Australian School of Business
Working Paper

Never Stand Still
Australian School of Business

Australian School of Business Research Paper No. 2013 ECON 17

Experimental Design to Persuade

Anton Kolotilin

Experimental Design to Persuade*

Anton Kolotilin†

This Version: May, 2013

Abstract

A sender chooses ex ante how information will be disclosed ex post. A receiver obtains public information and information disclosed by the sender. Then he takes one of two actions. The sender wishes to maximize the probability that the receiver takes the desired action. I show that the sender optimally discloses only whether the receiver's utility is above a cutoff. I derive necessary and sufficient conditions for the sender's and receiver's welfare to be monotonic in information. Most notably, the sender's welfare increases with the precision of the sender's potential information and decreases with the precision of public information.

JEL Classification: C44, D81, D82, D83

Keywords: information, disclosure, persuasion, stochastic orders

1 Introduction

Economists have long been interested in how an interested party can communicate her private information to a decision maker when their interests are imperfectly aligned (seminal contributions include Spence (1973), Milgrom (1981), and Crawford and Sobel (1982)). I

*This paper is a revised version of Chapter 1 of my dissertation at MIT. I thank Robert Gibbons and Muhamet Yildiz for their invaluable guidance and advice. I also thank Ricardo Alonso, Abhijit Banerjee, Gabriel Carroll, Glenn Ellison, Péter Eső, Gabriele Gratton, Richard Holden, Bengt Holmström, Sergei Izmalkov, Uliana Loginova, Niko Matouschek, Parag Pathak, Michael Powell, Andrea Prat, Jean Tirole, Juuso Toikka, Alexander Wolitzky, Juan Xandri, Luis Zermeño, seminar participants at NES, MIT, UNSW, as well as the participants at the 7$^{\text{th}}$ Pan-Pacific Conference on Game Theory for helpful comments and suggestions.

†University of New South Wales, School of Economics. Email: a.kolotilin@unsw.edu.au.

study a situation in which before obtaining information, the interested party can choose a mechanism that specifies what information will be disclosed to the decision maker. This situation has been largely unexplored until recently (the pioneering articles are Rayo and Segal (2010) and Kamenica and Gentzkow (2011)).

The drug approval process by the Food and Drug Administration (FDA) is a good example of such a situation. If a pharmaceutical company (manufacturer) wants a new drug to be approved, it has to submit a research protocol for all tests that are going to be undertaken. The FDA closely monitors the record keeping and the adherence to the research protocol. So the FDA essentially observes both the design and results of all tests. Finally, based on the results of these tests, the FDA either approves the drug or rejects it. Because of the large cost of the process and large benefits of approval, the manufacturer has strong incentives to optimally design tests to maximize the probability of the FDA's approval.[1] What is the optimal design of tests? How much and what types of information these tests should reveal? What determines the success rate of drug trials and what determines the average quality of approved drugs?

I give exhaustive answers to these important questions by considering the following sender-receiver game. The receiver has a binary action choice: 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 beliefs. He forms his beliefs based on public information and information disclosed by the sender. The sender chooses ex ante how information will be disclosed to the receiver ex post. Formally, she can publicly choose any conditional distribution of messages given information. I call this distribution a *mechanism*. The sender chooses the mechanism that maximizes her expected utility – the ex ante probability that the receiver will act. No monetary transfers between the sender and receiver are allowed.

This model is a special case of Kamenica and Gentzkow (2011) who consider a general model with an arbitrary set of actions, and arbitrary utility functions for the sender and receiver. They derive some interesting properties of the optimal mechanism. To completely characterize the optimal mechanism, I impose more structure that still fits many real-life examples well. Most importantly, however, I derive general monotone comparative statics results that relate the sender's and receiver's expected utilities to the probability distribution of information. Specifically, I provide necessary and sufficient conditions for

[1] The description of the drug approval process is taken from Lipsky and Sharp (2001).

the sender and receiver to prefer one distribution of information to another for all values of the receiver's opportunity cost of acting. I now present the main results of the paper using the drug approval process.

The manufacturer optimally chooses a test that produces two outcomes: positive and negative. The former makes the FDA indifferent between approving and rejecting the drug, and the latter makes the FDA strictly prefer to reject the drug. These results follow directly from Kamenica and Gentzkow (2011). On top of these results, I show that the positive outcome occurs if and only if the drug's quality is above a cutoff.

What factors affect the manufacturer's welfare (or equivalently the probability of the drug approval)? The manufacturer's welfare is higher if the manufacturer is able to design more informative tests (in the mean-preserving spread sense) and if better drugs enter the testing phase (in the first-order stochastic dominance sense). Interestingly, under the absence of public information, these two conditions are not only sufficient but also necessary if the manufacturer's welfare is required to be higher for all values of the FDA's opportunity cost of approving the drug. Under the presence of public information, the manufacturer's welfare is higher if (and under some additional conditions, also only if) public information is less precise and more positive about the drug's quality.

What factors affect the FDA's welfare (or equivalently the expected quality of approved drugs)? Surprisingly, the FDA's welfare remains the same if the manufacturer is able to design more informative tests. However, the FDA's welfare is higher if public information is more precise and more positive about the drug's quality. These two conditions are also necessary if the FDA's welfare is required to be higher for all values of the opportunity cost of approving the drug. Finally, the overall welfare of the manufacturer and FDA is increasing in the precision of potential information of the manufacturer but is not monotonic in the precision of public information.

Although the above monotone comparative statics results are intuitive, they do not hold in the large existing literature where the sender chooses what information to disclose when she already has her private information. In particular, they do not hold under cheap talk and verifiable communication (Green and Stokey (2007) and Ivanov (2010)). The difference is due to the sender's incentive compatibility constraint on information disclosure, which is absent in my model, because the sender chooses what information to reveal at the ex ante stage.

Public information in the model captures not only information that will literally become public, such as the results of required tests, but also any verifiable private information of the manufacturer or the FDA that they have at the ex ante stage, such as the results of preclinical trials or rival applications previously submitted to the FDA. Indeed, using an argument similar to Milgrom (1981), I show that such information gets fully disclosed. The manufacturer's unverifiable information, such as private opinions of its experts, on the contrary, can be ignored because the manufacturer can never credibly transmit it.

The most related paper is Kamenica and Gentzkow (2011) discussed above. Rayo and Segal (2010) and Kolotilin (2012) study optimal information disclosure when the receiver has unverifiable private information. Lerner and Tirole (2006), Brocas and Carrillo (2007), and Benoit and Dubra (2011) study information disclosure in environments similar to mine, but in their models, the sender is exogenously constrained in choosing a mechanism, so they do not characterize the optimal mechanism. Bergemann and Pesendorfer (2007) and Eso and Szentes (2007) study optimal information disclosure in environments in which monetary transfers are allowed. Finally, Athey and Levin (2001) derive monotone comparative statics results in certain single-person decision problems.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 completely characterizes the optimal information disclosure mechanism and presents monotone comparative statics results. Section 4 extends the model to allow the sender and receiver to have private information at the ex ante stage. Section 5 concludes. All proofs and technical details are relegated to the appendices.

2 Model

Consider a communication game between a female sender and a male receiver. The receiver takes a binary action $a = 0, 1$. Say that the receiver *acts* if he takes $a = 1$, and the receiver *does not act* if he takes $a = 0$. The sender's utility depends only on a, but the receiver's utility depends both on a and on (s, r), where components s and r denote the sender's type and public type, respectively. Without loss of generality, the sender's utility is a, and the receiver's utility is u_0 if $a = 0$ and s if $a = 1$.[2] Before (s, r) is realized, the sender can commit to a mechanism that sends a message m to the receiver as a (stochastic) function

[2]Defining the sender's preferred action as $a = 1$ and applying an affine transformation gives that his utility is a. Suppose now that the receiver's utility is $u_0(s, r)$ if $a = 0$ and $u_1(s, r)$ if $a = 1$. Because the action is binary, only the difference $u_1(s, r) - u_0(s, r)$ matters for the receiver's choice of action, so $u_0(s, r)$ can be normalized to u_0 (or even to 0). Further, for any given r, which is observed both by the sender and the receiver, the sender's type can be transformed according to $u_1(., r)$ to ensure that the receiver's utility from acting is s.

of (s, r); specifically, the sender chooses the conditional distribution $\phi(m|s, r)$ of m given (s, r).

Assume that the set of messages M contains at least two elements m_0 and m_1, the set of sender's types S is $[\underline{s}, \overline{s}]$, the set of public types R is an arbitrary set that satisfies mild regularity conditions that ensure that all conditional expectations exist.[3] The information (s, r) has some joint distribution. For simplicity, assume that all distributions admit strictly positive densities unless stated otherwise. In particular, the marginal distribution $G(r)$ of r and the conditional distribution $F(s|r)$ of s given r admit strictly positive densities $g(r)$ and $f(s|r)$.

