HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

Emiliano Catonini, Sergey Stepanov

ON THE OPTIMALITY OF FULL DISCLOSURE

Working Paper WP9/2023/02
Series WP9
Research of economics and finance

Moscow
2023

Catonini, E., Stepanov, S.

On the Optimality of Full Disclosure* [Electronic resource] : Working paper WP9/2023/02 / E. Catonini, S. Stepanov ; National Research University Higher School of Economics. – Electronic text data (295 Kb). – Moscow : HSE Publishing House, 2023. – (Series WP9 “Research of economics and finance”). – 44 p.

A privately informed sender can commit to any disclosure policy towards a receiver, whose actions affect the utility of the sender. We show that full disclosure is optimal under a sufficient condition with some desirable properties. First, it speaks directly to the utility functions of the parties, as opposed to the indirect utility function of the sender; this makes it easily interpretable and verifiable. Second, it does not require the sender's payoff to be a function of the posterior mean state. Third, it is weaker than the conditions on the sender's utility function that were recently provided for the case in which the receiver's reaction is linear in the expected state. With this, we show that full disclosure is optimal under some modeling assumptions commonly used in principal-agent papers.

JEL classification: D82, D83

Keywords: information design, Bayesian persuasion, full disclosure

Emiliano Catonini, NYU Shanghai; E-mail: ec1838@nyu.edu

Sergey Stepanov, HSE University, International College of Economics and Finance and Faculty of Economic Sciences, Russia. Postal address: Office S641, Pokrovsky Boulevard 11, 109028, Moscow, Russia; E-mail: sstepanov@hse.ru

*We thank Anton Kolotilin, Mikhail Drugov, and the participants of HSE Economic Theory workshop and EEA-ESEM 2022 congress, for helpful comments.

1 Introduction

We consider the classical problem of information transmission between a sender with private, payoff-relevant information and a receiver who takes actions which affect the sender's payoff. Following the Bayesian persuasion literature pioneered by Rayo and Segal (2010) and Kamenica and Gentzkow (2011), we suppose that the sender has commitment power over the information she reveals to the receiver. Without setting any restrictions on the possible persuasion strategies, we search for conditions under which full disclosure is optimal. Differently from other complicated schemes, just disclosing the truth seems to be a realistic goal in many scenarios — e.g., with transparency policies in organizations.

In our model, the state space can be a continuum, therefore the concavification approach of Kamenica and Gentzkow (2011) is not operational. Moreover, differently from most contributions in the field (e.g., Dworczak and Martini (2019), Dizdar and Kováč (2020), Gentzkow and Kamenica (2016), Kolotilin et al. (2021), Arieli et al. (2020)), we do not assume that the sender's payoff is a function of the posterior mean state (or any moments of the posterior distribution).[1] Despite this, we obtain a sufficient condition for the optimality of full disclosure that speaks directly to the underlying incentives of the parties, as opposed to the indirect utility function of the sender. This makes our condition easily interpretable and verifiable. In particular, it can be interpreted as a requirement of minimal alignment of incentives between the sender and the receiver. Notably, despite its level of generality, our condition is substantially weaker than the sufficient condition provided

[1]We will discuss three notable exceptions in detail.

by Kolotilin et al. (2022) for environments in which the receiver's optimal action is linear in the expected state.

To see why the effect of full disclosure may be non-trivial, consider a simple principal-agent setup, as an example. The agent generates an output, which he shares with the principal in a fixed proportion. The output is increasing in the agent's effort, and the agent bears the cost of effort. The state of nature determines the productivity of effort, with a higher state resulting in higher productivity. The principal knows the state, while the agent does not. At first sight, the principal would always want to commit to revealing the state to the agent, as both parties *seem* to benefit from effort more when the state is higher. Here is a simple argument why this may not be the case. Suppose that the agent is sufficiently risk averse. Then, good news about the productivity may actually depress effort. This is because a higher productivity implies that the agent reaches a higher income, hence a lower marginal utility, at lower levels of effort. If the principal is risk neutral, then the disclosure discourages the agent precisely when the principal benefits more from effort (and incentivizes the agent when the principal gains less from effort). In such a case, full disclosure is unlikely to be optimal. Note also that, even when this "income effect" does not prevail in the agent's incentives, full disclosure may still not be optimal. Even if the agent increases effort under the good news and reduces it under the bad news about the state, as the principal wants, the increase may be smaller than the decrease, to the point that the overall effect on the principal's utility is negative.

To see how we tackle these difficulties, stick to the principal-agent setup and consider a message that pools two equally likely states. The principal contemplates splitting this message into two messages that reveal the state. Then, given the optimal effort under the pooling message, the agent will discover that her marginal utility of effort is positive when one state is revealed, negative when the other state is revealed, and the two values have the same magnitude, just opposite signs. Thus, the agent will decrease effort under the first state and increase it under the second state. Two forces determine whether the principal gains from the split or not: the changes in the agent's effort and the changes in the principal's utility per unit of effort. Under each state, the agent modifies his effort until its marginal utility returns to zero. Then what matters is how much the principal's utility changes per unitary change of the agent's marginal utility. In particular, the principal benefits from the split if this measure of her marginal utility is larger when the agent wants to increase effort with respect to when he prefers to reduce effort. In this sense, ours is a condition of minimal alignment of interest between the two parties.

Our main result extends this argument to all possible messages in a general sender-receiver framework. Specifically, we show that *any* message with a non-singleton support can be split so as to improve the sender's welfare if an increase of action that decreases the receiver's marginal utility by one unit has a larger benefit for the sender when it also benefits the receiver, compared to when it harms him. This condition ensures the optimality of full disclosure.

Under some additional regularity assumptions, we also provide an analogous sufficient condition that is entirely expressed in terms of derivatives of the parties' utility functions ("derivatives condition"). This condition may be easier to check in some economic applications.

Finally, we also derive a sufficient condition for the *sub*optimality of full disclosure. While there remains a gap between this condition and our optimality condition (one is not a negation of the other), it helps to establish when full disclosure is definitely not optimal, as we will show in an example.

We then focus on the principal-agent setting we outlined before. Typically, in this application, the principal's utility cannot be represented as a function of the posterior mean. We discuss several examples demonstrating that our sufficient condition for full disclosure is easy to check and often satisfied. The first example (section 5.1) sheds light on the role of risk aversion for the optimality/suboptimality of full disclosure. We assume that both parties exhibit CRRA and the output is a product of the state and a concave power function of effort. Full disclosure turns out to be optimal when the agent is more risk averse than the principal (a typical textbook situation) but not too risk averse (with the coefficient of relative risk aversion below one). In this case, state and effort are complements for both parties, and then disclosing the state boosts effort exactly when the principal benefits from higher effort more. Instead, when the agent becomes too risk averse (while the principal remains moderately risk averse), full disclosure ceases to be optimal. As we discussed earlier, under high agent's risk aversion, good news about productivity *depress* effort, that is, effort and state become substitutes for the agent while remaining complements for the principal.

Another interesting case discussed in Section 5.1 is when the agent is sufficiently risk averse, and the principal is at least as risk averse as the agent. In that case, the average effort falls but the principal nevertheless gains from transparency. This happens because for the principal effort and state are even more substitutes than for the agent. Bad news about productivity encourages effort, and the principal benefits even more from effort in lower states than the agent does.

In the second example (section 5.2) we simplify the preferences by assuming risk neutrality for both parties and focus instead on the properties of the production function that ensure the optimality of full disclosure. By applying the "derivatives condition", we show that full disclosure is optimal under some commonly used functional forms for output.

Without assuming that the sender's payoff is a function of the expected state, Kolotilin (2018) and Kolotilin et al. (2022) establish that (under some assumptions on the utility functions) full disclosure is optimal if and only if, for any pair of states, the sender prefers revealing them to garbling.[2] So, under some conditions, the problem reduces to checking only messages with binary support. In some simple cases (for example, the receiver's optimal action depends only on the expected state and the sender's direct utility depends only on the action), the sender's indirect utility function becomes a function of only the posterior mean; then, the necessary and sufficient condition for the optimality of full disclosure boils down to requiring the convexity of this function. Kolotilin et al. (2022) make further progress by providing sufficient conditions for the optimality of full disclosure on the sender's (direct) utility function for the special case in which the receiver's optimal action is linear in the expected state (while the sender's utility is allowed to depend on the state as well as the action).

Except for the requirement that the receiver's utility is strictly concave in action and delivers an interior solution (along with some regularity assump-

[2]In more rigorous terms, the sender prefers to split any posterior with a binary support into two degenerate posteriors.

tions), we impose no restrictions on how state and action affect utilities. Despite this, we offer a sufficient condition for the optimality of full disclosure in terms of the primitives of the model: the (direct) utility functions of the sender and the receiver. In this way, compared to the condition in Kolotilin (2018) and Kolotilin et al. (2022), we gain operability and interpretability. Moreover, in contrast to Kolotilin (2018) and Kolotilin et al. (2022), we do not impose a single-crossing assumption on the receiver's utility. This allows applying our condition to environments where considering only binary support messages may not be without loss of generality. Despite this level of generality, our condition turns out to be substantially weaker than the sufficient condition of Kolotilin et al. (2022) for the case in which the receiver's action is linear in the expected case, as it requires neither convexity of the sender's payoff in action, nor its supermodularity in action and state.

