

# Persuasion with Verifiable Information<sup>★</sup>

Maria Titova<sup>\*,†</sup>

Kun Zhang<sup>‡</sup>

October 5, 2025

## Abstract

This paper studies a game in which an informed sender with state-independent preferences uses verifiable messages to convince a receiver to choose an action from a finite set. We characterize the equilibrium outcomes of the game and compare them with commitment outcomes in information design. We provide conditions under which a commitment outcome is an equilibrium outcome and identify environments in which the sender does not benefit from commitment power. Our findings offer insights into the interchangeability of verifiability and commitment in applied settings.

JEL CLASSIFICATION: C72; D82; D83

KEYWORDS: Communication, persuasion, disclosure, verifiable information, commitment

---

<sup>★</sup>This paper benefited immensely from the excellent comments and suggestions of the editor, Marzena Rostek, the associate editor, and two anonymous referees. We are indebted to Renee Bowen, Andreas Kleiner, and Joel Sobel for their guidance and support. We also thank Nageeb Ali, Arjada Bardhi, Simone Galperti, Germán Gieczewski, Navin Kartik, Jan Knoepfle, Frédéric Koessler, Remy Levin, Aleksandr Levkun, Elliot Lipnowski, Claudio Mezzetti, Denis Shishkin, Joel Watson, Mark Whitmeyer, as well as numerous seminar and conference audiences, for their helpful feedback. All remaining errors are our own.

<sup>\*</sup>Corresponding author.

<sup>†</sup>Department of Economics and Department of Political Science, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235, USA. Email: [motitova@gmail.com](mailto:motitova@gmail.com).

<sup>‡</sup>School of Economics, University of Queensland, Level 6, Colin Clark Building 39, Brisbane St. Lucia, QLD 4072, Australia. Email: [kun@kunzhang.org](mailto:kun@kunzhang.org).

# 1. INTRODUCTION

Persuasion with verifiable information plays an essential role in many economic settings, including courtrooms, electoral campaigns, product advertising, financial disclosure, and job market signaling. In a courtroom, a prosecutor tries to persuade a judge to convict a defendant by selectively presenting inculpatory evidence. In an electoral campaign, a politician carefully chooses which campaign promises he can credibly make to win over voters. In advertising, a firm convinces consumers to purchase its product by highlighting specific product characteristics. In finance, a CEO discloses certain financial statements and indicators to board members to obtain higher compensation. In a labor market, a job candidate lists specific certifications to make her application more attractive to an employer.

We consider the following model of persuasion with verifiable information. First, the sender (he/him) learns the state of the world. Second, the sender chooses a message, which is a verifiable statement about the state of the world, and sends it to the receiver (she/her). Verifiability requires that any feasible message contain the truth (the true state of the world) but not necessarily the whole truth; the message may also include other states. Upon observing the message, the receiver takes an action from a finite set. The sender's preferences are state-independent and strictly increasing in the receiver's action, whereas the receiver's preferences depend on both her action and the state.

Seminal papers in this literature (e.g., [Grossman, 1981](#), [Milgrom, 1981](#)) establish an “unraveling” result, which states that the sender fully reveals the state in every equilibrium. In these papers, the sender's preferences are strictly monotone in the receiver's action (e.g., he is maximizing quantity sold) and the receiver's action space is rich (e.g., she is choosing a perfectly divisible quantity to buy). The argument goes as follows: the sender who is privately informed about the quality of his product always wants to separate himself from all lower-quality senders, as this separation convinces the receiver to purchase a strictly higher quantity of the product. We note that if the receiver's action space is finite, the sender may not fully reveal the state in every equilibrium. This is easiest to see when the receiver's action space is binary, such as when she is choosing between buying and not buying. In this case high-quality senders may not mind pooling with some lower-quality senders as long as the receiver chooses to buy.