The timing of the communication game is as follows:

1. The sender publicly chooses a mechanism $\phi(m|s, r)$.
2. A triple (m, s, r) 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 beliefs and acts if and only if the conditional expectation $\mathbb{E}_\phi[s|m, r]$ of s given (m, r) is at least u_0. (Note that PBE requires that the receiver takes the sender's preferred action whenever he is indifferent between the two actions.) At the first stage, the sender chooses an *optimal mechanism* that maximizes her expected utility, the probability that the receiver acts.

Using the revelation principle, restrict attention to mechanisms that send only two messages: m_0 that persuades the receiver not to act and m_1 that persuades the receiver to act. Adopt the convention that $\phi(m_1|s, r)$ denotes the probability of the message m_1 given (s, r). Hereafter, all notions are in the weak sense. For example, increasing means not decreasing and higher means not lower.

To see that my model is a good approximation of the drug approval process, let us reinterpret the manufacturer as the sender and the FDA as the receiver. The FDA's approval decision is the receiver's action, and the research protocol is the sender's choice of a mechanism. Any information that can potentially be revealed by some tests is the

[3]For example, R is allowed to be a complete separable metric space endowed with the Borel sigma algebra (Theorems 1.4.12 and 4.1.6 in Durrett (1996)).

sender's information, the results of the required tests is public information, and the results of the remaining tests is a message. The manufacturer has a lot of freedom in choosing the design of tests. For example, it chooses dosage and characteristics of volunteer patients, such as gender, age, and health condition. Moreover, the manufacturer can make specifics of subsequent tests to be contingent on the results of the required tests. Due to the FDA's regulation and close monitoring, the FDA observes both the design and results of all tests, and then approves the drug if its benefits outweigh its costs and risks.

3 Analysis

I start this section by deriving the optimal mechanism, which has a simple cutoff structure. This result makes the model suitable for sharp comparative statics analysis. I first illustrate the main comparative statics results under the absence of public information. It is not trivial to generalize the results when the public information is present because each public type generates a different distribution of the sender's type.[4] At the end of the section, I present this generalization and discuss its practical importance using the drug approval process.

3.1 Optimal Mechanism

The optimal mechanism ϕ^* has a simple cutoff structure.

Theorem 1 *The optimal mechanism is given by*

$$\phi^* (m_1|s, r) = \begin{cases} 1 \text{ if } s \geq s^* (r), \\ 0 \text{ if } s < s^* (r). \end{cases} \quad (1)$$

If $\int_{\underline{s}}^{\overline{s}} sf(s|r)\, ds \geq u_0$, *then* $s^*(r) = \underline{s}$; *otherwise* $s^*(r) < u_0$ *is the unique solution to* $\int_{s^*(r)}^{\overline{s}} (s - u_0) f(s|r)\, ds = 0$.

Clearly, the optimal mechanism is conditioned on each piece of public information r. This implies that it does not matter whether the sender commits to a mechanism before or after the realization of r. I give the intuition for Theorem 1 conditional on some value r. If it is not possible to induce the receiver to always act, then the optimal mechanism induces the receiver to act if and only if his utility is above the cutoff. The cutoff is such

[4]Therefore, to compare two information structures one needs to compare two distributions of distributions of the sender's type instead of simply comparing two distributions of the sender's type as under the absense of public information.

that the receiver is indifferent between the two actions whenever he acts. Intuitively, the optimal mechanism has two defining features: (i) it makes the receiver indifferent between the two actions whenever he acts; and (ii) it makes the receiver know whether his utility is above the cutoff. If the first feature were violated, then the receiver would strongly prefer to act whenever he acts. Thus, it would be possible to increase the probability that the receiver acts by sending m_1 for a slightly larger set of types s. If the second feature were violated, then it would be possible to construct a mechanism that sends m_1 with the same total probability, but for higher types s. This mechanism would violate the first feature, so it would be possible to increase the probability that the receiver acts.

Theorem 1 and subsequent results extend when the distribution of (s, r) does not admit a density, as I show in Appendix C. The only difference is that the optimal mechanism may randomize over messages at the cutoff as the following example shows. Suppose that $u_0 = 0$, public information is absent, and F is a discrete distribution that assigns probabilities $1/3$ and $2/3$ to 1 and -1. The optimal mechanism sends the message m_1 if $s = 1$, and the messages m_1 and m_0 with equal probabilities if $s = -1$. As a result, the receiver who gets m_1 is indifferent between the two actions and the probability of m_1 is $2/3$.

Weaker versions of Theorem 1 appear in the literature. Lerner and Tirole (2006) show that the mechanism from Theorem 1 is optimal in a smaller class of feasible mechanisms in a more specific setting than mine. Kamenica and Gentzkow (2011) establish Theorem 1 for the above discrete example. For a more general setting than mine, they derive interesting properties of the optimal mechanism. In particular, these properties imply that m_1 makes the receiver indifferent between the two actions and that m_0 can only be sent to types $s < u_0$. However, they do not imply that that the optimal mechanism has a cutoff structure in that m_0 is sent if and only if $s < s^*(r)$. Moreover, my proof is simpler.

3.2 Comparative Statics without Public Information

In this section, assume the absence of public information. Theorem 2 presents monotone comparative statics results that relate the sender's and receiver's expected utilities under the optimal mechanism to the distribution of the sender's type. This theorem uses the standard definitions from the literature on stochastic orders. Let P_1 and P_2 be two distributions. P_2 *is higher than* P_1 *in the increasing convex order* if there exists a distribution P such that P_2 first-order stochastically dominates P and P is a mean-preserving spread of P_1.[5]

[5] See Definition 1 and Lemma 1 in Appendix A for more definitions and results on stochastic orders.

Theorem 2 *Let F_1 and F_2 be two distributions of s that do not depend on r.*

1. *The sender's expected utility under the optimal mechanism is higher under F_2 than under F_1 for all u_0 if and only if F_2 is higher than F_1 in the increasing convex order.*

2. *The receiver's expected utility under the optimal mechanism is higher under F_2 than under F_1 for all u_0 if and only if $\mathbb{E}_{F_2}[s] \geq \mathbb{E}_{F_1}[s]$.*

Part 2 holds because the optimal mechanism is as uninformative as possible from the receiver's perspective, as follows from Theorem 1. Indeed, under the optimal mechanism, if the receiver acts, then he either holds the prior belief or is indifferent between the two actions. Thus, the receiver's expected utility under the optimal mechanism is $\max\{\mathbb{E}[s], u_0\}$, which is equal to his expected utility under a mechanism that sends the same message regardless of s.

Part 1 is more interesting. It says that the sender's expected utility is higher if the distribution of s is (i) more favorable for acting (in the first-order stochastic dominance sense) and (ii) more variable (in the mean-preserving spread sense). Condition (i) is straightforward: more favorable distribution makes it easier for the sender to persuade the receiver. As to condition (ii), shifting probability weights to the ends of the support of $[\underline{s}, \overline{s}]$ decreases $\mathbb{E}[s|s < F^{-1}(p^*)]$ and increases $\mathbb{E}[s|s \geq F^{-1}(p^*)]$ allowing the sender to increase the probability $1-p^*$ that the receiver acts, as follows from Theorem 1. Interestingly, these two conditions are not only sufficient but also necessary if the sender's expected utility is required to be higher for any value of u_0.[6]

To get a deeper understanding of the results involving condition (ii) above, notice that the model has the following equivalent interpretation. There is an underlying binary state ω. The receiver's utility is ω if he acts (and u_0 if he does not). The sender's type s is a noisy signal about ω normalized to $\mathbb{E}[\omega|s]$. The sender chooses a mechanism $\phi(m|s)$, which determines how much information about s is disclosed to the receiver. Indeed, let ω take values $\underline{\omega} = \underline{s}$ and $\overline{\omega} = \overline{s}$ with probabilities $\underline{p} = (\overline{s} - \mathbb{E}[s]) / (\overline{s} - \underline{s})$ and $\overline{p} = 1 - \underline{p}$; let the density functions of s given ω be $h(s|\underline{\omega}) = f(s)(\overline{s} - s) / (\overline{s} - \mathbb{E}[s])$ and $h(s|\overline{\omega}) = f(s)(s - \underline{s}) / (\mathbb{E}[s] - \underline{s})$. For this construction, $\underline{p}h(s|\underline{\omega}) + \overline{p}h(s|\overline{\omega}) = f(s)$ and $\mathbb{E}[\omega|s] = s$, which establishes the equivalence.

Under this interpretation, variability condition (ii) corresponds to an informativeness of potential information s (Blackwell (1953)). The following corollary of Theorem 2 presents

[6] As can be seen from the proof, it is straightforward to write a strong version of Theorem 2 in which the sender's and receiver's expected utilities are strictly higher under F_2.

comparative statics with respect to such changes in informativeness.

