the following example shows:

[21] To nest my model into Myerson (1982), assume that the principal is the sender who designs a mechanism for two agents. The first agent has type s, has no action to take, and always gets zero utility. The second agent has type r, privately chooses $a=0,1$, and his utility is $au(r,s)$.

[22] To be precise, the proof of Proposition 5 assumes that the receiver's type is continuous. However, it is straightforward to generalize the proof of Proposition 5 if the receiver's type is discrete. This generalization nests the motivating binary example.

Example 1 *Let* $s = (s_1, s_2) \in \{0,1\} \times \{0,1\}$ *and* $r \in \{r_1, r_2, r_3\}$. *Moreover, let all combinations of* (s, r) *be equally likely. Finally, let* $u(s, r_1) = s_1 - 1$, $u(s, r_2) = s_2 - 1$, *and* $u(s, r_3) = 3/4 - s_1 s_2$.

Consider the following mechanism of the two-way communication game:

$$\gamma^*(m_1|s,n) = \begin{cases} 1 \text{ if } n = r_1 \text{ and } s_1 = 1, \text{ or if } n = r_2 \text{ and } s_2 = 1, \text{ or if } n = r_3, \\ 0 \text{ otherwise.} \end{cases}$$

Intuitively, this mechanism allows the receiver to learn at most one component of s. Clearly, under this mechanism, it is incentive compatible for the receiver to truthfully report r and act whenever he receives m_1. Thus, the probability that the receiver acts under γ^* is:

$$\Pr{}_{\gamma^*}(a = 1) = \Pr(r = r_1)\Pr(s_1 = 1) + \Pr(r = r_2)\Pr(s_2 = 1) + \Pr(r = r_3) = \frac{2}{3}.$$

However, the sender cannot induce the receiver to act with probability 2/3 in the benchmark model. To see this, note first that the receiver r_1 acts only if he is certain that $s_1 = 1$, and the receiver r_2 acts only if he is certain that $s_2 = 1$. Thus, under any mechanism ϕ, the probability that the receiver acts cannot exceed 2/3. The only possibility of how the sender could achieve this probability would be to reveal both s_1 and s_2, but in that case the receiver r_3 would not act when $(s_1, s_2) = (1, 1)$.

6 Conclusions

In this paper, I have studied optimal information disclosure mechanisms with two-side asymmetric information. The receiver bases his action not only on the information disclosed by the sender but also on his private information. Thus, from the sender's perspective, each message results in a stochastic action by the receiver. The analysis reveals an important quantity-quality tradeoff of messages. The optimal mechanism finds a balance between these two conflicting objectives. This balance is easiest to explain when the sender's information has a binary structure. In this case, the prior distribution of the sender's information imposes a budget constraint on the frequencies of various messages, whereas the distribution of the receiver's information determines the sender's expected utility, which is linear in the frequencies of various messages. The optimal mechanism sends messages with the highest marginal utility-price ratio.

I also derive interesting non-monotone comparative statics results with respect to the receiver's private information for the binary example in which the sender is perfectly informed but the receiver is partially informed. If the receiver's private information is either very precise or very imprecise, then the sender's expected utility decreases and the receiver's expected utility increases with the precision of the receiver's information. However, if the precision of the receiver's information is intermediate, then these results can be overturned. Surprisingly, the receiver may become worse off as his private information becomes more precise. Thus, if there is an earlier stage when the receiver can publicly choose how informed he will be, he may not want to be as informed as possible.

The paper also makes several technical contributions, which can be applied to other models of information disclosure. First, it identifies and characterizes the single-crossing assumption that is crucial for tractable results and for the quantity-quality tradeoff. Second, it provides a simple guess and verify method based on duality theory that allows to check that a candidate mechanism is optimal. Third, it makes a first step towards understanding when mechanisms with two-way communication outperform mechanisms with one-way communication.

Appendix: Proofs

Proof of Proposition 1. For any mechanism ϕ, $r = \mathbb{E}_{\phi}[s|r]$, which implies that F is a mean-preserving spread of H. Conversely, if F is a mean-preserving spread of H, then s has the same distribution as $r+z$ for some z such that $\mathbb{E}[z|r] = 0$. Define $\phi(\widetilde{r}, \widetilde{s}) = \Pr(r \leq \widetilde{r}, r+z \leq \widetilde{s})$ for all $(\widetilde{r}, \widetilde{s}) \in R \times S$. For this ϕ, the marginal distribution of s is F and $\mathbb{E}_{\phi}[s-r|r] = \mathbb{E}[z|r] = 0$. Therefore, ϕ is a feasible mechanism. Finally, $\int_{-\infty}^{\infty} G(m)\, dH(m)$ is simply the probability that the receiver acts. ■

Proof of Proposition 2. The proof relies on results from Section 4 and, therefore, should be read after Section 4. Clearly, this example satisfies the single-crossing assumption of Section 4 after the following change in variables: $\widetilde{r} = -r$. Thus, all results of Section 4 apply. With this change of variables, $r^*(s) = -s$, and $\overline{G}_{\widetilde{r}}(x) \equiv \Pr(\widetilde{r} > x)$ is equal to $G(-x)$. I prove each part in turn.

1. By part 1 of Theorem 1, all mechanisms are equivalent if and only if

$$s - r = b(r)(G(s) - G(r))$$

for some positive b and all $s \in S$, $r \in S$, which is equivalent to G being linear on R.

2. By part 2 of Theorem 1, ϕ_{full} is optimal if and only if the sender prefers to reveal s_1

and s_2 than to pool them at $\widetilde{r}$ where $\widetilde{r} \in (-s_2, -s_1)$, which is equivalent to

$$G(r) \leq \frac{s_2 - r}{s_2 - s_1} G(s_1) + \frac{r - s_1}{s_2 - s_1} G(s_2).$$

Clearly, this inequality holds if and only if G is convex on S.

3. By part 3 of Theorem 1, ϕ_{no} is optimal if and only if the sender prefers to pool s_1, s_2, s_3 at $\widetilde{r}_{no}$ than to pool s_1, s_2 at $\widetilde{r}$ and to reveal s_3 for all possible s_1, s_2, s_3, $\widetilde{r}$, which is equivalent to

$$G(r_{no}) \geq \frac{(s_3 - r_{no})}{(s_3 - r)} G(r) + \frac{(r_{no} - r)}{(s_3 - r)} G(s_3)$$

for all $r \in S$ and $s_3 \in S$ such that $(r - r_{no})(s_3 - r_{no}) < 0$. The change of variables $r_2 = s_3$ and $r_1 = r$ completes the proof.

4. Lemma 1 implies that the described mechanism is optimal if there exists feasible (η, ν) for (8):

$$\eta(s) + (s - r)\nu(r) \geq G(r) \text{ for all } (r, s) \in R \times S \tag{15}$$

such that weak duality condition (10) holds with equality. I now construct (η, ν) that satisfies (15). Note that condition (15) bounds ν only from one side for $r \notin S$. In particular, $\nu(r) \geq (G(r) - \eta(s)) / (s - r)$ if $r < \underline{s}$ and $\nu(r) \leq (\eta(s) - G(r)) / (r - s)$ if $r > \overline{s}$. Thus, we can set $\nu(r) = 0$ if $r < \underline{s}$ and $v(r) = -K$ if $r > \overline{s}$ where K is sufficiently large. (To see that $0 \geq (G(r) - \eta(s)) / (s - r)$ if $r < \underline{s}$, note that $\eta(s) \geq G(s)$ for all $s \in S$ as follows from (15) if $s = r$.) For $(r, s) \in S \times S$, we can set:

$$\eta(s) = \begin{cases} G(s) \text{ if } s \in [\underline{s}, s_c], \\ G(r_c) + g(r_c)(s - r_c) \text{ if } s \in (s_c, \overline{s}], \end{cases}$$

$$\nu(r) = \begin{cases} -g(r) \text{ if } r \in [\underline{s}, s_c], \\ -g(r_c) \text{ if } r \in (s_c, \overline{s}]. \end{cases}$$