Using a concept analogous to the concept of "virtual value" in the mechanism design literature, Mensch (2021) offers conditions for full disclosure jointly on the receiver's utility function and on a transformation of the sender's utility function that takes into account the incentive compatibility constraint of the receiver ("virtual utility"). His focus is on the importance of complementarities between states and actions, and whether these complementarities "point in the same direction" for the sender and the receiver. While Mensch's condition for full disclosure (Theorem 5) is insightful, it is rather abstract and not straightforward to apply, as it requires a derivation of the "virtual utility". Instead, our conditions are directly on primitives of the model, that is, the shape of the parties' utility functions.

The paper is organized as follows. Section 2 sets up the model. Section 3 derives the conditions for the optimality of full disclosure, as well as the condition for its suboptimality. In Section 4, we compare our sufficient condition with the conditions obtained in the literature for two special cases of the parties' preferences. Section 5 demonstrates how our conditions can be applied in a principal-agent setting and discusses the role of risk aversion and complementarity/substitutability between the action and the state. All proofs are relegated to the Appendix.

2 Model

There are a sender (she) and a receiver (he). The receiver needs to take a non-contractible action $a \in A$. There is a state of the world $\omega \in \Omega$ with common prior $p \in \Delta(\Omega)$. We assume that A and Ω are compact intervals in the real line; with this, we do not rule out that the possible states be discrete, because we do not impose restrictions on the support of p.

Action and state jointly determine the receiver's utility $U(\omega, a)$.[3] We assume that, for every $\omega \in \Omega$, $U(\omega, a)$ is twice differentiable and strictly concave in a, with $U_a(\omega, a) = 0$ for some finite $a \in A$, denoted by $a^*(\omega)$. We also assume that $U_a(\omega, a)$ is continuous in ω.

The sender's utility is $V(\omega, a)$, and we assume it to be differentiable in a, with $V_a(a, \omega)$ continuous in ω. Until Section 5, we abstract away from the origin of $U(\omega, a)$ and $V(\omega, a)$.

Before learning the state, the sender can commit to an information struc-

[3]The state may just represent whatever information is available to the sender about a "more primitive" state that affects payoffs; in this case U is an expected utility.

ture, whereby the receiver gets some information about the state before choosing the action. Formally, following the standard Bayesian persuasion framework, the sender commits to a mapping from the set of states Ω to distributions over messages that are sent to the receiver. The information structure chosen by the sender is common knowledge. The goal of the sender is to select an information structure that maximizes her expected utility.

After receiving message m, the receiver solves

$$\max_a \mathbb{E}(U(\omega, a)|m)$$

Due to our assumptions on $U(\omega, a)$, the receiver's optimal action is unique under every posterior belief about the state, and it is determined by the first-order condition

$$\frac{d\mathbb{E}(U(\omega, a)|m)}{da} = 0$$

By continuity of $U_a(\omega, a)$ in ω, the receiver's optimal action changes continuously in the posterior belief. With this, the persuasion problem of the sender is well-defined and has a solution.

3 Sufficient conditions

3.1 Main condition

Full disclosure is optimal if any message that pools or partially pools several states that induce different actions can be split into several more informative (in Blackwell sense) messages in a way that strictly increases the sender's expected utility (conditional on the original message). We will first consider messages that generate a posterior with binary support. The crucial and most insightful passage of our construction identifies a condition under which splitting a message with binary support into two messages that reveal the state weakly benefits the sender — we illustrate this passage in detail in the main text (and report a more formal proof in the Appendix). Then, we will sketch how we extend this argument to find a *strictly* profitable split of *any* message m under a slightly stronger condition, and finally establish the optimality of full disclosure under the original condition — the details of these two passages are deferred to the formal proof in the Appendix.

Consider two states, ω_1 and ω_2, such that the receiver's optimal action is higher under ω_2: $a_2^* := a^*(\omega_2) > a^*(\omega_1) =: a_1^*$. Let m be a message that (partially) pools ω_1 and ω_2, and let $\pi_1 := \Pr(\omega_1|m)$, $\pi_2 := \Pr(\omega_2|m)$, $\pi_2 = 1 - \pi_1$. The graph below depicts the receiver's utilities under ω_1, ω_2, and his expected utility under m: $U(\omega_1, a)$, $U(\omega_2, a)$, $\widetilde{U}(\omega, a)$. Action a^* denotes the receiver's optimal action under m. The sender's state-contingent utilities $V(\omega_1, a)$ and $V(\omega_2, a)$ are depicted increasing, with $V(\omega_2, a)$ above $V(\omega_1, a)$, for illustration purposes, but they do not have to be such.

Figure 1.

Conditional on m, the sender (weakly) benefits from disclosing $\omega_1,\ \omega_2$ instead of sending m if and only if

$$\pi_1 V(\omega_1, a_1^*) + \pi_2 V(\omega_2, a_2^*) \geq \pi_1 V(\omega_1, a^*) + \pi_2 V(\omega_2, a^*),$$

that is,

$$\pi_2[V(\omega_2, a_2^*) - V(\omega_2, a^*)] \geq \pi_1[V(\omega_1, a^*) - V(\omega_1, a_1^*)]. \tag{1}$$

Graphically, condition (1) means that the probability-weighted increase in the sender's payoff as we move from A to B exceeds the probability-weighted decrease as we move from C to D.[4]

If (1) holds for all possible $\omega_1,\ \omega_2$ and π_1, full disclosure is optimal. Stated in this way, the condition does not help much, as it does not provide a recipe to verify it for all possible $\omega_1,\ \omega_2$ and π_1.

Our idea is as follows. First, instead of comparing the total probability-weighted changes in the sender's state-contingent payoff, we are going to compare "marginal changes" (weighted with the corresponding probabilities) as we move from A to B and from C to D, "pointwise". We will define what it means for a change to be "marginal" in such a way that if any marginal change on the way from A to B is larger than on the way from C to D, the total change will be larger as well.

Second, notice that any a on the way from A to B (i.e., between a^* and a_2^*), is higher than any a on the way from C to D (i.e., between a^* and a_1^*). In addition, $U_a(\omega_2, a) > 0$ for any $a \in [a^*, a_2^*)$, and $U_a(\omega_1, a) < 0$ for any $a \in (a_1^*, a^*]$. Since these properties hold for any message with binary support, they allow us to formulate a sufficient condition that neither involves specific posterior probabilities nor requires computing the optimal receiver's action.

We start from defining the marginal changes. We cannot compare marginal changes in the space of a, because $[a_1^*, a^*]$ and $[a^*, a_2^*]$ have different lengths. Hence, we move to the space of probability-weighted receiver's marginal utilities: $x_1 := \pi_1 U_a(\omega_1, a)$ and $x_2 := -\pi_2 U_a(\omega_2, a)$. As a runs from a^* to a_1^* (for x_1) and from a^* to a_2^* (for x_2), both x_1 and x_2 run from the *same* con-

[4]We are saying "increase" and "decrease" to relate to the graph. But, as we have said, $V(\omega, a)$ does not have to be upward sloping, so, in general, it is more accurate to talk about comparing the change as we move from A to B with a negative of the change as we move from C to D, exactly as in (1).

stant, $k < 0$, to zero. That the starting point is the same stems from the first-order condition under m:

$$\begin{aligned} \pi_1 U_a(\omega_1, a^*) + \pi_2 U_a(\omega_2, a^*) &= 0 \\ &\Rightarrow \pi_1 U_a(\omega_1, a^*) = -\pi_2 U_a(\omega_2, a^*) =: k \end{aligned} \tag{2}$$

That the arrival point is zero is due to the first-order condition under ω_i: $U_a(\omega_i, a_i^*) = 0$.

Now, since x_1 and x_2 span the same intervals, comparing marginal changes in $V(\omega_1, a)$ and $V(\omega_2, a)$ in the space of x_1 and x_2 (respectively) is legitimate. Comparing a marginal gain from revealing ω_2 with a marginal loss from revealing ω_1[5] at given $a_1 \in (a_1^*, a^*)$ and $a_2 \in (a^*, a_2^*)$, is the same as comparing $\partial(\pi_2 V(\omega_2, a_2))/\partial x_2(a_2)$ with $-\partial(\pi_1 V(\omega_1, a_1))/\partial x_1(a_1)$ (we are using ∂ to emphasize that we are differentiating while holding ω_i and π_i fixed). Thus, if

$$\partial(\pi_2 V(\omega_2, a_2))/\partial x_2(a_2) \geq -\partial(\pi_1 V(\omega_1, a_1))/\partial x_1(a_1)$$
$$\text{for all } a_1 \in (a_1^*, a^*) \text{ and } a_2 \in (a^*, a_2^*),$$

inequality (1) will be satisfied.