Corollary 1 *Let the receiver's utility from acting* ω *be either* $\underline{\omega}$ *or* $\overline{\omega}$ *with probabilities* $\underline{p}$ *and* $\overline{p}$. *Let* $H_1(s|\omega)$ *and* $H_2(s|\omega)$ *be two conditional distributions of* s *given* ω *that admit densities* $h_1(s|\omega)$ *and* $h_2(s|\omega)$.

1. *The sender's expected utility under the optimal mechanism is higher under* H_2 *than under* H_1 *for all* u_0 *if and only if there exists a distribution* $Q(s_1|s_2)$ *of* s_1 *given* s_2 *such that for all* ω *and* s_1,
$$H_1(s_1|\omega) = \int Q(s_1|s_2)\, h_2(s_2|\omega)\, ds_2. \tag{2}$$

2. *The receiver's expected utility under the optimal mechanism is the same under* H_2 *and under* H_1 *for all* u_0.

Part 2 holds again because the optimal mechanism leaves no rent to the receiver, so the receiver's expected utility $\max\{\mathbb{E}[\omega], u_0\}$ does not depend on potential information s. In part 1, the distribution H_1 is a garbling of the distribution H_2 in that H_1 is obtained from H_2 by adding noise. Thus, any mechanism $\phi_1(m|s_1)$ under H_1 can be replicated under H_2 by $\phi_2(m|s_2) = \int \phi_1(m|s_1)\, dQ(s_1|s_2)$, which implies that the sender's expected utility is higher under a more informative distribution H_2. Based on this intuition, the comparative statics results can be extended beyond this model as long as the sender can choose any mechanism at the ex ante stage. This assumption, however, is critical for the results.

Under a cheap talk version of my model, the sender would not be able to disclose any information because she always prefers the receiver to act. Thus, the sender's expected utility would not change as her information becomes more precise. More generally, Green and Stokey (2007) and Ivanov (2010) show that the sender's expected utility may strictly decrease in the precision of her information. This happens because having less precise information may reduce the sender's incentive to misrepresent information.

Similarly, under a verifiable communication version of my model, the sender would disclose all her information by the unravelling argument due to Milgrom (1981). Thus, by Theorem 1, it is optimal for the sender to know only whether the receiver's utility is above the cutoff, which is less informative than knowing the receiver's utility exactly. That is, the sender's utility may strictly decrease as her information becomes more precise.

In the drug approval process, the manufacturer's welfare is the success rate of drug trials and the FDA's welfare is the average quality of approved drugs. Note that these variables can actually be observed in the data, so the theoretical comparative statics results are amenable to empirical analysis. Others being equal, the success rate is higher if the manufacturer has a better drug discovery process, or if the manufacturer has better testing capabilities. The average quality of approved drugs is also higher under the first scenario, but is the same under the second scenario.

3.3 Comparative Statics with Public Information

This section generalizes the previous section by introducing public information. Notice that each piece of public information r is associated with a distinct conditional distribution $F(.|r)$ of the sender's type s. For convenience, I identify r with $F(.|r)$ so that the only primitive of the model is a distribution G of r. Thus, to compare two environments, we need to compare two distributions G of r, where r is multidimensional because it is a distribution of s.

To avoid technical issues that arise due to multidimensionality, I start with the case of binary s. Specifically, assume that $s = \underline{s}, \overline{s}$ and $r = \Pr(\overline{s}|r)$. Proposition 1 presents monotone comparative statics results that relate the sender's and receiver's expected utilities to the primitive G. This theorem uses a new stochastic order. *P_2 is higher than P_1 in the increasing concave order* if there exists P such that P_2 first-order stochastically dominates P and P_1 is a mean-preserving spread of P.

Proposition 1 *Let the support of $F(.|r)$ consist of $\underline{s}$ and $\overline{s}$ where r is identified with* $\Pr(\overline{s}|r)$ *for all $r \in R = [0,1]$. Let G_1 and G_2 be two distributions of r.*

1. *The sender's expected utility under the optimal mechanism is higher under G_2 than under G_1 for all u_0 if and only if G_2 is higher than G_1 in the increasing concave order.*

2. *The receiver's expected utility under the optimal mechanism is higher under G_2 than under G_1 for all u_0 if and only if G_2 is higher than G_1 in the increasing convex order.*

Part 1 of Proposition 1 states that the sender's expected utility increases as the distribution of public information becomes (i) more favorable for acting (in the first-order stochastic dominance sense) and (ii) less variable (in the mean-preserving spread sense). The intuition for condition (i) is again straightforward: more favorable public information makes it easier for the sender to persuade the receiver. As to condition (ii), when public information is less polarized, the receiver has a weaker opinion about his best action, so it is easier for the sender to influence him.

Part 2 of Proposition 1 states that the receiver's expected utility increases as the distribution G of public information becomes (i) more favorable for acting and (ii) more variable. As before the receiver has the same expected utility under the optimal mechanism and the mechanism ϕ_0 that sends the same message regardless of s. Since the receiver's utility from not acting is fixed at u_0, the receiver is better off as G becomes more favorable for acting by the revealed preference argument. As to condition (ii), the receiver with a stronger opinion enjoys a higher expected utility from his preferred action. Interestingly, conditions (i) and (ii) for both parts of Proposition 1 are not only necessary but also sufficient if the sender and receiver are required to be better off under G_2 for all values of u_0.

Mathematically, Proposition 1 is exhaustive because it gives tight comparative statics results with respect to the primitive G. But economically, changes in G in the first-order stochastic dominance sense are not meaningful if the prior distribution of s is fixed. Indeed, if the prior probability of $\overline{s}$ is fixed at $\overline{p}$ and r is a public signal about s (normalized as before to $\Pr(\overline{s}|r)$), then any feasible distribution G of r must satisfy $\int r dG(r) = \overline{p}$. Therefore, changes in G should be mean-preserving under this interpretation of public information.

Mean-preserving changes in G correspond to variability conditions (ii), which in turn correspond to informativeness of public information (Blackwell (1953)). The following corollary of Proposition 1 presents comparative statics with respect to such changes.

Corollary 2 *Let s take only two values $\underline{s}$ and $\overline{s}$ where $\Pr(\overline{s}) = \overline{p}$. Let H_1 and H_2 be two distributions of public signals r_1 and r_2 given s. The sender's (receiver's) expected utility under the optimal mechanism is higher (lower) under H_2 than under H_1 for all u_0 if and only if there exists a distribution $Q(r_2|r_1)$ of r_2 given r_1 such that for all s and r_2,*

$$H_2(r_2|s) = \int Q(r_2|r_1) h_1(r_1|s) dr_1. \tag{3}$$

The corollary states that it is easier for the sender to persuade a less informed receiver, but the receiver is better off with more precise public information.[7] Indeed, public information is less precise under H_2 because H_2 only adds noise to H_1. Intuitively, under H_2, the sender can replicate any mechanism ϕ_1 available under H_1 by using a mechanism that sends two messages sequentially. The first stage of this mechanism will then make public information more precise, from H_2 to H_1, and the second stage will implement ϕ_1. The receiver, in contrast, prefers more precise public information H_1 because she can take a better informed action under H_1.

[7] Again, in the cheap talk literature, the sender's and receiver's expected utilities are not monotonic in the precision of public information (Chen (2012)).

Theorem 3 shows that in the general case of continuous s, changes in the distribution G according to conditions (i) and (ii) have the same qualitative effects on the sender's and receiver's expected utilities as in Proposition 1. However, in this case, r being a distribution of s becomes infinite dimensional, so we need to extend the stochastic orders to the multidimensional case and impose a partial order on R (see Appendix D for details).

Theorem 3 *Let R be the set of distributions on $[\underline{s}, \overline{s}]$ endowed with an increasing convex order in that r_2 is higher than r_1 if $F(.|r_2)$ is higher than $F(.|r_1)$ in the increasing convex order. Let G_1 and G_2 be two distributions of r.*

1. *The sender's expected utility under the optimal mechanism is higher under G_2 than under G_1 for all u_0 if G_2 is higher than G_1 in the increasing concave order.*

2. *The receiver's expected utility under the optimal mechanism is higher under G_2 than under G_1 for all u_0 if G_2 is higher than G_1 in the increasing convex order.*

An important implication of Theorem 3 is that the sender becomes worse off and the receiver better off as public information becomes more precise. However, the social welfare does not necessarily increase with the precision of public information even if it puts a very small weight on the sender's utility. Indeed, suppose that initially public information is absent. As public information appears, the marginal increase in the receiver's expected utility is 0 by the Envelope Theorem, as noted by Radner and Stiglitz (1984), but the marginal decrease in the sender's expected utility is strictly positive.