It is straightforward to verify that η is convex and greater than G and $-\nu$ is a subderivative of η. Thus, (15) holds. Further, (15) holds with equality if (r, s) lies in the support of the described mechanism. Thus, weak duality condition (10) holds. ■

Proof of Proposition 3. Mathematically, the statement that $v_H(r)$ crosses the horizontal axis at most once and from below means that $v_H(r_1) \geq 0$ and $r_2 \geq r_1$ imply $v_H(r_2) \geq 0$. Denote $\widetilde{u}_i(s)$ as the function $\widetilde{u}(r_i, s)$ of s. If $\widetilde{u}_2(s) \geq b\widetilde{u}_1(s)$ for all s and some constant $b \geq 0$, then $v_H(r)$ crosses the horizontal axis at most once and from below for all distributions H because $\int_S \widetilde{u}_1(s)\, dH(s) \geq 0$ and $r_2 \geq r_1$ imply:

$$\int_S \widetilde{u}_2(s)\, dH(s) \geq b \int_S \widetilde{u}_1(s)\, dH(s) \geq 0.$$

Conversely, suppose that $v_H(r_1) \geq 0$ and $r_2 \geq r_1$ imply $v_H(r_2) \geq 0$ and let us show that there exists $b \geq 0$ such that $\widetilde{u}_2(s) \geq b\widetilde{u}_1(s)$ for all s. This result is obvious if $\widetilde{u}_1(s) < 0$ for all s. Consider now the case in which $\widetilde{u}_1(s) \geq 0$ for some s. Suppose to get a contradiction that there does not exist the required $b \geq 0$. Then the function $\widetilde{u}_2(s)$ does not belong to the closed convex cone C defined as the set of functions which can be represented as $d\widetilde{u}_1(s) + v(s)$ for some constant $d \geq 0$ and some continuous positive function $v(s)$. By the Separating Hyperplane Theorem (Corollary 5.84 of Aliprantis and Border (2006)), there exists a continuous linear functional ψ satisfying $\psi(\widetilde{u}_2) < 0$ and $\psi(c) \geq 0$ for all $c \in C$. By the Riesz Representation Theorem (Theorem 6 in Section 36 of Kolmogorov and Fomin (1975)), ψ can be represented in the form $\psi(c) = \int_S c(s)\, d\Psi(s)$, where Ψ is a function of bounded variation on S. Define the function H as $H(s) = \Psi(s)/V(\Psi)$, where $V(\Psi) > 0$ denotes the total variation of Ψ on S. Recall that the set C contains all positive continuous functions and $\int_S c(s)\, d\Psi(s) \geq 0$ for all $c \in C$. Applying the Dominated Convergence Theorem (Theorem 11.21 of Aliprantis and Border (2006)) to an appropriate sequence of positive continuous functions converging to the indicator function $\mathbf{1}_{[s_1,s_2]}$ yields that $\Psi(s_2) - \Psi(s_1) \geq 0$ for all $s_2 > s_1$, which in turn implies that H is a distribution function on S. Recalling that $V(\Psi) > 0$, $\widetilde{u}_1 \in C$, $\psi(\widetilde{u}_2) < 0$ and $\psi(c) \geq 0$ for all $c \in C$ yields $v_H(r_1) \geq 0$ and $v_H(r_2) < 0$, which is a contradiction.

Suppose that $v_H(r_1) \geq 0$, $r_2 \geq r_1$, and the functions u and g satisfy the suppositions of the second part. The second part follows from:

$$\begin{aligned}
v_H(r_2) &= \int_S u(r_2, s)\, g(r_2|s)\, dH(s) \\
&\geq \int_S u(r_1, s)\, g(r_2|s)\, dH(s) \\
&\geq \frac{\int_S g(r_2|s)\, dH(s)}{\int_S g(r_1|s)\, dH(s)} \int_S u(r_1, s)\, g(r_1|s)\, dH(s) \\
&= \frac{\int_S g(r_2|s)\, dH(s)}{\int_S g(r_1|s)\, dH(s)} v_H(r_1) \\
&\geq 0.
\end{aligned}$$

The first inequality holds because u is increasing in r. Since g has the monotone likelihood ratio, the distribution of s given r_2 first-order stochastically dominates the distribution of s given r_1 in that:

$$\frac{\int_{\underline{s}}^{\widetilde{s}} g(r_2|s)\, dH(s)}{\int_S g(r_2|s)\, dH(s)} \leq \frac{\int_{\underline{s}}^{\widetilde{s}} g(r_1|s)\, dH(s)}{\int_S g(r_1|s)\, dH(s)} \text{ for all } \widetilde{s} \in S,$$

as Milgrom (1981) shows. Thus, the second inequality holds because the function $u(r_1, s)$ is increasing in s. ■

Proof of Lemma 1. The proof of similar results can be found in Anderson and Nash (1987). However, to make the paper self-contained, I prove this lemma.

Multiplying (6) by η and integrating over S gives

$$\int_S \eta(s) f(s) ds = \int_{R\times S} \eta(s) d\phi(r,s).$$

Multiplying (7) by ν and integrating over R gives

$$\int_{R\times S} \widetilde{u}(r,s)\nu(r) d\phi(r,s) = 0.$$

Summing up these two equalities gives

$$\int_S \eta(s) f(s) ds = \int_{R\times S} (\eta(s) + \widetilde{u}(r,s)\nu(r)) d\phi(r,s). \tag{16}$$

Integrating (9) over $R \times S$ gives

$$\int_{R\times S} \overline{G}(r|s) d\phi(r,s) \leq \int_{R\times S} (\eta(s) + \widetilde{u}(r,s)\nu(r)) d\phi(r,s). \tag{17}$$

Conditions (16) and (17) yield (10).

Suppose that inequality (10) holds with equality for some feasible (η, ν) and ϕ:

$$\int_S \eta(s) f(s) ds = \int_{R\times S} \overline{G}(r|s) d\phi(r,s). \tag{18}$$

Consider any other feasible $\widetilde{\phi}$. Inequality (10) implies

$$\int_{R\times S} \overline{G}(r|s) d\widetilde{\phi}(r,s) \leq \int_S \eta(s) f(s) ds.$$

Combining this inequality with (18) gives

$$\int_{R\times S} \overline{G}(r|s) d\widetilde{\phi}(r,s) \leq \int_{R\times S} \overline{G}(r|s) d\phi(r,s),$$

showing that ϕ is an optimal solution to the primal problem (5). An analogous argument proves that (η, ν) is optimal solutions to (8). Finally, combining (16) and (18) for optimal ϕ and (η, ν) gives (11). ■

Proof of Lemma 2. I omit the proof of this lemma because it is not used in the subsequent analysis and its proof essentially repeats that of Theorem 5.2 in Anderson and Nash (1987). ■

Proof of Proposition 4. I start by proving the first part. The receiver's expected utilities under any mechanism ϕ and the no revelation mechanism ϕ_{no} are:

$$\begin{aligned}\mathbb{E}_{\phi}[u] &= \int_{R\times S}\left(\int_{r}^{\overline{r}}\widetilde{u}(\widetilde{r},s)\,d\widetilde{r}\right)d\phi(r,s),\\ \mathbb{E}_{\phi_{no}}[u] &= \int_{S}\left(\int_{r_{no}}^{\overline{r}}\widetilde{u}(\widetilde{r},s)\,d\widetilde{r}\right)f(s)\,ds\\ &= \int_{R\times S}\left(\int_{r_{no}}^{\overline{r}}\widetilde{u}(\widetilde{r},s)\,d\widetilde{r}\right)d\phi(r,s).\end{aligned}$$

The first two lines hold because a message m induces the receiver r to act if and only if $r\geq m$ and ϕ_{no} sends the message r_{no} regardless of s. The third line holds because the marginal distribution of s for any mechanism ϕ coincides with the prior distribution of s. For a mechanism ϕ, denote the conditional distribution of s given a message r by $\phi(s|r)$ and the marginal distribution of a message r by $\phi(r)$. Fubini's Theorem (Theorem 11.27 of Aliprantis and Border (2006)) gives

$$\begin{aligned}\mathbb{E}_{\phi}[u]-\mathbb{E}_{\phi_{no}}[u] = &\int_{\underline{r}}^{r_{no}}\left[\int_{r}^{r_{no}}\left(\int_{S}\widetilde{u}(\widetilde{r},s)\,d\phi(s|r)\right)d\widetilde{r}\right]d\phi(r)\\ &-\int_{r_{no}}^{\overline{r}}\left[\int_{r_{no}}^{r}\left(\int_{S}\widetilde{u}(\widetilde{r},s)\,d\phi(s|r)\right)d\widetilde{r}\right]d\phi(r).\end{aligned} \tag{19}$$