Our first result characterizes (perfect Bayesian) equilibrium outcomes, which we define as mappings from the state space to a distribution over the receiver's actions. In [Theorem 1](#), we show that every equilibrium outcome must be incentive-compatible (for the sender, IC for short) and obedient (for the receiver). We say that an outcome is IC if the sender receives at least his complete information payoff in each state; otherwise,

he has a profitable deviation toward fully revealing the state. Obedience requires that if the receiver takes an action with positive probability in some states, it must maximize her expected utility. The second part of [Theorem 1](#) adds that if an outcome is deterministic, IC, and obedient, then it is an equilibrium outcome. A deterministic outcome is one in which the receiver takes some action with probability one in every state. Although not all equilibrium outcomes are deterministic, we show in [Lemma 1](#) that all equilibria in which the receiver does not mix (e.g., equilibria in which the receiver uses a predetermined tie-breaking rule) induce deterministic outcomes.

In our model, the sender does not have commitment power: he learns the state and then chooses a verifiable message that maximizes his expected payoff in that state. Our second goal is to understand when the sender can achieve the same payoff in equilibrium as he does in information design (e.g., [Kamenica and Gentzkow, 2011](#)). In information design, the sender commits to a disclosure strategy before learning the state; a commitment outcome is an obedient outcome that maximizes the sender’s ex-ante utility. Our second main result ([Theorem 2](#)) states that the commitment payoff is achievable in equilibrium if and only if there exists a commitment outcome that is deterministic and IC. Intuitively, commitment outcomes that are IC but nondeterministic are generally not equilibrium outcomes because in a commitment outcome, the receiver typically breaks ties in favor of the sender-preferred action. However, as we mentioned earlier, all equilibria in which the receiver does not mix induce deterministic outcomes.

To determine when a commitment outcome can be implemented as an equilibrium, we must ask when a deterministic and IC commitment outcome exists. When the state space is rich, we show that a deterministic commitment outcome always exists ([Proposition 2](#)). That commitment outcome is an equilibrium outcome if and only if the sender receives at least his complete information payoff by [Theorem 2](#). When the state space is finite, however, *all* commitment outcomes may be nondeterministic (e.g., in the seminal example of [Kamenica and Gentzkow, 2011](#)). We show that in a modified game in which the set of available verifiable messages is determined stochastically, it is possible to implement any IC commitment outcome (not only a deterministic one; see [Section 5](#)).

Throughout the paper, we illustrate our results for the special case in which the receiver chooses between two actions, a setting commonly used in applications.<sup>1</sup> For

---

<sup>1</sup>See, for example, [Kolotilin \(2015\)](#), in which pharmaceutical companies persuade the U.S. Food and Drug Administration to approve drugs; [Ostrovsky and Schwarz \(2010\)](#) and [Boleslavsky and Cotton \(2015\)](#), in which schools persuade employers to hire their graduates; [Alonso and Câmara \(2016\)](#) and [Bardhi and Guo \(2018\)](#), in which politicians persuade voters; and [Gehlbach and Sonin \(2014\)](#), in which governments persuade citizens.

this case we show that an IC commitment outcome always exists ([Proposition 1](#)). Thus, when the receiver has two actions, verifiability and commitment assumptions are interchangeable when the state space is sufficiently rich ([Propositions 3](#) and [4](#)).

## RELATED LITERATURE

The literature on verifiable disclosure (games in which the sender learns the state and then chooses a message out of a state-dependent message space) was pioneered by [Grossman and Hart \(1980\)](#), [Grossman \(1981\)](#), and [Milgrom \(1981\)](#); this paper uses the same mapping from states to available messages as in [Milgrom and Roberts \(1986\)](#), except in [Section 5](#).<sup>2</sup>

A few recent papers similarly characterize the equilibrium set (or the set of equilibrium payoffs of the sender) and assess the value of commitment in various verifiable disclosure models. [Zhang \(2022\)](#) focuses on a special case of our model, further assuming that the state space is a unit interval, the receiver has monotone preferences, and the receiver’s optimal action only depends on the expected state. Under these assumptions, the information design problem is known to have a bi-pooling solution, which always induces a deterministic commitment outcome. [Zhang \(2022\)](#) provides conditions under which this solution is implementable in equilibrium. [Ali, Kleiner, and Zhang \(2024\)](#) focus on settings in which the sender favors uncertainty: his preferences are state-dependent and deviations to full revelation are never profitable. They provide conditions under which the sets of equilibrium payoffs of the sender are virtually the same in the disclosure game as in information design. [Gieczewski and Titova \(2024\)](#) consider a generalized disclosure game with an arbitrary message mapping and focus on coalition-proof equilibria.