Continuing the drug approval process example, any commonly known information at the time the manufacturer designs drug trials or any information that the FDA requires to be revealed during drug trials can be viewed as public information. The above results mean that requiring the manufacturer to run more tests, increases the average quality of approved drugs but decreases the success rate of drug trials. In the next section, I show that even the manufacturer's and FDA's private information can be viewed as public information in certain cases, which implies that the derived comparative statics results have broader applications.

4 Extensions

In this section, I show that any verifiable ex ante private information of the sender and receiver gets fully disclosed at no cost to the sender in the unique equilibrium. This result has practical importance. It implies that the monotone comparative statics results from the previous section continue to hold after reinterpreting any verifiable information as public information. In particular, applying these results to the drug approval process gives that the average quality of approved drugs increases and the success rate decreases if the manufacturer carries out more thorough preclinical trials or if the FDA has more precise private information about the tested drug from other sources.

4.1 Receiver's Verifiable Private Information

In this section, the receiver has verifiable private information at the ex ante stage. As usual, verifiable information is the information that cannot be lied about but can be concealed. In this case, the sender extracts the receiver's information at no cost and then discloses her information optimally as if the receiver's type was public. Therefore, all results of Section 3 apply.

To illustrate this result, assume that the type r is privately known by the receiver rather than publicly known. In other respects, the environment is the same as in Section 2. In particular, players, actions, the information structure, and preferences are the same. In addition, assume that the set of receiver's types R is given by $[\underline{r}, \overline{r}]$ and is ordered in such a way that $s^*(r)$ is strictly increasing in r where $s^*(r)$ is given by Theorem 1.

Similarly to Milgrom (1981), assume that the set of receiver's reports is $N(r) = [\underline{r}, r]$. That is, the receiver can report any type that is lower than his true type. Intuitively, the report n can be viewed as the receiver's claim that his true type r is at least n and the receiver's claims are required to be truthful in that r must belong to $[n, \overline{r}]$.

Now a mechanism ϕ sends a message m to the receiver as a (stochastic) function of (s, n). Finally, the timing of the game is as follows: 1. The sender publicly chooses a mechanism $\phi(m|s, n)$. 2. The receiver's type r is drawn according to G. 3. The receiver makes a report n. 4. A pair (m, s) is drawn according to ϕ and F. 5. The receiver gets a message m and takes an action a. 6. Utilities are realized.

Again, the solution concept used is PBE. Theorem 4 characterizes the unique PBE.

Theorem 4 *In the unique PBE, the receiver reports his true type $n = r$ and the sender chooses the optimal mechanism ϕ^* given by Theorem 1.*

The proof of existence of fully revealing equilibrium is fairly standard (Milgrom (1981)). To show the uniqueness, I construct a mechanism which is arbitrarily close to ϕ^* and which makes the receiver strictly prefer to disclose his information. This theorem shows that without loss of generality we can view the receiver's verifiable private information as public information.

Note that the mechanism ϕ^* and truthful reporting of the receiver constitutes a PBE even if the sender has partial commitment in that she can choose a mechanism only after the receiver's report. However, this PBE is not unique in this new model. For example, there exists a PBE in which the receiver always reports $n = 0$.[8]

4.2 Sender's Ex Ante Private Information

In this section, the sender has private information before she chooses a mechanism. As a result, the sender discloses all of her verifiable information and none of her unverifiable information. Thus, without loss of generality, the sender's verifiable information can be viewed as public information, and the sender's ex ante unverifiable information can be integrated out. Again, all results of Section 3 apply.

In this section, the type r is privately known by the sender rather than publicly known. In other respects, the environment is the same as in Section 2. In addition, assume that R is given by $[\underline{r}, \overline{r}]$ and is ordered in such a way that $s^*(r)$ is strictly decreasing in r.

The timing of the game is as follows: 1. The sender's type r is drawn according to G. 2. The sender makes a report n. 3. The sender publicly chooses a mechanism $\phi(m|s, n)$. 4. A pair (m, s) is drawn according to ϕ and F. 5. The receiver gets a report n and a message m and takes an action a. 6. Utilities are realized.

Again, the solution concept is PBE. The sender's information is verifiable if the set of her reports is $N(r) = [\underline{r}, r]$, where the report n can be viewed as the receiver's truthful claim that his type is at least n. The sender's information is unverifiable if the set of her reports is $N = [\underline{r}, \overline{r}]$ regardless of r.[9] Theorem 5 characterizes the unique PBE for both cases of verifiable and unverifiable information of the sender.

Theorem 5 *If the sender's ex ante private information is verifiable, then in the unique PBE, the sender reports $n = r$ and chooses the optimal mechanism ϕ^* given by Theorem 1.*

[8]Indeed, suppose that the sender believes that each out-of-equilibrium report $n \neq 0$ is made by the receiver with type $r = n$. Note that under such a belief, the sender chooses a mechanism $\phi^*(m|s, n)$ for any $n \neq 0$. Thus, the receiver's interim expected utility from reporting $n \neq 0$ is $\max\{u_0, \mathbb{E}[s|r]\}$, which is smaller than that from reporting $n = 0$.

[9]Relatedly, the reader may wonder whether the main results of the paper would change if the message m generated by the mechanism ϕ was privately observed by the sender. Gentzkow and Kamenica (2012) show that the optimal mechanism does not change and m is fully disclosed by the sender if m is verifiable. Obviously, the sender cannot disclose any information if m is unverifiable, so all mechanisms are equivalent.

If the sender's ex ante private information is unverifiable, then in the unique PBE, the sender reports some fixed n regardless of r and chooses the optimal mechanism ϕ^{**} *given by Theorem 1 where* $F(s|r)$ *is replaced with* $\int F(s|r)\,dG(r)$ *for all* r*.*

The result that the sender discloses all her verifiable private information is again in spirit of the unravelling result. The sender conceals all her unverifiable private information, because regardless of her information, she always wants to pretend that she has the best news for the receiver. Note that if the sender could commit to a mechanism before the realization of r, then by Theorem 1, the optimal mechanism would be ϕ^{**} defined in Theorem 5. That is, the full commitment optimum is achieved as the equilibrium outcome if the sender's information is unverifiable. This observation is consistent with Theorem 3, which shows that the sender's expected utility decreases with the precision of public information.

5 Conclusions

In this paper, I have studied optimal information disclosure mechanisms. I have imposed the following key assumptions. First, at the ex ante stage, the sender can publicly choose how her information will be disclosed ex post; specifically, she can choose any conditional distribution of messages given her information. Second, the receiver has a binary action choice. Third, the sender's utility depends on the receiver's action but does not depend on information.

The model is highly tractable and can be used as a building block. Compared to no disclosure, the optimal disclosure mechanism gives the same expected utility to the receiver but persuades him to act with a higher probability. Thus, in a richer model with many receivers, the receivers will gain or lose from the optimal information disclosure mechanism depending only on whether externalities from acting are positive or negative.

The monotone comparative statics results imply that it is straightforward to enrich the model with strategic decisions that affect the information structure. Returning to the drug approval process example, the manufacturer can increase the success rate of drug trials (i) by improving the discovery process such that better drugs enter the testing phase and (ii) by improving the testing phase such that a better design of drug trials can be chosen. The FDA, on the other hand, can improve the quality of approved drugs by imposing more required tests that the manufacturer must carry out, thereby obtaining more precise information about the tested drug.

I have also shown that the sender and receiver disclose all of their verifiable private information at the ex ante stage, so all of my results apply after reinterpreting this private information as public. However, in this paper, I have not explored the possibility of the receiver having private information that cannot be elicited by the sender ex ante. Generically, the receiver does have such private information at least by the time he takes an action. For example, the FDA carries out an independent review after receiving the application from the manufacturer. Moreover, the manufacturer is uncertain about preferences and beliefs of the FDA regarding the safety and efficacy of a new drug. Since the optimal mechanism leaves no rent to the receiver if the receiver is uninformed, as a trivial result, the optimal mechanism is (weakly) more informative if the receiver is privately informed. The detailed analysis of this situation is my central goal in Kolotilin (2012).

Appendix A: Stochastic Orders

Definition 1 presents the unidimensional stochastic orders used in Section 3.2.