By the single-crossing assumption, we have $\int_{S}\widetilde{u}(\widetilde{r},s)\,d\phi(s|r)>0$ for $\widetilde{r}>r$. Therefore, $\int_{r}^{r_{no}}\left(\int_{S}\widetilde{u}(\widetilde{r},s)\,d\phi(s|r)\right)d\widetilde{r}>0$ for $r<r_{no}$. Since $\phi(r)$ of any mechanism ϕ that differs from ϕ_{no} puts strictly positive probability on messages in $[\underline{r},r_{no})$, the first integral in (19) is strictly positive. The analogous argument shows that the second integral in (19) is strictly negative; so $\mathbb{E}_{\phi}[u]-\mathbb{E}_{\phi_{no}}[u]>0$ for any ϕ that differs from ϕ_{no}.

I now prove the second part. The receiver's expected utility under ϕ_{full} is

$$\begin{aligned}\mathbb{E}_{\phi_{full}}[u] &= \int_{S}\left(\int_{r^*(s)}^{\overline{r}}\widetilde{u}(\widetilde{r},s)\,d\widetilde{r}\right)f(s)\,ds\\ &= \int_{R\times S}\left(\int_{r^*(s)}^{\overline{r}}\widetilde{u}(\widetilde{r},s)\,d\widetilde{r}\right)d\phi(r,s).\end{aligned}$$

Fubini's Theorem together with the condition $\widetilde{u}(r^*(s),s)=0$ gives

$$\begin{aligned}\mathbb{E}_{\phi_{full}}[u]-\mathbb{E}_{\phi}[u] = &\int_{S}\int_{r>r^*(s)}\left(\int_{r^*(s)}^{r}\widetilde{u}(\widetilde{r},s)\,d\widetilde{r}\right)d\phi(r,s)\\ &-\int_{S}\int_{r<r^*(s)}\left(\int_{r}^{r^*(s)}\widetilde{u}(\widetilde{r},s)\,d\widetilde{r}\right)d\phi(r,s).\end{aligned} \tag{20}$$

By the single-crossing assumption, we have $\widetilde{u}(\widetilde{r},s)>0$ for $\widetilde{r}>r^*(s)$; so $\int_{r^*(s)}^{r}\widetilde{u}(\widetilde{r},s)\,d\widetilde{r}>0$ for $r>r^*(s)$. Any ϕ that differs from ϕ_{full} puts strictly positive probability on the event $r>r^*(s)$, otherwise $\int_{R\times S}\widetilde{u}(r,s)\,d\phi(r,s)$ would be strictly negative rather than zero. Therefore, the first integral in (20) is strictly positive. The analogous argument shows that the second integral in (20) is strictly negative; so $\mathbb{E}_{\phi_{full}}[u] - \mathbb{E}_{\phi}[u] > 0$ for any ϕ that differs from ϕ_{full}. ■

Proof of Theorem 1. I prove each part in turn.

The "only if" part of 1. Suppose to get a contradiction that there exist s_1, s_2, and r such that the sender is not indifferent to reveal s_1 and s_2 or to pool them at r. Consider two mechanisms that differ only in that one sends different messages for $s \in [s_1, s_1 + \varepsilon_1]$ and $s \in [s_2 - \varepsilon_2, s_2]$, and the other sends the same message for $s \in [s_1, s_1 + \varepsilon_1] \cup [s_2 - \varepsilon_2, s_2]$ where ε_1 and ε_2 are sufficiently small and satisfy

$$\int_{s_1}^{s_1+\varepsilon_1} \widetilde{u}(r,s) f(s)\, ds + \int_{s_2-\varepsilon_2}^{s_2} \widetilde{u}(r,s) f(s)\, ds = 0.$$

Clearly, these two mechanisms are not equivalent. Therefore, if all mechanisms are equivalent, then the sender is indifferent to reveal s_1 and s_2 or to pool them at r for all $r \in (r^*(\overline{s}), r^*(\underline{s}))$ and all s_1, s_2 such that $r \in (r^*(s_2), r^*(s_1))$:

$$\frac{\overline{G}(r^*(s_1)|s_1) - \overline{G}(r|s_1)}{\widetilde{u}(r,s_1)} = \frac{\overline{G}(r^*(s_2)|s_2) - \overline{G}(r|s_2)}{\widetilde{u}(r,s_2)}.$$

Therefore, we can define the required b as $\widetilde{u}(r,s) / \left(\overline{G}(r^*(s)|s) - \overline{G}(r|s)\right)$, which is strictly positive and does not depend on s.

The "if" part of 1. Consider any mechanism ϕ. Substituting (14) into (7) gives

$$\int_{R\times S} \overline{G}(r|s)\, d\phi(r,s) = \int_{R\times S} \overline{G}(r^*(s)|s)\, d\phi(r,s).$$

Taking into account (6) gives

$$\int_{R\times S} \overline{G}(r|s)\, d\phi(r,s) = \int_{S} \overline{G}(r^*(s)|s) f(s)\, ds,$$

which implies that the probability that the receiver acts is the same for all mechanisms.

The "only if" part of 2. Suppose to get a contradiction that there exist s_1, s_2, and r such that it is strictly better to pool s_1 and s_2 at r than to reveal them. Consider the mechanism that differs from ϕ_{full} only in that it sends the same message for $s \in [s_1, s_1 + \varepsilon_1] \cup [s_2 - \varepsilon_2, s_2]$ where ε_1 and ε_2 are sufficiently small and satisfy

$$\int_{s_1}^{s_1+\varepsilon_1} \widetilde{u}(r,s) f(s)\, ds + \int_{s_2-\varepsilon_2}^{s_2} \widetilde{u}(r,s) f(s)\, ds = 0.$$

Clearly, this mechanism strictly dominates ϕ_{full}.

The "if" part of 2. The complementarity condition (11) suggests that

$$\eta(s)+\widetilde{u}(r^*(s),s)\nu(r)=\overline{G}(r^*(s)|s) \text{ for all } s\in S.$$

Taking into account that $\widetilde{u}(r^*(s),s)=0$ gives $\eta(s)=\overline{G}(r^*(s)|s)$ for all $s\in S$. Note that weak duality condition (10) is satisfied with equality for $\eta(s)=\overline{G}(r^*(s)|s)$. Therefore, by Lemma 1, ϕ_{full} is optimal if there exists ν such that

$$\overline{G}(r^*(s)|s)+\widetilde{u}(r,s)\nu(r)\geq\overline{G}(r|s) \text{ for all } (r,s)\in R\times S, \tag{21}$$

which is equivalent to

$$\frac{\overline{G}(r|s_2)-\overline{G}(r^*(s_2)|s_2)}{\widetilde{u}(r,s_2)}\leq\nu(r)\leq\frac{\overline{G}(r^*(s_1)|s_1)-\overline{G}(r|s_1)}{-\widetilde{u}(r,s_1)}$$

for all $r\in(r^*(\overline{s}),r^*(\underline{s}))$ and s_1, s_2 such that $r\in(r^*(s_2),r^*(s_1))$. (For $r\notin(r^*(\overline{s}),r^*(\underline{s}))$, the existence of ν is obvious because (21) bounds ν only from one side.) To summarize, (12) suffices for optimality of ϕ_{full}.