[5]Here again it would be more accurate to say "a marginal change" from revealing ω_2 and "negative of a marginal change" from revealing ω_1.

Now, notice that[6]

$$\begin{aligned}\frac{\partial(\pi_1 V(\omega_1,a))}{\partial x_1(a)} &= \frac{\partial V(\omega_1,a)}{\partial U_a(\omega_1,a)} = \frac{V_a(\omega_1,a)}{U_{aa}(\omega_1,a)},\\ \frac{\partial(\pi_2 V(\omega_2,a))}{\partial x_2(a)} &= \frac{\partial V(\omega_2,a)}{-\partial U_a(\omega_2,a)} = \frac{V_a(\omega_2,a)}{-U_{aa}(\omega_2,a)}.\end{aligned}$$

Moreover, notice that for all $a_1 \in (a_1^*, a^*)$, $U_a(\omega_1, a_1) < 0$, and for all $a_2 \in (a^*, a_2^*)$, $U_a(\omega_2, a_2) > 0$.

Consequently, if $-V_a(\omega_2,a_2)/U_{aa}(\omega_2,a_2) \geq -V_a(\omega_1,a_1)/U_{aa}(\omega_1,a_1)$ for any a_1, a_2, ω_1, ω_2 such that $a_1 < a_2$ and $U_a(\omega_1,a_1) < 0 < U_a(\omega_2,a_2)$, revealing the states in the support of any binary-support message benefits the sender. Hence, we arrive at the following sufficient condition for the optimality of splitting any message with binary support:

$$\text{For all } a_1,\ a_2,\ \omega_1,\ \omega_2,$$
$$\begin{cases} a_1 < a_2 \\ U_a(\omega_1,a_1) < 0 < U_a(\omega_2,a_2) \end{cases} \Rightarrow \frac{V_a(\omega_1,a_1)}{-U_{aa}(\omega_1,a_1)} \leq \frac{V_a(\omega_2,a_2)}{-U_{aa}(\omega_2,a_2)}. \qquad (3)$$

Condition (3) can be concisely phrased as the requirement that $V_a(\omega,a)/(-U_{aa}(\omega,a))$ goes up (or stays the same) whenever both a and $U_a(\omega,a)$ increase and $U_a(\omega,a)$ switches from negative to positive.

Lemma 1 *Under condition (3), for any message that generates a posterior with binary support, revealing the states in the support instead of sending*

[6] Formally, the second equality in each of the two lines below can be derived as follows. Let $y = U_a(\omega,a)$ and $a = U_a^{-1}(\omega,y)$ respectively. Then, holding ω fixed:

$$\frac{\partial V(\omega, U_a^{-1}(\omega,y))}{\partial y} = V_a(\omega, U_a^{-1}(\omega,y))\frac{\partial U_a^{-1}(\omega,y)}{\partial y} = \frac{V_a(\omega,U_a^{-1}(\omega,y))}{U_{aa}(\omega,U_a^{-1}(\omega,y))} = \frac{V_a(\omega,a)}{U_{aa}(\omega,a)}.$$

the message weakly increases the expected utility of the sender, conditional on the message. If the inequality between the ratios in (3) is strict, the expected utility of the sender strictly increases.

Kolotilin (2018) shows that, under certain assumptions, it is enough to consider only binary support messages to check for the optimality of full disclosure.[7] These assumptions are: (i) both A and Ω are compact intervals in $\mathbb{R}$, (ii) $U_a(\omega, a)$ and $V_a(\omega, a)$ are continuous in ω and continuously differentiable in a, (iii) for any posterior, the receiver's expected utility is single-peaked in a and his optimal a is interior, (iv) the receiver's optimal state-contingent action $a^*(\omega)$ is monotonic in ω ("single crossing").

We have milder requirements compared to (ii), and, more importantly, our framework does not impose (iv). So, we cannot rule out a priori that non-binary support messages be unneeded to optimize the sender's utility. Nonetheless, we are able to show that (3) is a sufficient condition for full disclosure, in the following way. First, we extend the argument of Lemma 1 to find a profitable split of any arbitrary message m. To start, we show that we can always split m into a message with binary support and a "complementary" message that both induce the same action as m. Then, if (3) holds as a strict inequality, it is tempting to say that a further split of the binary support message does the job and generates a welfare-improving ultimate

[7]Kolotilin (2018), Proposition 1, part (ii) and Corollary 1, part (ii). See also Kolotilin et al. (2022), Lemma 3, for a more explicit formulation. More precisely, both papers state that, under the assumptions that allow to focus on binary-support messages, full disclosure is optimal if and only if (1) holds for all possible ω_1, ω_2 and π_1. (By employing (2), Kolotilin (2018) expresses the condition in terms of $U_a(\omega_1, a^*)$ and $U_a(\omega_1, a^*)$ instead of π_1 and π_2.) As we argued in the Introduction, compared to these papers, our contribution consists of translating the necessary-and-sufficient but abstract condition (1) into a just sufficient but easily interpretable/verifiable condition, and extending it to settings where considering binary-support messages may not be enough.

split. However, with a continuous state space, the binary-support message may have a zero probability conditional on m, and then we cannot claim welfare improvement. We circumvent this problem by looking at arbitrarily small "neighborhoods" of the two states of the binary-support message. This allows us to claim that (3) with the *strict* instead of weak inequality is a sufficient condition for the optimality of full disclosure. The last step uses perturbations of the sender's utility function to claim that condition (3) is sufficient for the optimality of full disclosure. These steps are formalized in the proof of our main result:

Theorem 1 *Under condition (3) full disclosure is optimal for the sender.*

Condition (3) does not require computing the receiver's optimal response to a posterior and can be applied to a broad class of sender's and receiver's utility functions (Section 5 provides examples). Moreover, it can be interpreted as a requirement of minimal alignment of interest between the sender and the receiver. Suppose for a second that U_{aa} is a constant. Conditions $a_1 < a_2$ and $U_a(\omega_1, a_1) < 0 < U_a(\omega_2, a_2)$ mean that state ω_2 generates positive incentives for the receiver (i.e., the incentive to increase a_2) and state ω_1 generates negative incentives (i.e., the incentive to decrease a_1). Then, (3) requires that the sender's marginal benefit from an increase in action is (weakly) larger when such an increase is desirable for the receiver with respect to when it is not.

"Normalization" of V_a by U_{aa} in (3) can be understood as follows. It is important not only how strong the sender's utility reacts to marginal changes in action, but also how far the action moves once the state is revealed. The "speed of readjustment" is determined precisely by U_{aa}. When $-U_{aa}(\omega_2, a_2)$

is lower, a_2 increases slower, that is, it goes a longer way until it reaches the optimal value under ω_2. This implies a higher benefit for the sender from the revelation of ω_2 if $V_a(\omega_2, a_2)$ is positive (a higher loss if $V_a(\omega_2, a_2)$ is negative). Similarly, when $-U_{aa}(\omega_1, a_1)$ is lower, a_1 goes a longer way, but now this is a *decrease* towards the new optimal action, so there is a higher loss from the revelation of ω_1 if $V_a(\omega_1, a_1)$ is positive (a higher benefit if $V_a(\omega_2, a_2)$ is negative).

Note also that condition (3) is always trivially satisfied when $V = U$, that is, when the incentives of the parties are perfectly aligned. This is because $U_a(\omega_1, a_1) < 0 < U_a(\omega_2, a_2)$ implies

$$\frac{U_a(\omega_1, a_1)}{-U_{aa}(\omega_1, a_1)} < \frac{U_a(\omega_2, a_2)}{-U_{aa}(\omega_2, a_2)},$$

given that $U_{aa} < 0$.

3.2 Derivatives condition

A stronger but somewhat simpler condition than (3) is the following:

$$\text{For all } a_1,\ a_2,\ \omega_1,\ \omega_2,$$
$$\left\{ \begin{array}{c} a_1 < a_2 \\ U_a(\omega_1, a_1) < U_a(\omega_2, a_2) \end{array} \right. \Rightarrow \frac{V_a(\omega_1, a_1)}{-U_{aa}(\omega_1, a_1)} \leq \frac{V_a(\omega_2, a_2)}{-U_{aa}(\omega_2, a_2)}. \tag{4}$$

It is stronger than (3) because it requires that the relation between the ratios holds for a larger set of $(\omega_1, a_1), (\omega_2, a_2)$ pairs, where $U_a(\omega_1, a_1)$ does not have to be negative and $U_a(\omega_2, a_2)$ does not have to be positive.