Definition 1 *Let X_1 and X_2 be two random variables with distributions P_1 and P_2 on $[\underline{x}, \overline{x}]$. Say that*

1. *P_2 first-order stochastically dominates P_1 (denoted by $P_2 \geq_{st} P_1$) if $P_2(x) \leq P_1(x)$ for all x.*

2. *P_2 is a mean-preserving spread of P_1 (denoted by $P_2 \geq_{cx} P_1$) if there exist two random variables $\widehat{X}_2$ and $\widehat{X}_1$, defined on the same probability space, with distributions P_2 and P_1 such that $\mathbb{E}\left[\widehat{X}_2 | \widehat{X}_1\right] = \widehat{X}_1$.*

3. *P_2 is higher than P_1 in the increasing convex order (denoted by $P_2 \geq_{icx} P_1$) if there exists a distribution P such that $P_2 \geq_{st} P \geq_{cx} P_1$.*

4. *P_2 is higher than P_1 in the increasing concave order (denoted by $P_2 \geq_{icv} P_1$) if there exists a distribution P such that $P_2 \geq_{st} P$ and $P_1 \geq_{cx} P$.*

Lemma 1 gives useful equivalent representations of the above stochastic orders.

Lemma 1 *Let P_1 and P_2 be two distributions that admit densities on $[\underline{x}, \overline{x}]$.*

1. *$P_2 \geq_{st} P_1$ if and only if $\mathbb{E}[h(X_2)] \geq \mathbb{E}[h(X_1)]$ for all increasing functions h.*

2. $P_2 \geq_{cx} P_1$ *is equivalent to each of the following conditions:*

   (a) $\mathbb{E}[h(X_2)] \geq \mathbb{E}[h(X_1)]$ *for all convex functions* h*;*

   (b) $\int_x^{\overline{x}} P_2(\widetilde{x})\, d\widetilde{x} \leq \int_x^{\overline{x}} P_1(\widetilde{x})\, d\widetilde{x}$ *for all* $x \in [\underline{x}, \overline{x}]$ *with equality for* $x = \underline{x}$*;*

   (c) $\int_p^1 P_2^{-1}(\widetilde{p})\, d\widetilde{p} \geq \int_p^1 P_1^{-1}(\widetilde{p})\, d\widetilde{p}$ *for all* $p \in [0,1]$ *with equality for* $p = 0$*.*

3. $P_2 \geq_{icx} P_1$ *is equivalent to each of the following conditions:*

   (a) $\mathbb{E}[h(X_2)] \geq \mathbb{E}[h(X_1)]$ *for all increasing convex functions* h*;*

   (b) $\int_x^{\overline{x}} P_2(\widetilde{x})\, d\widetilde{x} \leq \int_x^{\overline{x}} P_1(\widetilde{x})\, d\widetilde{x}$ *for all* $x \in [\underline{x}, \overline{x}]$*;*

   (c) $\int_p^1 P_2^{-1}(\widetilde{p})\, d\widetilde{p} \geq \int_p^1 P_1^{-1}(\widetilde{p})\, d\widetilde{p}$ *for all* $p \in [0,1]$*.*

4. $P_2 \geq_{icv} P_1$ *is equivalent to each of the following conditions:*

   (a) $\mathbb{E}[h(X_2)] \geq \mathbb{E}[h(X_1)]$ *for all increasing convex functions* h*;*

   (b) $\int_{\underline{x}}^{x} P_2(\widetilde{x})\, d\widetilde{x} \leq \int_{\underline{x}}^{x} P_1(\widetilde{x})\, d\widetilde{x}$ *for all* $x \in [\underline{x}, \overline{x}]$*;*

   (c) $\int_0^p P_2^{-1}(\widetilde{p})\, d\widetilde{p} \geq \int_0^p P_1^{-1}(\widetilde{p})\, d\widetilde{p}$ *for all* $p \in [0,1]$*.*

Proof. See Shaked and Shanthikumar (2007) Section 1.A.1 for part 1, Section 3.A.1 for part 2, and Section 4.A.1 for parts 3 and 4. ■

Definition 2 presents the multidimensional stochastic orders used in Section 3.3.

Definition 2 *Let* $\mathcal{P}$ *be the set of distributions on* $[\underline{x}, \overline{x}]$ *endowed with some partial order* $\geq_P$*. Let* $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *be two random elements with distributions* Q_1 *and* Q_2 *on* $\mathcal{P}$*. Say that*

1. Q_2 *first-order stochastically dominates* Q_1 *(denoted by* $Q_2 \geq_{mst} Q_1$*) if* $\Pr_{Q_2}(\mathbf{X}_2 \in U) \geq \Pr_{Q_1}(\mathbf{X}_1 \in U)$ *for all measurable increasing sets* $U \subset \mathcal{P}$ *in that* $P \geq_P P'$ *and* $P' \in U$ *imply* $P \in U$*.*

2. Q_2 *is a mean-preserving spread of* Q_1 *(denoted by* $Q_2 \geq_{mcx} Q_1$*) if there exist two random elements* $\widehat{\mathbf{X}}_2$ *and* $\widehat{\mathbf{X}}_1$*, defined on the same probability space, with distributions* Q_2 *and* Q_1 *such that* $\mathbb{E}\left[\widehat{\mathbf{X}}_2 | \widehat{\mathbf{X}}_1\right] = \widehat{\mathbf{X}}_1$*.*

3. Q_2 *is higher than* Q_1 *in the increasing convex order (denoted by* $Q_2 \geq_{micx} Q_1$*) if there exists a distribution* Q *such that* $Q_2 \geq_{mst} Q \geq_{mcx} Q_1$*.*

4. Q_2 *is higher than* Q_1 *in the increasing concave order (denoted by* $Q_2 \geq_{micv} Q_1$*) if there exists a distribution* Q *such that* $Q_2 \geq_{mst} Q$ *and* $Q_1 \geq_{mcx} Q$*.*

Lemma 1 gives equivalent representations of the multidimensional stochastic orders.

Lemma 2 *Let* Q_1 *and* Q_2 *be two distributions on* $\mathcal{P}$.

1. $Q_2 \geq_{mst} Q_1$ *if and only if* $\mathbb{E}[h(\mathbf{X}_2)] \geq \mathbb{E}[h(\mathbf{X}_1)]$ *for all increasing functions* h *in that* $h(P_2) \geq h(P_1)$ *for all* $P_1, P_2 \in \mathcal{P}$ *such that* $P_2 \geq_P P_1$.

2. $Q_2 \geq_{mcx} Q_1$ *if and only if* $\mathbb{E}[h(\mathbf{X}_2)] \geq \mathbb{E}[h(\mathbf{X}_1)]$ *for all convex functions* h *in that* $h(\alpha P_1 + (1-\alpha) P_2) \leq \alpha h(P_1) + (1-\alpha) h(P_2)$ *for all* $P_1, P_2 \in \mathcal{P}$ *and all* $\alpha \in (0,1)$.

3. $Q_2 \geq_{micx} Q_1$ *if and only if* $\mathbb{E}[h(\mathbf{X}_2)] \geq \mathbb{E}[h(\mathbf{X}_1)]$ *for all increasing convex functions* h.

4. $Q_2 \geq_{micv} Q_1$ *if and only if* $\mathbb{E}[h(\mathbf{X}_2)] \geq \mathbb{E}[h(\mathbf{X}_1)]$ *for all increasing concave functions* h.

Proof. See Shaked and Shanthikumar (2007) Section 6.B.1 for part 1, and Section 7.A.1 for parts 2, 3, and 4. ■

Appendix B: Proofs

Proof of Theorem 1. The optimal mechanism ϕ^* solves

$$\underset{\phi(m_1|s,r)\in[0,1]}{\text{maximize}} \int_{S\times R} f(s|r)\, g(r)\, \phi(m_1|s,r)\, drds$$

subject to

$$\int_S (s-u_0)\, f(s|r)\, \phi(m_1|s,r)\, ds \geq 0 \text{ for all } r \in R$$

where the objective function is the probability that the receiver acts and the constraint requires that the receiver prefers to act whenever he receives m_1.