The "only if" part of 3. Suppose to get a contradiction that there exist s_1, s_2, s_3, r such that it is strictly better to pool s_1, s_2 at r and to reveal s_3 than to pool s_1, s_2, s_3 at r_{no}. Consider the mechanism that differs from ϕ_{no} only in that it sends one message if $s\in[s_1,s_1+\varepsilon_1]\cup[s_2,s_2+\varepsilon_2]$ and another message if $s\in[s_3-\varepsilon_3,s_3]$ where ε_1, ε_2, and ε_3 are sufficiently small and satisfy

$$\begin{aligned}\int_{s_1}^{s_1+\varepsilon_1}\widetilde{u}(r,s)f(s)\,ds+\int_{s_2}^{s_2+\varepsilon_2}\widetilde{u}(r,s)f(s)\,ds &= 0,\\ \sum_{i=1,2}\int_{s_i}^{s_i+\varepsilon_i}\widetilde{u}(r_{no},s)f(s)\,ds+\int_{s_3-\varepsilon_3}^{s_3}\widetilde{u}(r_{no},s)f(s)\,ds &= 0.\end{aligned}$$

Clearly, this mechanism strictly dominates ϕ_{no}.

The "if" part of 3. The complementarity condition (11) suggests that

$$\eta(s)+\widetilde{u}(r_{no},s)\nu(r_{no})=\overline{G}(r_{no}|s) \text{ for all } s\in S. \tag{22}$$

Note that weak duality condition (10) is satisfied with equality for $\eta(s)=\overline{G}(r_{no}|s)-\widetilde{u}(r_{no},s)\nu(r_{no})$. Therefore, by Lemma 1, ϕ_{no} is optimal if there exists ν such that

$$\overline{G}(r_{no}|s)-\widetilde{u}(r_{no},s)\nu(r_{no})+\widetilde{u}(r,s)\nu(r)\geq\overline{G}(r|s) \text{ for all } (r,s)\in R\times S. \tag{23}$$

Since $\widetilde{u}(r,s)$ is continuous and $\nu(r)$ is bounded, inequality (23) holds if it holds for $(r,s)\in R\times S$ such that $r\neq r^*(s)$:

$$\frac{\overline{G}(r|s_2)-\left(\overline{G}(r_{no}|s_2)-\widetilde{u}(r_{no},s)\nu(r_{no})\right)}{\widetilde{u}(r,s_2)}\leq\nu(r)\leq\frac{\left(\overline{G}(r_{no}|s_1)+\widetilde{u}(r_{no},s)\nu(r_{no})\right)-\overline{G}(r|s_1)}{-\widetilde{u}(r,s_1)} \tag{24}$$

for all $r \in (r^*(\overline{s}), r^*(\underline{s}))$, and $s_1, s_2 \in S$ such that $r \in (r^*(s_2), r^*(s_1))$. (For $r \notin (r^*(\overline{s}), r^*(\underline{s}))$, the existence of ν is obvious because (23) bounds ν only from one side.)

At $r = r_{no}$, both sides of (24) become $\nu(r_{no})$. Thus, for (24) to be satisfied everywhere, the derivatives of both sides of (24) with respect to r evaluated at $r = r_{no}$ must coincide, which gives

$$\nu(r_{no}) = \frac{g(r_{no}|s_2)\,\widetilde{u}(r_{no}, s_1) - g(r_{no}|s_1)\,\widetilde{u}(r_{no}, s_2)}{\frac{\partial \widetilde{u}(r_{no}, s_1)}{\partial r}\widetilde{u}(r_{no}, s_2) - \frac{\partial \widetilde{u}(r_{no}, s_2)}{\partial r}\widetilde{u}(r_{no}, s_1)}. \tag{25}$$

Taking the limit $s_2 \downarrow s_{no}$ in (25), where s_{no} is the unique s that solves $u(r_{no}, s) = 0$, gives

$$\nu(r_{no}) = -\frac{g(r_{no}|s_{no})}{\partial \widetilde{u}(r_{no}, s_{no})/\partial r}. \tag{26}$$

Substituting $\nu(r_{no})$ from (26) into (24) implies that ϕ_{no} is optimal if

$$\frac{\overline{G}(r|s_2) - \left(\overline{G}(r_{no}|s_2) + \frac{\widetilde{u}(r_{no}, s_2) g(r_{no}|s_{no})}{\partial \widetilde{u}(r_{no}, s_{no})/\partial r}\right)}{\widetilde{u}(r, s_2)} \leq \frac{\left(\overline{G}(r_{no}|s_1) + \frac{\widetilde{u}(r_{no}, s_1) g(r_{no}|s_{no})}{\partial \widetilde{u}(r_{no}, s_{no})/\partial r}\right) - \overline{G}(r|s_1)}{-\widetilde{u}(r, s_1)} \tag{27}$$

for all $r \in (r^*(\overline{s}), r^*(\underline{s}))$, and $s_1, s_2 \in S$ such that $r \in (r^*(s_2), r^*(s_1))$. Inequality (13) becomes (27) after taking the limit $s_3 \to s_{no}$. Since (13) holds for all s_3 by the supposition, (27) also holds because all functions are smooth; so ϕ_{no} is optimal. ■

Proof of Lemma 3. Consider any $\widetilde{r}$ in the support of ϕ. For a moment assume that $\phi(s|\widetilde{r})$ admits a density. Because $\mathbb{E}_\phi[u(\widetilde{r}, s)|r = \widetilde{r}] = 0$, we can construct a decreasing function $v_1(e)$ and an increasing function $v_2(e)$ defined on $[0,1]$ such that $\Pr_\phi(s \in [v_1(e), v_2(e)]|r = \widetilde{r}) = e$ and $\mathbb{E}_\phi[u(\widetilde{r}, s)|r = \widetilde{r}, s \in [v_1(e), v_2(e)]] = 0$. If $\phi(s|\widetilde{r})$ does not admit a density, then a similar result holds but with possible randomization at the boundaries $v_1(e)$ and $v_2(e)$. Formally, there exists a quadruple function (v_1, v_2, q_1, q_2) from $R \times [0,1]$ to $[\min_{s \in S} \widetilde{u}(r, s), 0] \times [0, \max_{s \in S} \widetilde{u}(r, s)] \times [0,1] \times [0,1]$ such that

$$\begin{aligned}
&\textstyle\int_{v_1(\widetilde{r}, e) < \widetilde{u}(\widetilde{r}, s) < v_2(\widetilde{r}, e)} \widetilde{u}(\widetilde{r}, s)\, d\phi(s|\widetilde{r}) \\
&+ \textstyle\sum_{i=1,2} v_i(\widetilde{r}, e)\, q_i(\widetilde{r}, e) \Pr{}_\phi(\widetilde{u}(\widetilde{r}, s) = v_i(\widetilde{r}, e)|r = \widetilde{r}) = 0, \\
&\Pr{}_\phi(v_1(\widetilde{r}, e) < \widetilde{u}(\widetilde{r}, s) < v_2(\widetilde{r}, e)|r = \widetilde{r}) \\
&+ \textstyle\sum_{i=1,2} q_i(\widetilde{r}, e) \Pr{}_\phi(\widetilde{u}(\widetilde{r}, s) = v_i(\widetilde{r}, e)|r = \widetilde{r}) = e
\end{aligned}$$

for all $(\widetilde{r}, e) \in R \times [0,1]$. Define distribution φ of $(\widetilde{r}, e, s)$ as follows. The marginal distribution of $\widetilde{r}$ for φ coincides with the marginal distribution of $\widetilde{r}$ for ϕ. The conditional distribution of e given $\widetilde{r}$ is uniform on the unit interval $[0,1]$. The conditional distribution of s given $\widetilde{r}$ and e puts probabilities p_1 and $1 - p_1$ on s_1 and s_2, where s_1 and s_2 satisfy $\widetilde{u}(\widetilde{r}, s_1) = v_1(\widetilde{r}, e)$ and $\widetilde{u}(\widetilde{r}, s_2) = v_2(\widetilde{r}, e)$, and p_1 solves $p_1 v_1(\widetilde{r}, e) + (1 - p_1) v_2(\widetilde{r}, e) = 0$. Clearly, this φ satisfies the required properties of Lemma 3. ■