Assuming that U_{aaa}, $U_{aa\omega}$ and V_{aa} exist, condition (4) can be expressed in terms of just derivatives of U and V. To see this, notice that (4) is equivalent to stating that, at each (ω, a), $-V_a(\omega, a)/U_{aa}(\omega, a)$ is increasing in all directions in which both a and $U_a(\omega, a)$ increase. So, by applying directional derivatives, one can show the lemma below. Namely, consider the following conditions:

$$\text{For each } (\omega, a) \text{ s.t. } U_{a\omega} > 0,$$
$$\begin{cases} U_{aa\omega} V_a \geq V_{a\omega} U_{aa} \\ V_a (U_{aaa} U_{a\omega} - U_{aa\omega} U_{aa}) \geq U_{aa} (V_{aa} U_{a\omega} - V_{a\omega} U_{aa}) \end{cases}, \tag{5}$$

and

$$\text{For each } (\omega, a) \text{ s.t. } U_{a\omega} < 0,$$
$$\begin{cases} U_{aa\omega} V_a \leq V_{a\omega} U_{aa} \\ V_a (U_{aaa} U_{a\omega} - U_{aa\omega} U_{aa}) \leq U_{aa} (V_{aa} U_{a\omega} - V_{a\omega} U_{aa}) \end{cases}. \tag{6}$$

Lemma 2 *Assume that U_{aaa}, $U_{aa\omega}$ and V_{aa} exist. Then condition (4) is equivalent to (5) and (6).*

Notice that (5) and (6) do not cover the case $U_{a\omega} = 0$. This is because, when $U_{a\omega} = 0$, there is simply no direction in which both a and U_a increase.

Subsection 5.2 will illustrate the application of the derivatives conditions.

3.3 Sufficient condition for suboptimality of full disclosure

Subsection 3.1 delivered a sufficient condition for the optimality of full disclosure. We can apply almost the same scheme of reasoning to derive a sufficient condition for the *sub*optimality of full disclosure. Instead of the existence of a welfare-improving split for any message with binary support, the suboptimality of full disclosure requires the existence of at least one pair of states that can be pooled (or partially pooled) so as to improve the sender's welfare.

Namely, fix a pair of states ω_1, ω_2 and consider the following condition

$$\text{For all } a_1,\ a_2,$$
$$\left\{\begin{array}{c} a_1 < a_2 \\ U_a(\omega_1, a_1) < 0 < U_a(\omega_2, a_2) \end{array}\right. \Rightarrow \frac{V_a(\omega_1, a_1)}{-U_{aa}(\omega_1, a_1)} > \frac{V_a(\omega_2, a_2)}{-U_{aa}(\omega_2, a_2)} \tag{7}$$

This condition resembles (3) except that it is formulated for given ω_1 and ω_2 and the sign of the inequality between the ratios flips.

Theorem 2 *If there exists a pair of states $\omega_1, \omega_2 \in \operatorname{supp} p$ such that (7) holds non-vacuously, full disclosure is suboptimal for the sender.*

Notice that Theorem 2 does not imply that (3) delivers a necessary and sufficient condition for the optimality of full disclosure. The fact that $-V_a(\omega_1, a_1)/U_{aa}(\omega_1, a_1) \leq -V_a(\omega_2, a_2)/U_{aa}(\omega_2, a_2)$ fails to hold for some $a_1,\ a_2,\ \omega_1,\ \omega_2$ such that $a_1 < a_2$ and $U_a(\omega_1, a_1) < 0 < U_a(\omega_2, a_2)$ does not mean that there will necessarily be a pair of states ω_1 and ω_2 for which $-V_a(\omega_1, a_1)/U_{aa}(\omega_1, a_1) > -V_a(\omega_2, a_2)/U_{aa}(\omega_2, a_2)$ for *all* $a_1,\ a_2$, such that $a_1 < a_2$ and $U_a(\omega_1, a_1) < 0 < U_a(\omega_2, a_2)$, as the relation between the ratios may change sign as a_1 and a_2 change.

4 Well-known special cases

In this section we compare our sufficient condition with the conditions derived in the literature for two specific cases.

4.1 "Linear case"

Much of the literature has focused on settings in which the sender's payoff from sending a certain message can ultimately be represented as a function of the posterior mean only. This is the case, for example, when the receiver's action only depends on the expected state, $\mathbb{E}(\omega|m)$, and the sender's utility only depends on the receiver's action: $V(\omega, a) = V(a)$. Then, given the posterior induced by message m, the sender's payoff is $V(a^*(\mathbb{E}(\omega|m))$, which can be represented as an indirect utility function, $\widehat{V}(\mathbb{E}(\omega|m))$. It is well known that the necessary and sufficient condition for the optimality of full disclosure in this case is that $\widehat{V}(\cdot)$ is convex on the set of admissible values for $\mathbb{E}(\omega|m)$.

A particularly simple case is the "linear case" (Kolotilin et al. (2022)), in which $V(\omega, a) = V(a)$ and $U_a(\omega, a) = \omega - a$.[8] This shape of U_a arises, for example, in the classical case of a quadratic loss function of the receiver: $U(\omega, a) = -\frac{1}{2}(a - \omega)^2$. Then $a^*(\mathbb{E}(\omega|m)) = \mathbb{E}(\omega|m)$, and the convexity of $\widehat{V}(\cdot)$ is equivalent to the convexity of $V(a)$. Following Kolotilin et al. (2022), assume $A = \Omega = [0, 1]$. Hence, the necessary and sufficient condition for the optimality of full disclosure in the "linear case" can be written as

$$V'(a_1) \leq V'(a_2) \text{ for any } a_1 \in (0, 1), a_2 \in (0, 1), \text{ such that } a_1 < a_2. \quad (8)$$

[8] More generally, $U_a(\omega, a)$ can be any linear function of ω and a.

In this context, our condition (3) becomes

$$\text{For all } a_1,\ a_2,\ \omega_1,\ \omega_2, \begin{cases} a_1 < a_2 \\ \omega_1 - a_1 < 0 < \omega_2 - a_2 \end{cases} \Rightarrow V'(a_1) \leq V'(a_2) \quad (9)$$

At first sight, (9) seems weaker than (8) due to the extra restriction before the implication sign, $\omega_1 - a_1 < 0 < \omega_2 - a_2$. Note however that if it were truly weaker, it would be wrong, because (8) is a necessary condition. But for any $a_1 \in (0, 1),\ a_2 \in (0, 1)$, such that $a_1 < a_2$, one can always pick ω_1 and ω_2 such that $\omega_1 - a_1 < 0 < \omega_2 - a_2$. Hence, $\omega_1 - a_1 < 0 < \omega_2 - a_2$ becomes redundant in (9). The bottom line is that our sufficient condition for optimality of full disclosure is in fact necessary and sufficient in the "linear case".

4.2 "Linear receiver case"

Another simple case is what Kolotilin et al. (2022) call the "linear receiver case": $U_a(\omega, a) = \omega - a$ but V may depend on ω. As Kolotilin et al. show, a sufficient condition for full disclosure to be optimal is that the sender's utility is convex in a and supermodular in (a, ω), that is

$$\begin{cases} V_a(\omega, a_1) \leq V_a(\omega, a_2) \text{ for any } \omega \text{ and } a_1 < a_2 \\ V_a(\omega_1, a) \leq V_a(\omega_2, a) \text{ for any } a \text{ and } \omega_1 < \omega_2 \end{cases} \quad (10)$$

In this context, (3) becomes

$$\text{For all } a_1,\ a_2,\ \omega_1,\ \omega_2, \begin{cases} a_1 < a_2 \\ \omega_1 - a_1 < 0 < \omega_2 - a_2 \end{cases} \Rightarrow V_a(\omega_1, a_1) \leq V_a(\omega_2, a_2) \tag{11}$$

Our condition is weaker because it requires $V_a(\omega, a)$ to (weakly) increase only when a grows *and* ω grows more than a, more precisely from being smaller to being larger than a. In particular, our condition requires neither convexity of $V(\omega, a)$ in a, nor its supermodularity in a and ω. For example, take the classical setting of Crawford and Sobel (1982) with $U(\omega, a) = -(\omega - a)^2$ and $V(\omega, a) = -(\omega - a - b)^2$ with $b \geq 0$. These preferences satisfy the assumptions of the "simple receiver case". The condition from Kolotilin et al. (2022) does not hold because the sender's utility is concave in a. Instead, our condition is satisfied, as $V_a(\omega_1, a_1) \leq V_a(\omega_2, a_2)$ becomes simply $\omega_1 - a_1 \leq \omega_2 - a_2$. Although there is a disagreement between the sender and the receiver regarding the optimal action in each state, full disclosure is nonetheless optimal, and our condition sheds light on why it is so: News about the state move the marginal utilities of the two parties in the same direction, therefore the decrease of action under "bad" news has a lower impact on the utility of the sender than the increase of action under "good" news.

5 Application to a principal-agent model

In this section we explore the implications of our results in the following principal-agent setting. An agent exerts effort a to produce output $y(\omega, a)$. He bears the cost of effort, which is normalized to be a (in other words, a

should be treated as disutility of effort). The agent receives wage $w(y)$, and the principal receives $y - w(y)$. The agent's and the principal's utilities of money are (weakly) concave functions $u(\cdot)$ and $v(\cdot)$ respectively. The agent does not know ω, while the principal does and can send a message to the agent before he chooses effort. So, the agent is the receiver and the principal is the sender.