The Lagrangian for this problem is given by:

$$\mathcal{L} = \int_{S\times R} (1 + [s-u_0]\, \lambda(r))\, f(s|r)\, g(r)\, \phi(m_1|s,r)\, drds,$$

where $\lambda(r)\, g(r)$ is a multiplier for the constraint. Since the choice variable $\phi(m_1|s,r)$ belongs to the unit interval, we have $\phi(m_1|s,r) = 1$ if $s \geq u_0 - \frac{1}{\lambda(r)}$ and $\phi(m_1|s,r) = 0$ if $s < u_0 - \frac{1}{\lambda(r)}$ where $\lambda(r)$ is 0 if $\mathbb{E}_F[s|r] > u_0$ and is such that the constraint is binding if $\mathbb{E}_F[s|r] \leq u_0$. ■

Proof of Theorem 2. I start by proving the first part. Let s_i^* be given by Theorem 1 where F is replaced with F_i. If $F_2 \geq_{icx} F_1$ (see Definition 1), then the sender can induce the receiver to act with a higher probability under F_2 than under F_1 because

$$\begin{aligned}\int_{F_2^{-1}\left(F_1\left(s_1^*\right)\right)}^{\overline{s}} (s-u_0)\, dF_2(s) &= \int_{F_1\left(s_1^*\right)}^{1} \left(F_2^{-1}(\widetilde{p}) - u_0\right) d\widetilde{p} \\ &\geq \int_{F_1\left(s_1^*\right)}^{1} \left(F_1^{-1}(\widetilde{p}) - u_0\right) d\widetilde{p} \\ &= \int_{s_1^*}^{\overline{s}} (s-u_0)\, dF_1(s) \geq 0,\end{aligned}$$

where the equalities hold by the appropriate change of variables, the first inequality holds by Lemma 1 part 3 (c), and the last inequality holds by Theorem 1. Conversely, if $F_2 \not\geq_{icx} F_1$, then by Lemma 1 part 3 (c), there exists p such that $\int_p^1 F_2^{-1}(\widetilde{p})\, d\widetilde{p} < \int_p^1 F_1^{-1}(\widetilde{p})\, d\widetilde{p}$. Setting $u_0 = \int_{F_2^{-1}(p)}^1 s dF_2(s) / (1-p)$ and using an analogous argument, we get that the receiver acts with a strictly higher probability under F_1 than under F_2:

$$\begin{aligned}\int_{F_1^{-1}(p)}^{\overline{s}} (s-u_0)\, dF_1(s) &= \int_{p}^{1} \left(F_1^{-1}(\widetilde{p}) - u_0\right) d\widetilde{p} \\ &> \int_{p}^{1} \left(F_2^{-1}(\widetilde{p}) - u_0\right) d\widetilde{p} \\ &= \int_{F_2^{-1}(p)}^{\overline{s}} (s-u_0)\, dF_2(s) = 0.\end{aligned}$$

Now I prove the second part. The receiver's expected utility under F_i is $\max\{\mathbb{E}_{F_i}[s], u_0\}$ by Theorem 1. Clearly, if $\mathbb{E}_{F_2}[s] \geq \mathbb{E}_{F_1}[s]$, then $\max\{\mathbb{E}_{F_2}[s], u_0\} \geq \max\{\mathbb{E}_{F_1}[s], u_0\}$ for all u_0. Conversely, if $\mathbb{E}_{F_2}[s] < \mathbb{E}_{F_1}[s]$, then $\max\{\mathbb{E}_{F_2}[s], u_0\} < \max\{\mathbb{E}_{F_1}[s], u_0\}$ for any $u_0 \in (\mathbb{E}_{F_2}[s], \mathbb{E}_{F_1}[s])$. ■

Proof of Corollary 1. The distribution of the posterior $\Pr(\overline{\omega}|s)$ under H_2 is a mean-preserving spread of that under H_1 if and only if there exists Q such that (2) holds, as Blackwell (1953) shows. Since the posterior $\Pr(\overline{\omega}|s) = (s - \underline{\omega}) / (\overline{\omega} - \underline{\omega})$ is linear in s, the distribution of the posterior $\Pr(\overline{\omega}|s)$ under H_2 is a mean-preserving spread of that under H_1 if and only if $F_2 \geq_{cx} F_1$ where F_i is the distribution of s under H_i, given by $F_i(s) = H_i(s|\underline{\omega})\, \underline{q} + H_i(s|\overline{\omega})\, \overline{q}$ for $i = 1, 2$. Part 1 then follows by repeating all steps of the proof of Theorem 2 with the only difference that Lemma 1 part 2 (c) is used instead of Lemma 1 part 3 (c). Part 2 holds because the receiver's expected utility is $\max\{\mathbb{E}[\omega], u_0\}$ by Theorem 1, and $\mathbb{E}[\omega] = \underline{\omega q} + \overline{\omega q}$ does not depend on H. ■

Proof of Proposition 1. The sender's expected utility under the optimal mechanism is:

$$\begin{aligned}U_S &= \int_R \min\left\{\frac{\overline{s}-\underline{s}}{u_0-\underline{s}}\Pr\left(\overline{s}|r\right),1\right\}dG\left(r\right)\\ &= \int_0^1 \min\left\{\frac{\overline{s}-\underline{s}}{u_0-\underline{s}}r,1\right\}dG\left(r\right)\\ &= 1-\frac{\overline{s}-\underline{s}}{u_0-\underline{s}}\int_0^{\frac{u_0-\underline{s}}{\overline{s}-\underline{s}}} G\left(r\right)dr, \qquad (4)\end{aligned}$$

where the first equality holds by Theorem 6, the second by convention $\Pr\left(\overline{s}|r\right)=r$, and the third by integration by parts. Part 1 of the proposition follows immediately by Lemma 1 part 4 (b).

The receiver's expected utility under the optimal mechanism is:

$$\begin{aligned}U_R &= \int_R \max\left\{u_0,\mathbb{E}\left[s|r\right]\right\}dG\left(r\right)\\ &= \int_0^1 \max\left\{u_0,\underline{s}+\left(\overline{s}-\underline{s}\right)r\right\}dG\left(r\right)\\ &= \overline{s}-\left(\overline{s}-\underline{s}\right)\int_{\frac{u_0-\underline{s}}{\overline{s}-\underline{s}}}^1 G\left(r\right)dr, \qquad (5)\end{aligned}$$

where the first equality holds by Theorem 6, the second by convention $\Pr\left(\overline{s}|r\right)=r$, and the third by integration by parts. Part 2 of the proposition follows immediately by Lemma 1 part 3 (b). ■

Proof of Corollary 2. The distribution of the posterior $\Pr\left(\overline{s}|r\right)$ under H_1 is a mean-preserving spread of that under H_2 if and only if there exists Q such that (3) holds, as Blackwell (1953) shows. Applying Lemma 1 part 2 (b) to (4) and (5) proves the corollary.

■

Proof of Theorem 3. The probability that the receiver acts is $\int_R p^*\left(r\right)dG\left(r\right)$ where the conditional probability $p^*\left(r\right)$ that the receiver acts is given by $1-F\left(s^*\left(r\right)|r\right)$ with $s^*\left(r\right)$ given by Theorem 1. The function p^* is increasing in r in the increasing convex order by Theorem 2 part 1. Moreover, p^* is concave in r, as I show in the next paragraph. Therefore, part 1 of the theorem follows by Lemma 2 part 4.

For concavity of p^*, it suffices to show that there exists a mechanism ϕ that induces the receiver to act with probability $\alpha p^*\left(r_1\right)+\left(1-\alpha\right)p^*\left(r_2\right)$ when the distribution of s is $\alpha F\left(s|r_1\right)+\left(1-\alpha\right)F\left(s|r_2\right)$. Without loss of generality, suppose that $s^*\left(r_1\right)\geq s^*\left(r_2\right)$. The required mechanism is simply a mechanism that implements ϕ_1^* and ϕ_2^* with probabilities α and $1-\alpha$. Specifically, if $s\geq s^*\left(r_1\right)$, the receiver gets the message m_1. If $s<s^*\left(r_2\right)$, the

receiver gets the message m_0. Finally, if $s \in [s^*(r_2), s^*(r_1))$, the receiver gets the messages m_1 and m_0 with probabilities p_1 and $1 - p_1$ where

$$p_1 \equiv \frac{(1-\alpha)\left(F\left(s^*(r_1)|r_2\right) - F\left(s^*(r_2)|r_2\right)\right)}{\alpha\left(F\left(s^*(r_1)|r_1\right) - F\left(s^*(r_2)|r_1\right)\right) + (1-\alpha)\left(F\left(s^*(r_1)|r_2\right) - F\left(s^*(r_2)|r_2\right)\right)}.$$

The receiver's expected utility under the optimal mechanism is

$$U_R = \int_R \max\{u_0, \mathbb{E}[s|r]\}\, dG(r).$$

The function $\mathbb{E}[s|r]$ is linear in r because the expectation is linear in $F(.|r)$, which is identified with r. Moreover, $\mathbb{E}[s|r]$ is increasing in r by Lemma 1 part 3 (a) because $h(s) = s$ is increasing and convex in s and R is endowed with an increasing convex order. Thus, the function $\max\{u_0, \mathbb{E}[s|r]\}$ is increasing and convex in r. Part 2 of the theorem follows by Lemma 2 part 3. ■

Proof of Theorem 4. I start by showing that the described strategies constitute a PBE. If the receiver reports $n = r$, then his interim expected utility is $\max\{u_0, \mathbb{E}[s|r]\}$ as follows from Theorem 1. If the receiver reports $n < r$, then his interim expected utility is again $\max\{u_0, \mathbb{E}[s|r]\}$ because $s^*(r)$ is increasing in r. Thus, given the mechanism ϕ^*, it is a best response for the receiver to report his true type $n = r$. To see that it is optimal for the sender to choose ϕ^* at the first stage, note that ϕ^* is the optimal mechanism in the relaxed problem where r is publicly known, so ϕ^* gives a higher expected utility to the sender than any other feasible mechanism.