Proof of Proposition 5. Take any mechanism γ and define $\phi(r|s) \equiv \gamma(m_1|r,s)$. To prove this proposition, it is sufficient to show that ϕ is a distribution function that satisfies constraints (6), (7), and the sender's expected utility coincides with (5). Since $\phi(r|s) = 1$ for $r \geq r^*(\underline{s})$ and $\phi(r|s) = 0$ for $r < r^*(\overline{s})$, constraint (6) is satisfied. The receiver r prefers to report $n = r$ rather than r' only if:

$$\sum_{s=\underline{s},\overline{s}} \widetilde{u}(r,s)\left(\phi(r|s) - \phi(r'|s)\right)\Pr(s) \geq 0. \tag{28}$$

Writing (28) for $(r,r') = (r_2,r_1)$ and $(r,r') = (r_1,r_2)$ with $r_1, r_2 \in (r^*(\overline{s}), r^*(\underline{s}))$ yields:

$$-\frac{\widetilde{u}(r_2,\underline{s})}{\widetilde{u}(r_2,\overline{s})}\Delta(r_2,r_1,\underline{s}) \leq \Delta(r_2,r_1,\overline{s}) \leq -\frac{\widetilde{u}(r_1,\underline{s})}{\widetilde{u}(r_1,\overline{s})}\Delta(r_2,r_1,\underline{s}) \tag{29}$$

where $\Delta(r_2,r_1,s) \equiv (\phi(r_2|s) - \phi(r_1|s))\Pr(s)$. If $r_2 \geq r_1$, then $\widetilde{u}(r_2,s) \geq b\widetilde{u}(r_1,s)$ for some $b \geq 0$ (see Part 1 of Proposition 3); so

$$0 < -\frac{\widetilde{u}(r_2,\underline{s})}{\widetilde{u}(r_2,\overline{s})} \leq -\frac{\widetilde{u}(r_1,\underline{s})}{\widetilde{u}(r_1,\overline{s})}. \tag{30}$$

Combining (29) and (30) gives $\phi(r_2|s) \geq \phi(r_1|s)$ for $s = \underline{s},\overline{s}$, $r_1, r_2 \in (r^*(\overline{s}), r^*(\underline{s}))$, and $r_2 \geq r_1$. Thus, ϕ is a distribution function. Since $\phi(r|s)$ is increasing in r, it is differentiable in r almost everywhere. Thus, taking the limits $r' \uparrow r$ and $r' \downarrow r$ in (28) and then integrating over $\widetilde{R}$ gives (7). Finally, the sender's expected utility coincides with (5) by integration by parts:

$$\begin{aligned}\sum_{s=\underline{s},\overline{s}} \Pr(s)\int_R \phi(r|s)\, g(r|s)\, dr &= \sum_{s=\underline{s},\overline{s}} \Pr(s)\left[-\phi(r|s)\,\overline{G}(r|s)\Big|_{\underline{r}}^{\overline{r}} + \int_R \overline{G}(r|s)\, d\phi(r|s)\right] \\ &= \sum_{s=\underline{s},\overline{s}} \Pr(s)\int_R \overline{G}(r|s)\, d\phi(r|s)\,.\end{aligned}$$

■

References

Aliprantis, Charalambos D. and Kim C. Border (2006) *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Berlin: Springer-Verlag.

Anderson, Edward J. and Peter Nash (1987) *Linear Programming in Infinite-Dimensional Spaces: Theory and Applications*, New York: John Wiley and Sons.

Arvey, Richard D. and James E. Campion (1982) "The Employment Interview: A Summary and Review of Recent Research," *Personnel Psychology*, **35** (2), pp. 281–322.

de Barreda, Ines Moreno (2012) "Cheap Talk with Two-Sided Private Information." Mimeo, University of Oxford.

Blackwell, David (1953) "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, **24** (2), pp. 265–272.

Chen, Ying (2009) "Communication with Two-sided Asymmetric Information." Mimeo, Arizona State University.

——— (2012) "Value of Public Information in Sender–Receiver Games," *Economics Letters*, **114** (3), pp. 343–345.

Crawford, Vincent P. and Joel Sobel (1982) "Strategic Information Transmission," *Econometrica*, **50** (6), pp. 1431–1451.

Golosov, Mikhail, Vasiliki Skreta, Aleh Tsyvinski, and Andrea Wilson (2013) "Dynamic Strategic Information Transmission." Mimeo, NYU.

Green, Jerry R. and Nancy L. Stokey (2007) "A Two-Person Game of Information Transmission," *Journal of Economic Theory*, **135** (1), pp. 90–104.

Ivanov, Maxim (2010) "Informational Control and Organization Design," *Journal of Economic Theory*, **145** (2), pp. 721–751.

Kamenica, Emir and Matthew Gentzkow (2011) "Bayesian Persuasion," *American Economic Review*, **101** (6), pp. 2590–2615.

Kolmogorov, Andrey N. and Sergey V. Fomin (1975) *Introductory Real Analysis*, New York: Dover Publications.

Kolotilin, Anton (2013) "Experimental Design to Persuade." Mimeo, UNSW.

Lai, Ernest K. (2013) "Expert Advice for Amateurs." Mimeo, Lehigh University.

Milgrom, Paul (1981) "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, **12** (2), pp. 380–391.

Myerson, Roger B. (1982) "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, **10** (1), pp. 67–81.

Ostrovsky, Michael and Michael Schwarz (2010) "Information Disclosure and Unraveling in Matching Markets," *American Economic Journal: Microeconomics*, **2** (2), pp. 34–63.

Rayo, Luis and Ilya Segal (2010) "Optimal Information Disclosure," *Journal of Political Economy*, **118** (5), pp. 949–987.

Seidmann, Daniel J (1990) "Effective Cheap Talk with Conflicting Interests," *Journal of Economic Theory*, **50** (2), pp. 445–458.

Watson, Joel (1996) "Information Transmission when the Informed Party is Confused," *Games and Economic Behavior*, **12** (1), pp. 143–161.

Supplemental Appendix: Binary Case

When the receiver has private information, in general, the problem of finding the optimal mechanism becomes complicated, as Section 4 suggests. This appendix fully characterizes the optimal mechanism when s and r are binary. More formally, assume that F puts strictly positive probabilities only on s_1 and s_2 and that $G(.|s_1)$ and $G(.|s_2)$ put strictly positive probabilities only on r_1 and r_2.

The binary case splits into two subcases. In the first subcase, one sender's signal is more favorable for acting than the other, regardless of r. The analysis of this subcase is two-fold. First, it provides formal proofs for the motivating example of Section 3.1. Second, it shows that the quantity-quality tradeoff of the motivating example carries on to a more general setting. In the second subcase, different sender's signals are favorable for acting depending on r. In this subcase, the single-crossing assumption does not hold, so the analysis and results are very different from those in the paper.

Using the revelation principle, for any mechanism, we can find an equivalent mechanism that sends at most four messages: (i) $m_{\emptyset}$ that induces the receiver not to act for all r, (ii) m_1 that induces the receiver to act only if $r = r_1$, (iii) m_2 that induces the receiver to act only if $r = r_2$, and (iv) $m_{1,2}$ that induces the receiver to act for all r.

For notational simplicity, this appendix uses different notation. In particular, denote $p_j \equiv \Pr(s_j)$, $p_{i|j} \equiv \Pr(r_i|s_j)$, $u_{ij} \equiv u(r_i, s_j)$, $\widetilde{u}_{ij} \equiv u_{ij} p_{i|j}$, $k_i = \widetilde{u}_{i1} / (\widetilde{u}_{i1} - \widetilde{u}_{i2})$, and $\phi^j_K \equiv \Pr_\phi(m = m_K, s = s_j)$ for $i, j = 1, 2$ and $K = \{\emptyset\}, \{1\}, \{2\}, \{1, 2\}$. Indexes i and j are reserved for r and s, respectively. Note that k_i is the cutoff posterior belief $\Pr(s_2)$ at which the receiver r_i is indifferent to act because

$$\mathbb{E}[u|r_i] = \frac{\widetilde{u}_{i1}(1 - \Pr(s_2)) + \widetilde{u}_{i2}\Pr(s_2)}{\Pr(r_i)} = 0.$$

Aligned Preferences

If one sender's signal is more favorable for acting than the other for all r, then the analysis is analogous to that of the binary example in Section 3.1. In particular, the sender faces the quantity-quality tradeoff of messages, which is resolved by the choice of a mechanism that sends messages with the highest marginal utility-price ratio.