For simplicity, we assume that the wage is linear, that is, the agent receives a fixed share δ of the output. While we take the compensation scheme for the agent as given, the conclusions about the optimality of full disclosure will not depend on δ, as we will see. However, allowing for a non-linear wage schedule and jointly solving for the optimal wage schedule and disclosure policy could be an interesting avenue for future research.

We will first examine the implications of the parties' risk-aversion for the optimality of full disclosure, given a simple and meaningful production function. Then we will simplify the parties' preferences by assuming their risk-neutrality and focus on the properties of the production function instead.

5.1 Effects of risk aversion in a simple setting

Consider the following setting:

$$\begin{aligned} y(\omega, a) &= \omega a^{\kappa}, \ \kappa \in (0,1), \ w(y) = \delta y \\ u(x) &= \frac{x^{1-\gamma}}{1-\gamma}, \ v(x) = \frac{x^{1-\rho}}{1-\rho} \end{aligned}$$

That is, both the agent and the principal exhibit CRRA with coefficients γ and ρ respectively, where both γ and ρ are non-negative and different from 1. Assume that the upper boundary of A is large enough to ensure the interior solution of the agent's problem.

We can compute:

$$\begin{aligned}
U(\omega, a) &= \frac{1}{1-\gamma}(\delta\omega)^{1-\gamma} a^{\kappa(1-\gamma)} - a, \\
U_a(\omega, a) &= \kappa(\delta\omega)^{1-\gamma} a^{\kappa(1-\gamma)-1} - 1, \\
U_{aa}(\omega, a) &= (\kappa(1-\gamma) - 1)\kappa(\delta\omega)^{1-\gamma} a^{\kappa(1-\gamma)-2}, \\
V(\omega, a) &= \frac{1}{1-\rho}((1-\delta)\omega)^{1-\rho} a^{\kappa(1-\rho)}, \\
V_a(\omega, a) &= \kappa((1-\delta)\omega)^{1-\rho} a^{\kappa(1-\rho)-1}.
\end{aligned}$$

Notice that the principal's utility cannot be expressed as a function of the posterior mean, so we cannot use the familiar convexity/non-convexity argument to establish the optimality/suboptimality of full disclosure.

With some algebra, one can derive

$$\frac{V_a(\omega, a)}{-U_{aa}(\omega, a)} = const \cdot (U_a(\omega, a) + 1)^{\frac{\gamma-\rho}{1-\gamma}} \cdot a^{\frac{1-\rho}{1-\gamma}},$$

where *const* is a positive constant.

It is straightforward to check that the ratio is increasing as both U_a and a go up when $\rho \leq \gamma < 1$ or $\rho \geq \gamma > 1$. Hence, in this case, (3) holds, and full disclosure is optimal (see Figure 2). At the same time, under $\rho < 1 < \gamma$ or $\gamma < 1 < \rho$, the ratio is decreasing when both U_a and a increase. According to Theorem 2, full disclosure is then suboptimal. In all other cases, the ratio is decreasing in U_a and increasing in a. Then, neither (3) nor (7) is satisfied, and our analysis is inconclusive in such cases.

Figure 2.

We can notice that full disclosure fails to be optimal when ρ and γ are on the opposite sides from 1. This is related to the fact that, in this case, state end effort are complements for one party and substitutes for the other, which can be seen by examining the expressions for $U_a(\omega, a)$ and $V_a(\omega, a)$. In contrast, when ρ and γ are both smaller or both greater than 1, the direction of interaction between state and effort is the same for both parties, and, thus, full disclosure gets a chance.

For example, consider a typical textbook situation with a risk neutral principal ($\rho = 0$) and a risk averse agent. If the agent is not too risk averse

($\gamma < 1$), full disclosure is optimal. Since state and effort are complements for both parties, the principal benefits more from effort exactly when the agent has higher incentives to exert effort. Instead, when the agent becomes too risk averse ($\gamma > 1$), state and effort become substitutes for the agent. As a result, good news about productivity *depress* effort, while the principal benefits more from effort in higher states. As a result, full disclosure ceases to be optimal.

When the principal is highly risk averse ($\rho > 1$) the story is reversed: now *insufficient* risk aversion of the agent ($\gamma < 1$) implies that full disclosure is suboptimal. This is because now the principal benefits more from effort under *lower* states, while for the agent state and effort are complements. One needs to make the agent sufficiently risk averse ($\gamma > 1$) to align the interaction of effort and state between the two parties, so that full disclosure can be optimal.

What is interesting about the case of a highly risk averse principal is that full disclosure can be optimal despite lowering the expected effort and can be harmful despite raising the expected effort. Indeed, one can easily derive that the disclosure of states in the support of any given message increases the expected effort under $\gamma < 1$ and lowers it under $\gamma > 1$. This observation demonstrates that an increase (decrease) in the average effort due to disclosure is not sufficient to make full disclosure optimal (suboptimal), as the direction and strength of the interaction between state and effort in the principal's payoff matters too.

The role of complementarity/substitutability between the action and the state can also be observed if one carefully looks at our general condition (3). The interaction between the action and the state for the two parties matters because it affects whether $V_a(\omega, a)$ comoves with $U_a(\omega, a)$ when both a and $U_a(\omega, a)$ increase. Specifically, when action and state are complementary for the receiver, higher $U_a(\omega, a)$ together with higher a imply higher ω, meaning that $\omega_2 > \omega_1$ in (3). Then, if action and state are complementary for the sender as well, higher ω pushes $V_a(\omega, a)$ upwards for given a, thereby relaxing (3). In contrast, if action and state are substitutes for the sender, higher ω pushes $V_a(\omega, a)$ downward for given a, thereby tightening (3). By similar logic, if action and state are substitutes for the receiver, (3) is more (less) likely to be satisfied when they are substitutes (complements) for the sender. A word of caution: Although the fact that action and state are complements (or substitutes) for both parties helps to satisfy (3), it generally implies neither (3), nor that full disclosure is optimal.[9]

5.2 Risk neutral agent and principal, separable production function

Sometimes it is more convenient to use the derivatives conditions (5) or (6) instead of (3). This section illustrates how to apply them in a simple setting. In the previous subsection, we assumed a simple production function and played with risk aversion of the parties. Let us now assume that both parties' utilities are linear in output and examine different production functions instead. Linearity in output for both parties would arise, for example, in a setting where both parties are risk neutral and the wage is linear in

[9]For example, if $y = \omega\varphi(a)$ and both parties are risk-neutral, one can show that the sender's payoff can be represented as a function of just the posterior mean and then derive that full disclosure is optimal if and only if $\varphi'''(a)\varphi'(a) \geq (\varphi'(a))^2$. Hence, despite complementarity between the state and the action for both parties, full disclosure may be suboptimal. See also Mensch (2021) for a discussion on the role of complementarities for the optimality of full disclosure.

output.

The utilities of the agent and the principal under these assumptions are: $U(\omega, a) = \delta y(\omega, a) - a$ and $V(\omega, a) = (1-\delta)y(\omega, a)$, respectively, where δ is a positive constant.

Suppose[10]

$$y(\omega, a) = \beta(\omega)\varphi(a) + \xi(a), \tag{12}$$

with $\beta(\cdot) > 0,\ \beta'(\cdot) > 0,\ \varphi(\cdot) > 0,\ \varphi'(\cdot) > 0,\ \xi'(\cdot) \geq 0,\ \xi''(\cdot) + \varphi''(\cdot) < 0,$ $\xi''(\cdot)\varphi''(\cdot) \geq 0$ ($\xi''(\cdot)+\varphi''(\cdot) < 0$ ensures strict concavity of $y(\omega, a)$). Assume also $y_a(\omega, a)|_{a=\sup A} < 1/\delta$ to ensure that the agent's choice of a is interior. This output function could be called "multiplicatively-additively" separable in state and effort; we will call it just "separable", for simplicity. Special cases of this form (such as $\omega\sqrt{a}$ employed in the previous subsection) are commonly used in the literature.[11]

Due to our assumptions on $\beta(\cdot)$ and $\varphi(\cdot)$, state and effort are complements ($U_{a\omega} > 0$). Thus, the relevant condition is (5), which becomes:

$$\text{For each } (\omega, a), \begin{cases} y_{aa\omega}y_a \geq y_{a\omega}y_{aa} \\ y_{aaa}y_{a\omega} \geq y_{aa\omega}y_{aa} \end{cases}, \tag{13}$$

---

[10] y may also contain a term "$\alpha(\omega)$" that only depends on ω, but it would be irrelevant for both parties' choice problems.

[11] It is fair to note that our sufficient condition is not the only way to check for the optimality of full disclosure in this setting. One can show that the sender's payoff can eventually be represented as a function of expected $\beta(\omega)$ and then try to check for the convexity of this function. However, because the function turns out to be cumbersome, this is a daunting task, in general. For example, it is hard to use when $\varphi(\cdot)$ and $\xi(\cdot)$ are arbitrary concave power functions, while our condition is easy to apply, as we demonstrate below.