To complete the proof, I show that in all PBEs, the sender chooses ϕ^* and the receiver reports $n = r$. Suppose to get a contradiction that there exists another PBE. In this PBE, the sender's expected utility is strictly less than in the above PBE because ϕ^* is the optimal mechanism in the relaxed problem. Consider a mechanism $\widetilde{\phi}$ that sends the message m_1 if and only if $s \geq s^*(r) + \delta$ where $\delta > 0$ is sufficiently small. Under this mechanism, the receiver strictly prefers to report his true type r and the sender's expected utility is arbitrarily close to that under ϕ^*. A contradiction. ■

Proof of Theorem 5. Suppose that given the sender's report n, r is distributed according to G_n. Given this report, the receiver believes that s is distributed according to $F_n(s) = \int_R F(s|r)\, dG_n(r)$. By sequential rationality, at the third stage, the sender chooses the optimal mechanism ϕ_n^* that sends m_1 if and only if $s \geq s_n^*$ where s_n^* is given by Theorem 1 where $F(s|r)$ is replaced with $F_n(s)$.

I start by considering the case where the sender's information is verifiable. In this case, the sender r can make a report n only if $r \geq n$. Thus, the support of G_n does not intersect

$[\underline{r}, n)$. Suppose to get a contradiction that there exists an equilibrium report n such that H_n is supported on $[\underline{r}_n, \overline{r}_n]$ with $\overline{r}_n > n$. This means that with a strictly positive probability, the sender $\overline{r}_n$ makes the report n and induces the receiver to act with probability $1 - F(s_n^* | \overline{r}_n)$. If this sender made the report $\overline{r}_n$ instead, then she would induce the receiver to act with a strictly higher probability because $s_{\overline{r}_n}^* < s_n^*$, as I show in the next paragraph. Thus, G_n assigns probability one to $r = n$, meaning that the sender reports $n = r$ for all r.

The inequality $s_{\overline{r}_n}^* < s_n^*$ holds because $s_{\overline{r}_n}^* \leq s^*(\overline{r}_n) < s_n^*$. Suppose to get a contradiction that $s_n^* \leq s^*(\overline{r}_n)$, then

$$\int_{s_n^*}^{\overline{s}} (s - u_0)\, dF_n(s) = \int_{\underline{r}_n}^{\overline{r}_n} \left(\int_{s_n^*}^{\overline{s}} (s - u_0)\, dF(s|r) \right) dG_n(r) < 0,$$

contradicting the definition of s_n^*. The equality holds by Fubini's Theorem and the definition of $F_n(s)$. The inequality holds because

$$\int_{s_n^*}^{\overline{s}} (s - u_0)\, dF(s|r) < \int_{s^*(r)}^{\overline{s}} (s - u_0)\, dF(s|r) = 0$$

for all $r < \overline{r}_n$ as follows from $s_n^* < s^*(r) < u_0$ which is implied by the supposition $s_n^* \leq s^*(\overline{r}_n)$ and the assumption that $s^*(r)$ is strictly decreasing in r. Noting that the support of $G_{\overline{r}_n}$ does not intersect $[\underline{r}, \overline{r}_n)$ and using a similar argument gives $s_{\overline{r}_n}^* \leq s^*(\overline{r}_n)$.

Now, I consider the case where the sender's information is unverifiable. Suppose to get a contradiction that there exist two equilibrium reports n_1 and n_2 such that $s_{n_1}^* < s_{n_2}^*$. Then the sender would always prefer to report n_1 regardless of r. A contradiction. ■

Appendix C: Discontinuous Distributions

This appendix relaxes the assumption that all distributions are continuous. Instead, assume that $G(r)$ and $F(s|r)$ are arbitrary distributions whose supports are subsets of R and $S = [\underline{s}, \overline{s}]$. Theorem 6, a generalization of Theorem 1, characterizes the optimal mechanism.

Theorem 6 *The optimal mechanism is given by*

$$\phi^*(m_1|s,r) = \begin{cases} 1 \ \textit{if}\ s > s^*(r), \\ \pi^*(r) \ \ \textit{if}\ s = s^*(r), \\ 0 \ \textit{if}\ s < s^*(r). \end{cases} \tag{6}$$

If $\int_{\underline{s}}^{\overline{s}} s dF(s|r) \geq u_0$, *then* $s^*(r) = \underline{s}$ *and* $\pi^*(r) = 1$; *otherwise* $s^*(r) \leq 0$ *and* $q^*(r) \equiv \pi^*(r) \Pr(s = s^*(r)|r) \in [0, \Pr(s = s^*(r)|r)]$ *are the unique solution to*

$$\mathbb{E}_{\phi^*}[s - u_0|m_1] = \int_{(s^*(r),\overline{s}]} (s - u_0)\, dF(s|r) + (s^*(r) - u_0)\, q^*(r) = 0. \tag{7}$$

Proof. By Fubini's Theorem, the optimal mechanism ϕ^* solves

$$\underset{\phi(m_1|s,r)\in[0,1]}{\text{maximize}} \int_R \left(\int_S \phi(m_1|s,r)\, dF(s|r) \right) dG(r)$$

subject to

$$\int_S (s-u_0)\, \phi(m_1|s,r)\, dF(s|r) \geq 0 \text{ for all } r \in R.$$

We can see that the problem is separable; specifically, for each r, the optimal mechanism ϕ^* maximizes the inside integral subject to the constraint. Therefore, $\phi^*(m_1|s,r) = \widetilde{\phi}^*(m|s)$, where $\widetilde{\phi}^*(m|s)$ is the optimal mechanism in the model in which r is fixed, and the distribution of s is given by $F(s|r)$. Thus, we can omit r from consideration as if it was fixed.

Now I prove that if $\int_{\underline{s}}^{\overline{s}} s dF(s) < u_0$, then the optimal mechanism ϕ^* satisfies (6) where (s^*, π^*) solves (7). The remaining parts of Theorem 6 are immediate. Suppose to get a contradiction that there exists a mechanism $\widetilde{\phi}$ that results in a higher probability that the receiver acts: $\Pr_{\widetilde{\phi}}(m_1) > \Pr_{\phi^*}(m_1)$. In the next paragraph, I show that $F_{\phi^*}(s|m_1) \leq F_{\widetilde{\phi}}(s|m_1)$ for all $s \in [\underline{s}, \overline{s}]$ with strict inequality for $s \in [s^*, \overline{s})$, and, thus, $\mathbb{E}_{\phi^*}[s|m_1] > \mathbb{E}_{\widetilde{\phi}}[s|m_1]$ by the well-known result (a strong version of Lemma 1 part 1). Therefore, $\mathbb{E}_{\widetilde{\phi}}[s|m_1] < u_0$ because $\mathbb{E}_{\phi^*}[s|m_1] = u_0$ by (7). The conclusion that $\mathbb{E}_{\widetilde{\phi}}[s|m_1] < u_0$ contradicts the assumption that the message m_1 induces the receiver to act.

To complete the proof, I show that $F_{\phi^*}(s|m_1) \leq F_{\widetilde{\phi}}(s|m_1)$ for all $s \in [\underline{s}, \overline{s}]$ with strict inequality for $s \in [s^*, \overline{s})$. The inequality trivially holds for $s < s^*$ because $F_{\phi^*}(s|m_1) = 0$ and for $s = \overline{s}$ because $F_{\phi^*}(\overline{s}|m_1) = F_{\widetilde{\phi}}(\overline{s}|m_1) = 1$. Denote the joint distribution of m and s by $\phi(m,s)$. The following sequence of equalities and inequalities proves that $F_{\phi^*}(s|m_1) < F_{\widetilde{\phi}}(s|m_1)$ for $s \in [s^*, \overline{s})$:

$$\begin{aligned}
1 - F_{\phi^*}(s|m_1) &= \frac{\phi^*(m_1,\overline{s}) - \phi^*(m_1,s)}{\Pr_{\phi^*}(m_1)} \\
&= \frac{F(\overline{s}) - F(s)}{\Pr_{\phi^*}(m_1)} \\
&= \frac{\widetilde{\phi}(m_1,\overline{s}) - \widetilde{\phi}(m_1,s)}{\Pr_{\phi^*}(m_1)} + \frac{\widetilde{\phi}(m_0,\overline{s}) - \widetilde{\phi}(m_0,s)}{\Pr_{\phi^*}(m_1)} \\
&\geq \frac{\widetilde{\phi}(m_1,\overline{s}) - \widetilde{\phi}(m_1,s)}{\Pr_{\phi^*}(m_1)} \\
&> \frac{\widetilde{\phi}(m_1,\overline{s}) - \widetilde{\phi}(m_1,s)}{\Pr_{\widetilde{\phi}}(m_1)} \\
&= 1 - F_{\widetilde{\phi}}(s|m_1).
\end{aligned}$$