To make the analysis non-redundant, assume that $u_{i1} < 0 < u_{i2}$ for $i = 1, 2$, $k_2 < k_1$, and $p_2 < k_1$. Strict inequalities rule out non-generic cases. Inequalities $u_{i2} > u_{i1}$ and $k_2 < k_1$ can be obtained by relabelling elements of S and R, respectively. If u_{i1} and u_{i2} had the same sign for some i, then the receiver r_i would take the same action regardless of the mechanism and the analysis would be as if the receiver was uninformed (Kolotilin (2013)). Finally, if $p_2 \geq k_1$, the no revelation mechanism would induce the receiver to act for all r, and, thus, it would be optimal.

Under these assumptions, the optimal mechanism can take the three forms that were identified in Section 3, as follows from:

Proposition 6 *If $u_{i1} < 0 < u_{i2}$, $k_2 < k_1$, and $p_2 < k_1$, then the optimal mechanism sends two messages.*

1. *If $p_{1|2} + p_{2|1}\widetilde{u}_{22}/\widetilde{u}_{21} \geq \widetilde{u}_{12}/\widetilde{u}_{11}$, it sends $m_{1,2}$ and $m_{\emptyset}$: $m_{1,2}$ with certainty if $s = s_2$ and with a non-trivial probability if $s = s_1$.*

2. *If $p_{1|2} + p_{2|1}\widetilde{u}_{22}/\widetilde{u}_{21} < \widetilde{u}_{12}/\widetilde{u}_{11}$ and $p_2 < k_2$, it sends m_2 and $m_{\emptyset}$: m_2 with certainty if $s = s_2$ and with a non-trivial probability if $s = s_1$.*

3. *If $p_{1|2} + p_{2|1}\widetilde{u}_{22}/\widetilde{u}_{21} < \widetilde{u}_{12}/\widetilde{u}_{11}$ and $p_2 \geq k_2$, it sends m_2 and $m_{1,2}$: m_2 with a non-trivial probability both if $s = s_2$ and if $s = s_1$.*

In all cases, $m_{\emptyset}$ reveals s_1 in that $\Pr_{\phi}(s_2|m_{\emptyset}) = 0$; m_2 makes the receiver r_2 indifferent to act in that $\Pr_{\phi}(s_2|m_2) = k_2$; and $m_{1,2}$ makes the receiver r_1 indifferent to act in that $\Pr_{\phi}(s_2|m_{1,2}) = k_1$. The receiver's expected utility under the optimal mechanism is strictly greater than that under the no revelation mechanism only in case 1.[23]

[23] If s and r are independent, then $p_{i|2} = p_{i|1} = \Pr(r_i)$, so $\widetilde{u}_{ij}$ can be replaced with u_{ij} for all $i, j = 1, 2$ in all expressions.

The intuition for Proposition 6 is analogous to that of the binary example in Section 3.1. The receiver r_i acts upon receiving a message m under a mechanism ϕ if $\Pr_\phi(s_2|m) \geq k_i$. If the message m persuades the receiver r_1 to act, it also persuades the receiver r_2 to act because $k_2 < k_1$ by assumption. Thus, we can restrict attention to mechanisms with the three messages $m_\emptyset$, m_2, and $m_{1,2}$. To maximize the probability that the receiver acts, each message of the optimal mechanism either makes the receiver exactly indifferent to act for some r ($\Pr_\phi(s_2|m_{1,2}) = k_1$ and $\Pr_\phi(s_2|m_2) = k_2$) or makes the receiver certain that $s = s_1$ so that it is optimal not to act ($\Pr_\phi(s_2|m_\emptyset) = 0$).

Thus, the sender's problem is to maximize the probability that the receiver acts:

$$\left(k_2 p_{2|2} + (1-k_2)\, p_{2|1}\right) q_2 + q_{1,2}$$

over probabilities $q_\emptyset$, q_2, and $q_{1,2}$ of the messages $m_\emptyset$, m_2, and $m_{1,2}$ subject to the constraint imposed by the prior distribution of s:

$$k_2 q_2 + k_1 q_{1,2} = p_2.$$

Similar to Section 3, we can interpret k_2 and k_1 as unit prices of sending m_2 and $m_{1,2}$, and the probabilities $\left(k_2 p_{2|2} + (1-k_2)\, p_{2|1}\right)$ and 1 as the marginal utilities of sending m_2 and $m_{1,2}$. If $p_{1|2} + p_{2|1}\widetilde{u}_{22}/\widetilde{u}_{21} \geq \widetilde{u}_{12}/\widetilde{u}_{11}$, then the marginal utility-price ratio is highest for $m_{1,2}$, and the sender prefers to send $m_{1,2}$ than m_2, so the optimal mechanism sends $m_{1,2}$ and $m_\emptyset$. If $p_{1|2} + p_{2|1}\widetilde{u}_{22}/\widetilde{u}_{21} < \widetilde{u}_{12}/\widetilde{u}_{11}$, then the ratio is highest for m_2, and the sender prefers to send m_2 than $m_{1,2}$. The optimal mechanism then depends on whether the no revelation mechanism induces the receiver r_2 to act or not. If so ($p_2 \geq k_2$), then it sends the messages m_2 and $m_{1,2}$, otherwise it sends the messages m_2 and $m_\emptyset$.

Proof of Proposition 6. The optimal mechanism ϕ maximizes

$$\Pr{}_\phi(a=1) = p_{2|1}\phi_2^1 + p_{2|2}\phi_2^2 + \phi_{1,2}^1 + \phi_{1,2}^2$$

subject to

$$\begin{aligned}
&\phi_K^j \geq 0 \text{ for } j = 1,2 \text{ and } K = \{\emptyset\}, \{2\}, \{1,2\}, \\
&\phi_\emptyset^j + \phi_2^j + \phi_{1,2}^j = p_j \text{ for } j = 1,2, \\
&\widetilde{u}_{21}\phi_2^1 + \widetilde{u}_{22}\phi_2^2 \geq 0, \\
&\widetilde{u}_{11}\phi_{1,2}^1 + \widetilde{u}_{12}\phi_{1,2}^2 \geq 0, \\
&\widetilde{u}_{21}\phi_\emptyset^1 + \widetilde{u}_{22}\phi_\emptyset^2 < 0 \text{ or } \phi_\emptyset^1 = \phi_\emptyset^2 = 0, \\
&\widetilde{u}_{11}\phi_2^1 + \widetilde{u}_{12}\phi_2^2 < 0 \text{ or } \phi_2^1 = \phi_2^2 = 0.
\end{aligned} \tag{31}$$

Consider the relaxed problem that omits the last two constraints with strict inequalities. The solution to the relaxed problem satisfies $\phi_\emptyset^2 = 0$, $\widetilde{u}_{21}\phi_2^1 + \widetilde{u}_{22}\phi_2^2 = 0$, and $\widetilde{u}_{11}\phi_{1,2}^1 + \widetilde{u}_{12}\phi_{1,2}^2 =$

0, otherwise we can increase $\Pr_\phi(a=1)$ by the following changes to the mechanism. If $\phi_\emptyset^2 \neq 0$, change $\widetilde{\phi}_{1,2}^2 = \phi_{1,2}^2 + \phi_\emptyset^2$ and $\widetilde{\phi}_\emptyset^2 = 0$; if $\widetilde{u}_{11}\phi_{1,2}^1 + \widetilde{u}_{12}\phi_{1,2}^2 > 0$, change $\widetilde{\phi}_{1,2}^1 = \phi_{1,2}^1 + \varepsilon$ and either $\widetilde{\phi}_2^1 = \phi_2^1 - \varepsilon$ or $\widetilde{\phi}_\emptyset^1 = \phi_\emptyset^1 - \varepsilon$; if $\widetilde{u}_{21}\phi_2^1 + \widetilde{u}_{22}\phi_2^2 > 0$, change $\widetilde{\phi}_{1,2}^2 = \phi_{1,2}^2 + \varepsilon$ and $\widetilde{\phi}_2^2 = \phi_2^2 - \varepsilon$ where ε is a small positive number. These observations together with $k_2 < k_1$ imply that the solution to the relaxed problem satisfies the last two constraints and, therefore, it also solves the original problem. The original problem simplifies to the maximization of

$$\Pr{}_\phi(a=1) = \left(1 - \frac{\widetilde{u}_{12}}{\widetilde{u}_{11}}\right) p_2 - \left(p_{1|2} + \frac{\widetilde{u}_{22}}{\widetilde{u}_{21}} p_{2|1} - \frac{\widetilde{u}_{12}}{\widetilde{u}_{11}}\right) \phi_2^2$$

over ϕ_2^2 subject to

$$\begin{aligned} &\left(\tfrac{\widetilde{u}_{12}}{\widetilde{u}_{11}} - \tfrac{\widetilde{u}_{22}}{\widetilde{u}_{21}}\right) \phi_2^2 \leq p_1 + \tfrac{\widetilde{u}_{12}}{\widetilde{u}_{11}} p_2. \\ &0 \leq \phi_2^2 \leq p_2. \end{aligned}$$