Outside of verifiable disclosure models, our paper also relates to the informed information design (IID) literature pioneered by [Perez-Richet \(2014\)](#), especially [Koessler and Skreta \(2023; KS henceforth\)](#) and [Zapechelnyuk \(2023; Z henceforth\)](#). In IID, the sender chooses a Blackwell experiment like in information design, except he observes the state of the world before making the choice. Therefore, in IID, a sender faces additional incentive-compatibility constraints relative to (uninformed) information design, much like in verifiable disclosure. The key difference between IID and disclosure games is that the sender can use stochastic evidence in IID, whereas his evidence in verifiable disclosure is deterministic. The differences in equilibrium sets between IID and our

---

<sup>2</sup>For detailed surveys of this literature, see, for example, [Milgrom \(2008\)](#) and [Dranove and Jin \(2010\)](#).

model highlight the value of stochastic evidence.<sup>3</sup> In unconstrained IID (KS), an obedient outcome is an equilibrium outcome if and only if it is IC.<sup>4</sup> In IID constrained to nondegenerate experiments (Z), every obedient outcome is an equilibrium outcome. We show that in Milgrom-Roberts’s verifiable disclosure, an obedient outcome is an equilibrium outcome if and only if it is IC and *deterministic* (assuming that the receiver uses a pure strategy, as in KS and Z). Thus, the sender values stochastic evidence when the state space is finite but not when it is rich. An IID problem can also be interpreted as a verifiable disclosure game with random certification (where the randomization between messages is done by a machine, not the sender).<sup>5</sup> We formalize this observation in [Section 5](#) by introducing a verifiable disclosure game with a stochastic message mapping and showing that its equilibrium set is the same as in unconstrained IID. We describe the relationship between our results and those of KS in more detail throughout the paper.

While we study when the sender does not benefit from commitment power, a growing body of literature examines how much the receiver gains from commitment power by comparing equilibrium outcomes with those of optimal mechanisms in sender-receiver games with verifiable information. When the sender’s preferences are state-independent, [Glazer and Rubinstein \(2004, 2006\)](#) and [Sher \(2011\)](#) find that the receiver does not need commitment power to reach the optimal mechanism outcome. [Hart, Kremer, and Perry \(2017\)](#) and [Ben-Porath, Dekel, and Lipman \(2019\)](#) provide conditions under which the equilibrium and optimal mechanism outcomes are equivalent.

[Chakraborty and Harbaugh \(2010\)](#), [Lipnowski and Ravid \(2020\)](#), and [Lipnowski \(2020\)](#) study cheap-talk games in which the sender has state-independent preferences; the latter two compare equilibrium outcomes in one-shot cheap-talk games with commitment outcomes. In cheap-talk games, the sender’s messages are not verifiable: in every state, the sender has access to the same (sufficiently rich) set of messages. The

---

<sup>3</sup>Equilibrium concepts differ across all aforementioned papers; for a direct comparison of our results to those of KS and Z, we use perfect Bayesian equilibrium (PBE) with the refinement of the principle of preeminence of tests, which requires that “every out-of-equilibrium posterior belief must assign probability one to each event that is revealed as certain by the test” (Z, p. 1061). The principle of preeminence of tests rules out non-IC PBE because the receiver learns the state when the sender sends a fully informative experiment, and this deviation must be unprofitable. Note that PBE without refinements has no predictive power in IID, meaning that every obedient outcome is a PBE outcome (KS, p. 3197).