Using (12), condition (13) can be rewritten as:

$$\text{For each } (\omega, a),$$
$$\begin{cases} \varphi''(a)\xi'(a) \geq \varphi'(a)\xi''(a) \\ \beta(\omega)\left[\varphi'''(a)\varphi'(a) - (\varphi''(a))^2\right] \geq \xi''(a)\varphi''(a) - \xi'''(a)\varphi'(a) \end{cases}. \tag{14}$$

Now let us check (14) for some specific functional forms of $\varphi(\cdot)$ and $\xi(\cdot)$. As a first example, assume that both $\varphi(\cdot)$ and $\xi(\cdot)$ are weakly concave power functions: $\varphi(a) = ha^{\kappa}$, $\xi(a) = la^{\tau}$ with $h > 0,\ l > 0,\ \kappa \in (0, 1],\ \tau \in [0, 1]$, such that κ and τ are not both 1 (to ensure the strict concavity of the output). It is straightforward to derive that the first inequality boils down to $\kappa \geq \tau$, and the second inequality always holds. Thus, $\kappa \geq \tau$ is a sufficient condition for the optimality of full disclosure.

As another example, consider $\varphi(a) = h \cdot \ln a$ and $\xi(a) = l \cdot \ln a$. Then the first inequality holds as an equality, and it can be easily checked that the second one is always satisfied. Hence, full disclosure is always optimal in such a case.

6 Conclusion

In this paper, we have addressed the following question: When is it optimal for a privately-informed sender to commit to full disclosure of her information to the receiver? We answer with a sufficient condition that can be interpreted as a minimal alignment of incentives between the sender and the receiver.

Several recent papers have derived conditions for the optimality of full disclosure in terms of the sender's *indirect* utility function, assuming that it only depends on the posterior mean. Our condition, instead, speaks directly to the primitive incentives of the parties and does not rely on any assumption on how the state affects them. For this reason, it can be easily interpreted and verified in applications.

In a principal-agent setting where the principal is privately informed of a state that affects the productivity of the agent's effort, the optimal effort of a risk-averse agent depends on the entire shape of his posterior belief. As a consequence, given a disclosure policy, the indirect utility function of the principal does not only depend on the posterior mean, and the conditions that require this cannot be applied. Our condition, along with an analogous sufficient condition for *sub*optimality of full disclosure that we derive, can instead be used to study when full disclosure is optimal and when it is not, and to interpret the results in light of the risk aversion of the parties. For instance, we find that full transparency is optimal under the common modeling assumptions of risk-neutrality of the principal and risk-aversion of the agent, provided that the agent is not too risk averse (CRRA with the coefficient of relative risk aversion below one).

One interesting question is: In a principal-agent relationship, how does the optimality of full disclosure depend on the compensation scheme for the agent? More generally, how to jointly determine the optimal compensation scheme and disclosure policy? This is an avenue for future research.

7 Appendix

Proof of Lemma 1. Consider two states, ω_1 and ω_2, and a message m with support $\{\omega_1, \omega_2\}$. Let $\pi_1 := \Pr(\omega_1|m)$, $\pi_2 := \Pr(\omega_2|m)$, $\pi_2 = 1 - \pi_1$. Let the receiver's optimal actions in states ω_1, ω_2 and under message m

be, respectively, a_1^*, a_2^*, and a^*. Due to our assumptions on $U(\omega, a)$, each of a_1^*, a_2^* and a^* is unique and determined by the corresponding first-order condition.

If $a_1^* = a_2^*$, revealing the states is inconsequential. So, without loss of generality, let $a_2^* > a_1^*$. Then, from the receiver's first-order condition under m and strict concavity of $U_a = (\omega, a)$ in a, we get $a_2^* > a^* > a_1^*$.

The sender (weakly) benefits from disclosing ω_1, ω_2 instead of sending m if and only if

$$\pi_1 V(\omega_1, a^*) + \pi_2 V(\omega_2, a^*) \leq \pi_1 V(\omega_1, a_1^*) + \pi_2 V(\omega_2, a_2^*),$$

that is,

$$\pi_1[V(\omega_1, a^*) - V(\omega_1, a_1^*)] \leq \pi_2[V(\omega_2, a_2^*) - V(\omega_2, a^*)]. \tag{15}$$

Write (15) as

$$\int_{a_1^*}^{a^*} \pi_1 V_a(\omega_1, a)da \leq \int_{a^*}^{a_2^*} \pi_2 V_a(\omega_2, a)da. \tag{16}$$

Let $x_1(a) := \pi_1 U_a(\omega_1, a)$ and $x_2(a) := -\pi_2 U_a(\omega_2, a)$. Due to the first-order conditions for the receiver under ω_1, and ω_2, we have: $x_1(a_1^*) = x_2(a_2^*) = 0$. In addition, the receiver's first-order condition under message m yields:

$$\begin{aligned} \pi_1 U_a(\omega_1, a^*) + \pi_2 U_a(\omega_2, a^*) &= 0 & (17) \\ &\Rightarrow x_1(a^*) = x_2(a^*) =: k < 0; \end{aligned}$$

$k < 0$ comes from $U_a(\omega_i, a_i^*) = 0$, $a_2^* > a^* > a_1^*$, and strict concavity of $U_a(\omega, a)$ in a. Then, given that $dx_1 := \pi_1 U_{aa}(\omega_1, a)da$ and $dx_2 := -\pi_2 U_{aa}(\omega_2, a)da$, (16) is equivalent to

$$-\int_k^0 \frac{V_a(\omega_1, a_1(x_1))}{U_{aa}(\omega_1, a_1(x_1))} dx_1 \leq \int_k^0 \frac{V_a(\omega_2, a_2(x_2))}{-U_{aa}(\omega_2, a_2(x_2))} dx_2, \tag{18}$$

where $a_i(x_i)$ is the value of a derived from the definition of x_i, i.e.,

$$a_1(x_1) := U_a^{-1}(\omega_1, x_1/\pi_1),\ a_2(x_2) := U_a^{-1}(\omega_2, -x_2/\pi_2).$$

So, if $-V(\omega_1, a_1(x_1))/U_{aa}(\omega_1, a_1(x_1)) \leq -V_a(\omega_2, a_2(x_2))/U_{aa}(\omega_2, a_2(x_2))$ for any $x_1 = x_2 \in (k, 0)$, then (18) (hence, (15)) is satisfied.

For any $x_1 \in (k, 0),\ x_2 \in (k, 0)$, we have $a_1(x_1) \in (a_1^*, a^*),\ a_2(x_2) \in (a^*, a_2^*)$, that is, $a_1(x_1) < a_2(x_2)$ and $U_a(\omega_1, a_1(x_1)) < 0 < U_a(\omega_2, a_2(x_2))$. This means that (15) holds for any $\omega_1,\ \omega_2$, and π_1 if the following condition is satisfied:

$$\text{For all } a_1,\ a_2,\ \omega_1,\ \omega_2,$$
$$\begin{cases} a_1 < a_2 \\ U_a(\omega_1, a_1) < 0 < U_a(\omega_2, a_2) \end{cases} \Rightarrow \frac{V_a(\omega_1, a_1)}{-U_{aa}(\omega_1, a_1)} \leq \frac{V_a(\omega_2, a_2)}{-U_{aa}(\omega_2, a_2)},$$

which is condition (3).

To ensure that the sender strictly benefits from the split, we need that (15) holds as a strict inequality. Clearly, for this, we only need that $\leq$ turns into $<$ in the above condition. ■

Proof of Theorem 1. We first prove that, under (3) with strict instead of weak inequality, full disclosure is optimal for the sender. We do so by showing that any message m^* with non-singleton support Ω^* that pools states that induce different actions is suboptimal.

Let π denote the posterior probability distribution conditional on m^*. Given a function f of states and actions, given a message m and an action a, we let $\widetilde{f}(m,a)$ denote the expected value of $f(\omega,a)$ conditional on m. Let a^* denote the agent's optimal action upon receiving m^*. It is obtained by solving the first-order condition $\widetilde{U}_a(m^*,a) = 0$.

If revealing the states in the support of m^* can change the receiver's action, then there exist $\omega_1^*, \omega_2^* \in \Omega^*$ such that[12]

$$U_a(\omega_1^*, a^*) < 0 < U_a(\omega_2^*, a^*).$$

Then, by continuity of $U_a(\omega,a)$ in ω, there exists $\varepsilon > 0$ such that, for all intervals Ω_1, Ω_2 of length smaller than ε whose interiors contain ω_1^*, ω_2^*,

$$\forall\,(\omega_1,\omega_2) \in \Omega_1 \times \Omega_2, \quad U_a(\omega_1,a^*) < 0 < U_a(\omega_2,a^*). \tag{19}$$

Note that, since $\omega_1^*, \omega_2^* \in \Omega^*$, $\pi(\Omega_1)\pi(\Omega_2) > 0$.