The solution to this problem is:

$$\phi_2^2 = \begin{cases} 0 \text{ if } p_{1|2} + \frac{\widetilde{u}_{22}}{\widetilde{u}_{21}} p_{2|1} \geq \frac{\widetilde{u}_{12}}{\widetilde{u}_{11}}; \\ p_2 \text{ if } p_{1|2} + \frac{\widetilde{u}_{22}}{\widetilde{u}_{21}} p_{2|1} < \frac{\widetilde{u}_{12}}{\widetilde{u}_{11}} \text{ and } \widetilde{u}_{21}p_1 + \widetilde{u}_{22}p_2 < 0; \\ \widetilde{u}_{21} \frac{\widetilde{u}_{11}p_1 + \widetilde{u}_{12}p_2}{\widetilde{u}_{12}\widetilde{u}_{21} - \widetilde{u}_{11}\widetilde{u}_{22}} \text{ if } p_{1|2} + \frac{\widetilde{u}_{22}}{\widetilde{u}_{21}} p_{2|1} < \frac{\widetilde{u}_{12}}{\widetilde{u}_{11}} \text{ and } \widetilde{u}_{21}p_1 + \widetilde{u}_{22}p_2 \geq 0. \end{cases}$$

Finally, $\phi_{1,2}^2 = p_2 - \phi_2^2$, $\phi_2^1 = -\phi_2^2\widetilde{u}_{22}/\widetilde{u}_{21}$, $\phi_{1,2}^1 = -\phi_{1,2}^2\widetilde{u}_{12}/\widetilde{u}_{11}$, $\phi_\emptyset^2 = 0$, $\phi_\emptyset^1 = p_1 - \phi_2^1 - \phi_{1,2}^1$.

Under ϕ_{no}, the receiver's expected utility is

$$\mathbb{E}\left[\max_a \mathbb{E}_{\phi_{no}}\left[au(r,s)\,|r\right]\right] = \max\left\{0, \widetilde{u}_{21}p_1 + \widetilde{u}_{22}p_2\right\}.$$

Under ϕ, the receiver's expected utility is

$$\begin{aligned} \mathbb{E}\left[\max_a \mathbb{E}_\phi\left[au(r,s)\,|r,m\right]\right] &= \left(\widetilde{u}_{11}\phi_{1,2}^1 + \widetilde{u}_{12}\phi_{1,2}^2\right) + \left(\widetilde{u}_{21}\phi_{1,2}^1 + \widetilde{u}_{22}\phi_{1,2}^2\right) + \left(\widetilde{u}_{21}\phi_2^1 + \widetilde{u}_{22}\phi_2^2\right) \\ &= \widetilde{u}_{21}\phi_{1,2}^1 + \widetilde{u}_{22}\phi_{1,2}^2 \\ &= \begin{cases} \left(\frac{\widetilde{u}_{11}\widetilde{u}_{22} - \widetilde{u}_{12}\widetilde{u}_{21}}{\widetilde{u}_{11}}\right) p_2 \text{ if } p_{1|2} + \frac{\widetilde{u}_{22}}{\widetilde{u}_{21}} p_{2|1} \geq \frac{\widetilde{u}_{12}}{\widetilde{u}_{11}}; \\ 0 \text{ if } p_{1|2} + \frac{\widetilde{u}_{22}}{\widetilde{u}_{21}} p_{2|1} < \frac{\widetilde{u}_{12}}{\widetilde{u}_{11}} \text{ and } \widetilde{u}_{21}p_1 + \widetilde{u}_{22}p_2 < 0; \\ \widetilde{u}_{21}p_1 + \widetilde{u}_{22}p_2 \text{ if } p_{1|2} + \frac{\widetilde{u}_{22}}{\widetilde{u}_{21}} p_{2|1} < \frac{\widetilde{u}_{12}}{\widetilde{u}_{11}} \text{ and } \widetilde{u}_{21}p_1 + \widetilde{u}_{22}p_2 \geq 0. \end{cases} \end{aligned}$$

The second equality holds because $\widetilde{u}_{11}\phi_{1,2}^1 + \widetilde{u}_{12}\phi_{1,2}^2 = \widetilde{u}_{21}\phi_2^1 + \widetilde{u}_{22}\phi_2^2 = 0$. The first case holds because $\phi_{1,2}^2 = p_2$ and $\phi_{1,2}^1 = -p_2\widetilde{u}_{12}/\widetilde{u}_{11}$. The second case holds because $\phi_{1,2}^1 = \phi_{1,2}^2 = 0$. The third case holds because $\phi_{1,2}^1 = p_1 - \phi_2^1$, $\phi_{1,2}^2 = p_2 - \phi_2^2$, and $\widetilde{u}_{21}\phi_2^1 + \widetilde{u}_{22}\phi_2^2 = 0$. Therefore, the receiver's expected utilities under ϕ and ϕ_{no} differ if and only if $p_{1|2} + p_{2|1}\widetilde{u}_{22}/\widetilde{u}_{21} \geq \widetilde{u}_{12}/\widetilde{u}_{11}$.

■

Misaligned Preferences

The main goal of this section is to illustrate the variety of possible optimal mechanisms in the case where different sender's signals are more favorable for acting depending on the receiver's type. For example, a school may know whether a student is good at natural sciences or liberal arts, but it may be unsure which of these two qualities are valued by the employer. Note that this case violates the single-crossing assumption of Section 4.

All forms that the optimal mechanism can take are characterized by Proposition 7. Similar to the previous subcase, to make the analysis non-redundant, I impose certain assumptions.

Proposition 7 *If $u_{12} < 0 < u_{11}$, $u_{21} < 0 < u_{22}$, and $p_2 > k_1$, then the optimal mechanism sends at most two messages.*

1. *If $k_2 \leq k_1$, it sends m_2 and $m_{1,2}$. The message m_2 reveals s_2 in that $\Pr_\phi(s_2|m_2) = 1$ and the message $m_{1,2}$ makes the receiver r_1 indifferent to act in that $\Pr_\phi(s_2|m_{1,2}) = k_1$.*

2. *If $k_2 > k_1$, then depending on parameters, it sends either only m_2 or both m_2 and m_1. If it sends both m_2 and m_1, there are four cases in which each message m_i either reveals s_i in that $\Pr_\phi(s_i|m_i) = 1$, or makes the receiver r_i indifferent to act in that $\Pr_\phi(s_2|m_i) = k_i$.*

I only sketch the intuition for this proposition because it is tedious and involves many cases. Note that a message m that assigns a higher probability to s_2 is more persuasive for the receiver r_2, and less persuasive for the receiver r_1. The messages m_1 and m_2 are always feasible because revealing s_1 induces the receiver r_1 to act, and revealing s_2 induces the receiver r_2 to act. However, if $k_2 \leq k_1$ (part 1 of Proposition 7), then the message $m_{1,2}$ is feasible, but the message $m_\emptyset$ is not. In this case, the sender wants to send $m_{1,2}$ as often as possible. As a result, the optimal mechanism sends two types of messages: those that give minimal possible evidence to make the receiver act regardless of his signal, and those that reveal s. In contrast, if $k_1 < k_2$ (part 2 of Proposition 7), then the message $m_\emptyset$ is feasible, but the message $m_{1,2}$ is not. In this case, the optimal mechanism can take five different forms, which, in particular, include the full revelation and no revelation mechanisms.