The first and last equalities hold by Bayes' rule. The second equality holds by (6), which defines $\phi^*(m, s)$. The third equality holds by the consistency condition: $\phi(m_1, s) + \phi(m_0, s) = F(s)$ for all mechanisms ϕ and all $s \in [\underline{s}, \overline{s}]$. The first inequality holds because $\phi(m_0, .)$ is a distribution function of s. The second inequality holds by the assumption that $\Pr_{\widetilde{\phi}}(m_1) > \Pr_{\phi^*}(m_1)$. ■

Theorem 2 holds regardless of whether F_1 and F_2 admit densities. The original proof of the second part of Theorem 2 applies to arbitrary F_1 and F_2. To prove the first part of Theorem 2, one should replace the inverse functions with the quantile functions in Lemma 1 part 3 (c) and in the original proof. Specifically, for an arbitrary distribution P, the quantile function is defined as $\varphi(p) \equiv \inf\{x : p \leq P(x)\}$. If $F_2 \geq_{icx} F_1$, then the receiver acts with a higher probability under F_2 than under F_1 because

$$\int_{F_1(s_1^*)-q_1^*}^{1} \varphi_2(\widetilde{p})\, d\widetilde{p} \geq \int_{F_1(s_1^*)-q_1^*}^{1} \varphi_1(\widetilde{p})\, d\widetilde{p} = \int_{(s_1^*, \overline{s}]} s dF_1(s) + s_1^* q_1^*.$$

Conversely, if $F_2 \ngeq_{icx} F_1$, there exists p such that $\int_p^1 \varphi_2(\widetilde{p})\, d\widetilde{p} < \int_p^1 \varphi_1(\widetilde{p})\, d\widetilde{p}$, so the receiver acts with a strictly higher probability under F_1 than under F_2 if $u_0 = \int_p^1 \varphi_2(\widetilde{p})\, d\widetilde{p}/(1-p)$. Using similar logic, it is straightforward to extend all results to the case of arbitrary distribution functions.

Appendix D: Discussion of Theorem 3

Theorem 3 extends Proposition 1 to the general case of continuous s. Since r is multidimensional in this case, the theorem relies on multidimensional stochastic orders presented in Appendix A (see Definition 2 and Lemma 2). For first-order stochastic dominance and any other stochastic order based on it, we need to introduce a partial order on R. In the case of binary s, the set R is the unit interval $[0, 1]$, a totally ordered set. But what order can we impose on R when R is the set of distributions on $[\underline{s}, \overline{s}]$? To answer this question, consider two degenerate distributions G_1 and G_2 that assign probability 1 to $r_1 = F_1$ and $r_2 = F_2$, respectively. Theorem 2 implies that the sender and receiver are better off under G_2 if $F_2 \geq_{icx} F_1$. To be able to compare such G_1 and G_2, Theorem 3 uses an increasing convex order as a partial order on R.

We lose necessity in Theorem 3 because an increasing convex order is not a total order when s takes more than two values. In part 2 of Theorem 3, we can actually use a total order on R and regain necessity. By Theorem 1, only the distribution of $\mathbb{E}[s|r]$ matters for the receiver. Identifying r with $\mathbb{E}[s|r]$, we obtain the following result. The receiver's expected utility under the optimal mechanism is higher under G_2 than under G_1 for all u_0 if and only if $G_2 \geq_{icx} G_1$, as follows from Lemma 1 part 3 (b) and

$$U_R = \int_R \max\{u_0, \mathbb{E}[s|r]\} \, dG(r) = \int_{\underline{r}}^{\overline{r}} \max\{u_0, r\} \, dG(r) = \overline{r} - \int_{u_0}^{\overline{r}} G(r) \, dr.$$

For the results concerning the sender's expected utility, we must use an increasing convex order on R because of Theorem 2 part 1. But the fact that the sender's expected utility is higher under G_2 than under G_1 for all u_0 does not imply that $G_2 \geq_{micv} G_1$, so necessity cannot be regained. To see this, consider the following counterexample. Let G_1 assign probabilities $\left(\frac{2}{3}, \frac{1}{3}, 0\right)$ to $\left(r^A, r^B, r^C\right)$, and G_2 assign probabilities $(0, 0, 1)$ to $\left(r^A, r^B, r^C\right)$ where r^A assigns probabilities $\left(r_1^A, r_2^A, r_3^A\right) = \left(0, \frac{1}{2}, \frac{1}{2}\right)$ to $(s_1, s_2, s_3) = \left(0, \frac{1}{2}, 1\right)$, r^B assigns probabilities $\left(0, \frac{7}{8}, \frac{1}{8}\right)$ to (s_1, s_2, s_3), and r^C assigns probabilities $\left(\frac{3}{8}, 0, \frac{5}{8}\right)$ to (s_1, s_2, s_3).

By Theorem 6, the receiver acts with probability $\min\left\{\frac{1}{4u_0 - 2}, 1\right\}$ under r^A, with probability $\min\left\{\frac{1}{16u_0 - 8}, 1\right\}$ under r^B, and with probability $\min\left\{\frac{5}{8u_0}, 1\right\}$ under r^C. By considering all cases ($u_0 \leq \frac{9}{16}$, $\frac{9}{16} < u_0 \leq \frac{5}{8}$, $\frac{5}{8} < u_0 \leq \frac{3}{4}$, $\frac{3}{4} < u_0 \leq 1$, and $u_0 > 1$), it is straightforward to check that the sender's expected utility is always higher under G_2 than under G_1.

By Lemma 1 part 3 (b), for any r and r' supported on (s_1, s_2, s_3), we have $r \geq_{icx} r'$ if and only if $r_3 \geq r_3'$ and $r_2 s_2 + r_3 s_3 \geq r_2' s_2 + r_3' s_3$. Thus, the function $h(r) = 10\left(\frac{r_2}{2} + r_3\right) + r_3$ is increasing in r in the increasing convex order. Moreover, h is concave in r because it is linear in r. However, the expectation of h is strictly higher under G_1 than under G_2, which implies that $G_2 \not\geq_{micv} G_1$.

Similarly to Proposition 1, using Blackwell (1953), we can obtain the following corollary of Theorem 3.

Corollary 3 *Let the prior density of s be given by f. Let h_1 and h_2 be two densities of public signals r_1 and r_2 given s. The sender's (receiver's) expected utility under the optimal mechanism is higher (lower) under h_2 than under h_1 for all u_0 if there exists a density $q(r_1|r_2)$ of r_1 given r_2 such that for all s and r_2,*

$$h_2(r_2|s) = \int q(r_2|r_1) \, h_1(r_1|s) \, dr_1.$$

References

Athey, Susan and Jonathan Levin (2001) "The Value of Information in Monotone Decision Problems." Mimeo, MIT.

Benoit, Jean-Pierre and Juan Dubra (2011) "Apparent Overconfidence," *Econometrica*, **79** (5), pp. 1591–1625.

Bergemann, Dirk and Martin Pesendorfer (2007) "Information Structures in Optimal Auctions," *Journal of Economic Theory*, **137** (1), pp. 580–609.

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

Brocas, Isabelle and Juan D. Carrillo (2007) "Influence through Ignorance," *RAND Journal of Economics*, **38** (4), pp. 931–947.

Chen, Ying (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.

Durrett, Richard (1996) *Probability: Theory and Examples*, Belmont, CA: Duxbury Press.

Eso, Peter and Balazs Szentes (2007) "Optimal information disclosure in auctions and the handicap auction," *Review of Economic Studies*, **74** (3), p. 705.

Gentzkow, Matthew and Emir Kamenica (2012) "Disclosure of Endogenous Information." Mimeo, University of Chicago.

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.

Kolotilin, Anton (2012) "Optimal Information Disclosure: Quantity vs. Quality." Mimeo, UNSW.

Lerner, Josh and Jean Tirole (2006) "A Model of Forum Shopping," *American Economic Review*, **96** (4), pp. 1091–1113.

Lipsky, Martin S. and Lisa K. Sharp (2001) "From Idea to Market: The Drug Approval Process," *Journal of the American Board of Family Medicine*, **14** (5), pp. 362–367.

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

Radner, Roy and Joseph E. Stiglitz (1984) "A Nonconcavity in the Value of Information," in M Boyer and R. Kihlstrom eds. *Bayesian Models in Economic Theory*, New York: North-Holland.

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

Shaked, Moshe and J. George Shanthikumar (2007) *Stochastic Orders*, New York: Springer.

Spence, Michael (1973) "Job Market Signaling," *Quarterly Journal of Economics*, **87** (3), pp. 355–374.