For all intervals Ω_1, Ω_2 of length smaller than ε whose interiors contain ω_1^*, ω_2^*, let us decompose m^* into three messages as follows: m_1, m_2 with supports contained in Ω_1, Ω_2, plus a complementary message m_c that induces action a^* (i.e., $\widetilde{U}_a(m_c, a^*) = 0$) such that $\pi(m_1) > 0$, $\pi(m_2) > 0$, $\pi(m_c) > 0$. To be precise, by "decomposition" we mean that these messages are never sent in states outside Ω^*, and, for each $\omega \in \Omega^*$, conditional on m^* being drawn, one of the three messages is sent instead of

[12]If a state ω_1^* in the support of m^* induces a lower action than a^*, then there must also be a state ω_2^* in the support of m^* that induces a higher action for a^* in order to satisfy the first-order condition after m^*, and vice versa.

m^*, so that $\Pr(m_1|m^*,\omega) + \Pr(m_2|m^*,\omega) + \Pr(m_c|m^*,\omega) = 1$. Obviously, $\pi(m_1) + \pi(m_2) + \pi(m_c) = 1$.

Such messages can be constructed because for any decomposition of m into m_1, m_2, m_c,

$$\widetilde{U}_a(m^*, a^*) \equiv \pi(m_1)\widetilde{U}_a(m_1, a^*) + \pi(m_2)\widetilde{U}_a(m_2, a^*) + \pi(m_c)\widetilde{U}_a(m_c, a^*) = 0,$$

and if the supports of m_1, m_2 are contained in Ω_1, Ω_2, m_1 and m_2 satisfy $\widetilde{U}_a(m_1, a^*) < 0$ and $\widetilde{U}_a(m_2, a^*) > 0$ by (19). Hence, we can always adjust m_1 and m_2 so that

$$\pi(m_1)\widetilde{U}_a(m_1, a^*) + \pi(m_2)\widetilde{U}_a(m_2, a^*) = 0 \text{ (and hence } \widetilde{U}_a(m_c, a^*) = 0\text{).} \quad (20)$$

For every sequence of pairs of intervals Ω_1, Ω_2 that contain ω_1^* and ω_2^* in their interiors and have length smaller than ε and converging to 0, consider the corresponding sequence of messages. For each point (m_1, m_2, m_c) of the sequence, consider the relative probabilities

$$\frac{\pi(m_1)}{\pi(m_1) + \pi(m_2)}, \frac{\pi(m_2)}{\pi(m_1) + \pi(m_2)}.$$

The sequence of these probabilities lives in the compact square $[0,1]^2$, therefore it has a subsequence that converges to two values $p_1^*, p_2^* \in [0,1]$ with $p_1^* + p_2^* = 1$. Let $(m_1^n, m_2^n, m_c^n)_{n>0}$ denote the corresponding subsequence of messages. For each $n > 0$, recall from (20) that

$$-\frac{\pi(m_1^n)}{\pi(m_1^n) + \pi(m_2^n)}\widetilde{U}_a(m_1^n, a^*) = \frac{\pi(m_2^n)}{\pi(m_1^n) + \pi(m_2^n)}\widetilde{U}_a(m_2^n, a^*).$$

By continuity of $U_a(\omega, a^*)$ in ω, we have $\lim_{n\to\infty} \widetilde{U}_a(m_i^n, a^*) = U_a(\omega_i^*, a^*)$ for each $i = 1, 2$. Hence, we get

$$-p_1^* U_a(\omega_1^*, a^*) = p_2^* U_a(\omega_2^*, a^*). \tag{21}$$

Note that this also implies $p_1^*, p_2^* \neq 0$.

For each $n > 0$, call a_1^n, a_2^n the receiver's optimal actions under m_1^n, m_2^n, i.e., $\widetilde{U_a}(m_1^n, a_1^n) = \widetilde{U_a}(m_2^n, a_2^n) = 0$. Note that $a_1^n < a^* < a_2^n$, as $\widetilde{U_a}(m_1^n, a^*) < 0 < \widetilde{U_a}(m_2^n, a^*)$ and U is strictly concave in a. The sender's expected utility increases after the decomposition of m^* into m_1^n, m_2^n, m_c^n if the following inequality holds:

$$\begin{aligned}
\widetilde{V}(m^*, a^*) &= \pi(m_1^n)\widetilde{V}(m_1^n, a^*) + \pi(m_2^n)\widetilde{V}(m_2^n, a^*) + \Pr(m^n)\widetilde{V}(m_c^n, a^*) \\
&< \pi(m_1^n)\widetilde{V}(m_1^n, a_1^n) + \pi(m_2^n)\widetilde{V}(m_2^n, a_2^n) + \Pr(m^n)\widetilde{V}(m_c^n, a^*).
\end{aligned}$$

Rewrite the inequality as

$$\pi(m_1^n)\left[\widetilde{V}(m_1^n, a^*) - \widetilde{V}(m_1^n, a_1^n)\right] < \pi(m_2^n)\left[\widetilde{V}(m_2^n, a_2^n) - \widetilde{V}(m_2^n, a^*)\right],$$

and then as

$$\pi(m_1^n)\int_{a_1^n}^{a^*} \widetilde{V}_a(m_1^n, a)da < \pi(m_2^n)\int_{a^*}^{a_2^n} \widetilde{V}_a(m_2^n, a)da. \tag{22}$$

Call a_1^*, a_2^* the receiver's optimal actions under ω_1^* and ω_2^*. For each $i = 2$, by continuity of U_a in ω, we have $\lim_{n\to\infty} a_i^n = a_i^*$, and by continuity of V_a

in ω, we have $\lim_{n\to\infty} \widetilde{V}_a(m_i^n, a) = V_a(\omega_i^*, a)$. Therefore,

$$\begin{aligned}
\lim_{n\to\infty} \pi(m_1^n) \int_{a_1^n}^{a^*} \widetilde{V}_a(m_1^n, a)da &= p_1^* \int_{a_1^n}^{a^*} V_a(\omega_1^*, a)da, \\
\lim_{n\to\infty} \pi(m_2^n) \int_{a^*}^{a_2^n} \widetilde{V}_a(m_2^n, a)da &= p_2^* \int_{a^*}^{a_2^n} V_a(\omega_2^*, a)da.
\end{aligned}$$

So, it is enough to show

$$p_1^* \int_{a_1^*}^{a^*} V_a(\omega_1^*, a)da < p_2^* \int_{a^*}^{a_2^*} V_a(\omega_2^*, a)da; \tag{23}$$

then, for sufficiently large n, the decomposition satisfies (22).

Thus, we have reduced the problem to checking if decomposing a hypothetical message with binary support $\{\omega_1^*, \omega_2^*\}$ and relative probabilities p_1^* and $p_2^* = 1 - p_1^*$ of the two states strictly benefits the sender. To see it, notice that, if we replace p_i^* with π_i, and ω_i^* with ω_i, (23) and (21) become (16) and (17) from the proof of Lemma 1, except that (23) is a strict inequality while (16) is a weak inequality. Hence, we arrive at the same sufficient condition as Lemma 1 delivers, except that (as noted at the end of the proof of the lemma) the inequality between $\dfrac{V_a(\omega_1, a_1)}{-U_{aa}(\omega_1, a_1)}$ and $\dfrac{V_a(\omega_2, a_2)}{-U_{aa}(\omega_2, a_2)}$ becomes strict.

The last step of the proof is showing that, if full disclosure is optimal when condition (3) holds with strict inequality between the ratios, so it is when it holds with weak inequality. Suppose by contradiction that full disclosure is suboptimal and condition (3) holds. Let Δ denote the difference between the expected utility of the sender under the optimal communication scheme and under full disclosure. Fix $\gamma \in (0,1)$ and let

$$\hat{V}(\omega, a) = V(\omega, a) - \gamma \exp\left(U_a(\omega, a)\right).$$

So we have

$$\begin{aligned}
\hat{V}_a(\omega, a) &= V_a(\omega, a) - \gamma U_{aa}(\omega, a) \exp\left(U_a(\omega, a)\right), \\
\frac{\hat{V}_a(\omega, a)}{-U_{aa}(\omega, a)} &= \frac{V_a(\omega, a)}{-U_{aa}(\omega, a)} + \gamma \exp\left(U_a(\omega, a)\right).
\end{aligned}$$

Since $\exp\left(U_a(\omega, a)\right)$ is strictly increasing in $U_a(\omega, a)$, the following holds:

$$\text{For all } a_1,\ a_2,\ \omega_1,\ \omega_2 \text{ such that } U_a(\omega_1, a_1) < U_a(\omega_2, a_2),$$
$$\frac{V_a(\omega_1, a_1)}{-U_{aa}(\omega_1, a_1)} \leq \frac{V_a(\omega_2, a_2)}{-U_{aa}(\omega_2, a_2)} \Rightarrow \frac{\hat{V}_a(\omega_1, a_1)}{-U_{aa}(\omega_1, a_1)} < \frac{\hat{V}_a(\omega_2, a_2)}{-U_{aa}(\omega_2, a_2)}.$$