Proof of Proposition 7. The optimal mechanism ϕ maximizes

$$\Pr{}_\phi(a=1) = p_{1|1}\phi_1^1 + p_{1|2}\phi_1^2 + p_{2|1}\phi_2^1 + p_{2|2}\phi_2^2 + \phi_{1,2}^1 + \phi_{1,2}^2$$

subject to

$$\begin{aligned}
&\phi_K^j \geq 0 \text{ for } j = 1, 2 \text{ and } K = \{\emptyset\}, \{1\}, \{2\}, \{1,2\}, \\
&\phi_\emptyset^j + \phi_1^j + \phi_2^j + \phi_{1,2}^j = p_j \text{ for } j = 1, 2, \\
&\widetilde{u}_{i1}\phi_i^1 + \widetilde{u}_{i2}\phi_i^2 \geq 0 \text{ for } i = 1, 2, \\
&\widetilde{u}_{i1}\phi_{1,2}^1 + \widetilde{u}_{i2}\phi_{1,2}^2 \geq 0 \text{ for } i = 1, 2, \\
&\widetilde{u}_{i1}\phi_{3-i}^1 + \widetilde{u}_{i2}\phi_{3-i}^2 < 0 \text{ or } \phi_{3-i}^1 = \phi_{3-i}^2 = 0 \text{ for } i = 1, 2, \\
&\widetilde{u}_{i1}\phi_\emptyset^1 + \widetilde{u}_{i2}\phi_\emptyset^2 < 0 \text{ or } \phi_\emptyset^1 = \phi_\emptyset^2 = 0 \text{ for } i = 1, 2.
\end{aligned}$$

Note that $\widetilde{u}_{11} \Pr_\phi (s_1|m) + \widetilde{u}_{12} \Pr_\phi (s_2|m) \geq 0$ is equivalent to $\Pr_\phi (s_2|m) \leq k_1$, and $\widetilde{u}_{21} \Pr_\phi (s_1|m) + \widetilde{u}_{22} \Pr_\phi (s_2|m) \geq 0$ is equivalent to $\Pr_\phi (s_2|m) \geq k_2$. Therefore, the receiver r_1 acts if $\Pr_\phi (s_2|m) \leq k_1$, and the receiver r_2 acts if $\Pr_\phi (s_2|m) \geq k_2$. If $k_2 \leq k_1$, then no mechanism can send the message $m_\emptyset$ because $\Pr_\phi (s_2|m) < k_2$ and $\Pr_\phi (s_2|m) > k_1$ cannot both hold. On the contrary, if $k_2 > k_1$, then no mechanism can send the message $m_{1,2}$ because $\Pr_\phi (s_2|m) \geq k_2$ and $\Pr_\phi (s_2|m) \leq k_1$ cannot both hold. Consider these two cases in turn.

Let $k_2 \leq k_1$ and, thus, $\phi_\emptyset^1 = \phi_\emptyset^2 = 0$. Consider the relaxed problem with the constraints $\phi_K^j \geq 0$, $\phi_1^j + \phi_2^j + \phi_{1,2}^j = p_j$, $\widetilde{u}_{11}\phi_1^1 + \widetilde{u}_{12}\phi_1^2 \geq 0$, and $\widetilde{u}_{11}\phi_{1,2}^1 + \widetilde{u}_{12}\phi_{1,2}^2 \geq 0$ for all K and j. Note that the last two constraints imply $\widetilde{u}_{11}\left(\phi_1^1 + \phi_{1,2}^1\right) + \widetilde{u}_{12}\left(\phi_1^2 + \phi_{1,2}^2\right) \geq 0$, so the solution to the relaxed problem satisfies $\phi_1^1 = \phi_1^2 = 0$, otherwise we can increase $\Pr_\phi (a = 1)$ by the following changes to the mechanism: $\widetilde{\phi}_{1,2}^j = \phi_{1,2}^j + \phi_1^j$ and $\widetilde{\phi}_1^j = 0$ for $j = 1, 2$. Substituting $\phi_{1,2}^j = p_j - \phi_2^j$, the relaxed problem simplifies to: ϕ_2^1 and ϕ_2^2 maximize

$$\Pr{}_\phi (a = 1) = 1 - p_{1|1}\phi_2^1 - p_{1|2}\phi_2^2$$

subject to

$$\begin{aligned}
&\phi_2^j \in [0, p_j] \text{ for } j = 1, 2, \\
&\widetilde{u}_{11}\phi_2^1 + \widetilde{u}_{12}\phi_2^2 \leq \widetilde{u}_{11}p_1 + \widetilde{u}_{12}p_2.
\end{aligned}$$

The solution to this problem is $\left(\phi_2^1, \phi_2^2\right) = \left(0, \left(\widetilde{u}_{11}p_1 + \widetilde{u}_{12}p_2\right)/\widetilde{u}_{12}\right)$. It is also the solution to the original problem because it satisfies all constraints of the original problem.

Let $k_2 > k_1$ and, thus, $\phi_{1,2}^1 = \phi_{1,2}^2 = 0$. In the optimal mechanism, $\phi_\emptyset^1 = \phi_\emptyset^2 = 0$, otherwise we can increase $\Pr_\phi (a = 1)$ by the following changes to the mechanism: $\widetilde{\phi}_i^i = \phi_i^i + \phi_\emptyset^i$ and $\widetilde{\phi}_\emptyset^i = 0$ for $i = 1, 2$. Consider the relaxed problem with the constraints $\phi_1^j, \phi_2^j \geq 0$, $\phi_1^j + \phi_2^j = p_j$, $\widetilde{u}_{11}\phi_1^1 + \widetilde{u}_{12}\phi_1^2 \geq 0$, and $\widetilde{u}_{21}\phi_2^1 + \widetilde{u}_{22}\phi_2^2 \geq 0$ for all $j = 1, 2$. Substituting $\phi_1^j = p_j - \phi_2^j$, the relaxed problem simplifies to: ϕ_2^1 and ϕ_2^2 maximize

$$\Pr{}_\phi (a = 1) = p_{1|1}p_1 + p_{1|2}p_2 + \left(1 - 2p_{1|1}\right)\phi_2^1 + \left(1 - 2p_{1|2}\right)\phi_2^2$$

subject to

$$\begin{aligned}
&\phi_2^j \in [0, p_j] \ \text{ for } j = 1, 2, \\
&\widetilde{u}_{11}\phi_2^1 + \widetilde{u}_{12}\phi_2^2 \leq \widetilde{u}_{11}p_1 + \widetilde{u}_{12}p_2, \\
&\widetilde{u}_{21}\phi_2^1 + \widetilde{u}_{22}\phi_2^2 \geq 0.
\end{aligned}$$

The coefficients $1 - 2p_{1|1}$ and $1 - 2p_{1|2}$ in the objective function can have any sign and, therefore, any extreme point of the constraints can be a solution to this problem. If $p_2 \geq k_2$, the extreme points of $\left(\phi_2^1, \phi_2^2\right)$ are $(0, p_2)$, $(0, (\widetilde{u}_{11}p_1 + \widetilde{u}_{12}p_2) / \widetilde{u}_{12})$, and (p_1, p_2). If $p_2 < k_2$, the extreme points of $\left(\phi_2^1, \phi_2^2\right)$ are $(0, p_2)$, $(0, (\widetilde{u}_{11}p_1 + \widetilde{u}_{12}p_2) / \widetilde{u}_{12})$, $(-p_2\widetilde{u}_{22}/\widetilde{u}_{21}, p_2)$, $(\widetilde{u}_{22}, -\widetilde{u}_{21}) \cdot (\widetilde{u}_{11}p_1 + \widetilde{u}_{12}p_2) / (\widetilde{u}_{11}\widetilde{u}_{22} - \widetilde{u}_{12}\widetilde{u}_{21})$. All these extreme points can be a solution to the original problem because they satisfy all constraints of the original problem. ■