Therefore, if condition (3) holds with V, it holds with strict inequality with $\hat{V}$. For sufficiently small γ, the expected utility of the sender with V and $\hat{V}$ differ in absolute value by less than $\Delta/2$ no matter the communication scheme, and hence full disclosure remains suboptimal with $\hat{V}$. But we have shown above that full disclosure is optimal when condition (3) holds with strict inequality, a contradiction. ■

Proof of Lemma 2. Let $h(\omega, a) := \dfrac{V_a(\omega, a)}{-U_{aa}(\omega, a)}$. Condition (4) is equivalent to the statement that $h(\omega, a)$ weakly increases in all directions in the $\Omega \times A$ space in which a and $U_a(\omega, a)$ jointly increase. So, let us define a direction through a function $\omega(a)$ and take the full derivative of $h(\omega(a), a)$

with respect to a:

$$\frac{dh}{da} = \frac{-\frac{dV_a}{da}U_{aa} + \frac{dU_{aa}}{da}V_a}{(U_{aa})^2}.$$

We want to show that $\frac{dh}{da} \geq 0$, which is equivalent to

$$\frac{dU_{aa}}{da}V_a - \frac{dV_a}{da}U_{aa} \geq 0, \tag{24}$$

for *all* $\omega(a)$ such that $\dfrac{dU_a}{da} > 0$, i.e., all directions in which U_a increases as well. As $\dfrac{dU_a}{da} = U_{aa} + U_{a\omega}\dfrac{d\omega}{da}$, we have that $\dfrac{dU_a}{da} > 0$ is equivalent to

$$\begin{cases} \dfrac{d\omega}{da} > -\dfrac{U_{aa}}{U_{a\omega}} \text{ if } U_{a\omega} > 0 \\ \dfrac{d\omega}{da} < -\dfrac{U_{aa}}{U_{a\omega}} \text{ if } U_{a\omega} < 0 \end{cases} \tag{25}$$

If $U_{a\omega} = 0$, $\dfrac{dU_a}{da}$ cannot be positive, as $U_{aa} < 0$ by assumption.

Taking into account that

$$\begin{aligned} \frac{dV_a}{da} &= V_{aa} + V_{a\omega}\frac{d\omega}{da} \\ \frac{dU_{aa}}{da} &= U_{aaa} + U_{aa\omega}\frac{d\omega}{da} \end{aligned}$$

inequality (24) becomes

$$\begin{aligned} & U_{aaa}V_a + U_{aa\omega}V_a\frac{d\omega}{da} - \left(V_{aa}U_{aa} + V_{a\omega}U_{aa}\frac{d\omega}{da}\right) \\ \equiv\ & U_{aaa}V_a - V_{aa}U_{aa} + (U_{aa\omega}V_a - V_{a\omega}U_{aa})\frac{d\omega}{da} \geq 0. \end{aligned} \tag{26}$$

Consider first the case when $U_{a\omega} > 0$. Then, the necessary and sufficient conditions for (26) to hold for *all* $\omega(a)$ such that $\frac{dU_a}{da} > 0$, given that by (25) $\frac{d\omega}{da}$ can take *all* values above $-\frac{U_{aa}}{U_{a\omega}}$, are the following:

$$\begin{cases} U_{aa\omega}V_a - V_{a\omega}U_{aa} \geq 0 \\ U_{aaa}V_a - V_{aa}U_{aa} - (U_{aa\omega}V_a - V_{a\omega}U_{aa})\dfrac{U_{aa}}{U_{a\omega}} \geq 0 \end{cases},$$

which becomes

$$\begin{cases} U_{aa\omega}V_a \geq V_{a\omega}U_{aa} \\ V_a(U_{aaa}U_{a\omega} - U_{aa\omega}U_{aa}) \geq U_{aa}(V_{aa}U_{a\omega} - V_{a\omega}U_{aa}) \end{cases}. \tag{27}$$

Consider now the case when $U_{a\omega} < 0$. Then, following the same steps we get

$$\begin{cases} U_{aa\omega}V_a \leq V_{a\omega}U_{aa} \\ V_a(U_{aaa}U_{a\omega} - U_{aa\omega}U_{aa}) \leq U_{aa}(V_{aa}U_{a\omega} - V_{a\omega}U_{aa}) \end{cases},$$

■

Proof of Theorem 2. Suppose that there exists a pair of states ω_1, ω_2 with $a_1^*(\omega_1) < a_2^*(\omega_2)$ such that, for all $a_1,\ a_2$ satisfying $a_1 < a_2$ and $U_a(\omega_1, a_1) < 0 < U_a(\omega_2, a_2)$,

$$\frac{V_a(\omega_1, a_1)}{-U_{aa}(\omega_1, a_1)} > \frac{V_a(\omega_2, a_2)}{-U_{aa}(\omega_2, a_2)}.$$

For such states, (18) does not hold, and hence (15) does not hold, which means that pooling those states is better than revealing them. Then, by continuity, pooling intervals whose interiors contain (respectively) ω_1 and ω_2 is better than revealing the states in the intervals. Since ω_1 and ω_2 are in the support of the prior, such intervals have positive measure, and thus full disclosure is suboptimal. ■

References

[1] Arieli I., Y. Babichenko, R. Smorodinsky, and T. Yamashita, "Optimal persuasion via bi-pooling," working paper, available at https://ssrn.com/abstract=3511516.

[2] Crawford, V.P., and J. Sobel, 1982, "Strategic information transmission," *Econometrica*, 50(6): 1431-51.

[3] Dizdar D., and E. Kováč, 2020, "A simple proof of strong duality in the linear persuasion problem," *Games and Economic Behavior*, 122: 407-12.

[4] Dworczak, P., and A. Kolotilin, 2019, "The persuasion duality," working paper, available at https://ssrn.com/abstract=3474376.

[5] Dworczak P, and G. Martini, 2019, "The simple economics of optimal persuasion," *Journal of Political Economy*, 127(5): 1993-2048.

[6] Gentzkow M, and E. Kamenica, 2016, "A Rothschild-Stiglitz approach to Bayesian persuasion," *American Economic Review*, 106(5): 597-601.

[7] Kamenica E., and M. Gentzkow, 2011, "Bayesian persuasion," *American Economic Review*, 101(6): 2590-615.

[8] Kolotilin, A., 2018, "Optimal information disclosure: A linear programming approach," *Theoretical Economics*, 13(2) :607-35.

[9] Kolotilin, A., R. Corrao, and A. Wolitzky, 2022, "Persuasion with non-linear preferences," UNSW Economics Working Paper 2022-03, available at https://ssrn.com/abstract=4204824.

[10] Kolotilin, A., T. Mylovanov, and A. Zapechelnyuk, 2021, "Censorship as optimal persuasion," working paper.

[11] Mensch J., 2021, "Monotone persuasion," *Games and Economic Behavior*, 130: 521-42.

[12] Rayo L, and I. Segal, 2010, "Optimal information disclosure," *Journal of Political Economy*, 118(5): 949-87.

Катонини, Э., Степанов, С.

Об оптимальности полного раскрытия информации* [Электронный ресурс] : препринт WP9/2023/02 / Э. Катонини, С. Степанов ; Нац. исслед. ун-т «Высшая школа экономики». – Электрон. текст. дан. (295 Кб). – М. : Изд. дом Высшей школы экономики, 2023. – (Серия WP9 «Исследования по экономике и финансам»). – 44 с. (На англ. яз.)

Мы рассматриваем задачу, в которой отправитель информации способен придерживаться любой политики раскрытия информации получателю информации, и действия последнего влияют на полезность отправителя. Мы получаем достаточное условие оптимальности полного раскрытия информации, обладающее рядом желательных свойств. Во-первых, наше условие записывается в терминах прямых предпочтений игроков, а не в терминах неявной функции полезности отправителя; это делает его легко интерпретируемым и проверяемым. Во-вторых, наша модель не требует, чтобы выигрыш отправителя был функцией апостериорного матожидания состояния мира. В-третьих, наше условие слабее условий на функцию полезности отправителя, которые были недавно получены в литературе для случая, когда действие получателя линейно зависит от ожидаемого состояния. Используя наши результаты, мы демонстрируем, что полное раскрытие информации будет оптимальным в ряде стандартных моделей типа принципал – агент.

Эмилиано Катонини, Нью-Йоркский университет в Шанхае; E-mail: ec1838@nyu.edu

Сергей Степанов, НИУ ВШЭ, Международный институт экономики и финансов и факультет экономических наук; Россия, 109028, Москва, Покровский бульвар, д. 11, офис S641; E-mail: sstepanov@hse.ru

*Мы благодарим Антона Колотилина, Михаила Другова и участников семинара по экономической теории НИУ ВШЭ и конгресса ЕЭА-ЭСЕВ 2022 за полезные комментарии.

Препринты Национального исследовательского университета «Высшая школа экономики» размещаются по адресу: http://www.hse.ru/org/hse/wp

Препринт WP9/2023/02
Серия WP9
Исследования по экономике и финансам

Катонини Эмилиано, Степанов Сергей

Об оптимальности полного раскрытия информации
(*на английском языке*)

Публикуется в авторской редакции

Изд. № 2696