<sup>4</sup>KS focuses on interim optimal (IO) outcomes, which are PBE outcomes with a restriction on R’s off-path beliefs: any such belief must assign positive probability only to states in which the sender strictly benefits from the deviation.

<sup>5</sup>We thank Frédéric Koessler for pointing this out.

verifiability requirement faced by our sender significantly impacts the set of equilibrium outcomes.<sup>6</sup> Kamenica and Lin (2024) show that in standard cheap-talk games with finitely many actions and states (where “standard” means that the receiver is uninformed and the sender has no actions other than the choice of a message), generically the commitment payoff is achieved in an equilibrium if and only if there exists a deterministic commitment outcome. Our Theorem 2 provides a similar result for verifiable disclosure games.

## 2. MODEL

We study a game of persuasion with verifiable information between a sender (S, he/him) and a receiver (R, she/her). Below we describe the timing of the game along with the assumptions:<sup>7</sup>

1. S observes the state of the world,  $\theta \in \Theta$ .  
The state space  $\Theta$  is either finite ( $\Theta = \{1, \dots, N\}$ ,  $N \geq 2$ ) or rich ( $\Theta$  is a convex and compact subset of  $\mathbb{R}^n$ ). The state of the world is drawn from a common prior  $\mu_0 \in \Delta(\Theta)$  with  $\text{supp } \mu_0 = \Theta$ . If the state space is rich, we assume that the prior is atomless.
2. S sends message  $m \in M$  to R, where  $M$  is the collection of nonempty Borel subsets of  $\Theta$ . Since each message is a subset of the state space, we interpret it as a statement about the state of the world. S’s messages are *verifiable* in the sense that every message must contain the truth: the set of messages available to S in state  $\theta \in \Theta$  is  $\{m \in M \mid \theta \in m\}$ .<sup>8</sup>
3. R observes the message (but not the state) and takes an action from a finite set  $J := \{1, \dots, K\}$  with  $K \geq 2$ .
4. The game ends, and payoffs are realized.  
S’s payoff  $v : J \rightarrow \mathbb{R}$  depends only on R’s action. Without loss, we assume that actions are ordered such that  $v$  is increasing in  $j \in J$ . For ease of exposition, we

---

<sup>6</sup>Verifiability of his messages may help or hurt the sender depending on the preferences of the players. In fact, every equilibrium of the verifiable information game may be ex-ante better for the sender than every cheap-talk equilibrium and vice versa.

<sup>7</sup>For a topological space  $Y$ , let  $\Delta(Y)$  denote the set of Borel probability measures on  $Y$ . For  $\gamma \in \Delta Y$ , let  $\text{supp } \gamma$  denote the support of  $\gamma$ . We say that  $\gamma \in \Delta Y$  is degenerate if  $\text{supp } \gamma$  is a singleton, and non-degenerate otherwise.

<sup>8</sup>We borrow from Milgrom and Roberts (1986) the definition of a verifiable message as a subset of the state space that includes the realized state. This method satisfies normality of evidence (Bull and Watson, 2007), which makes it consistent with both major ways of modeling hard evidence in the literature.

also assume that  $v$  is strictly increasing.

R's preferences are described by a bounded measurable utility function  $u : J \times \Theta \rightarrow \mathbb{R}$ . We define R's *complete information action- $j$  set* as  $A_j := \{\theta \in \Theta \mid u(j, \theta) \geq u(j', \theta) \text{ for all } j' \in J\}$  to include all the states of the world in which she prefers to take action  $j$  under complete information.

We consider perfect Bayesian equilibria (henceforth equilibria) of this game. First, S's strategy is a function  $\sigma : \Theta \rightarrow \Delta_0 M$ , where  $\Delta_0 M$  is the set of probability measures on  $M$  with a finite support.<sup>9</sup> Second, R's strategy is a function  $\tau : M \rightarrow \Delta J$ . Finally, R's belief system  $q : M \rightarrow \Delta \Theta$  describes R's beliefs about the state after any observed message.

DEFINITION 1. A triple  $(\sigma, \tau, q)$  is an equilibrium if

- (i) for all  $\theta \in \Theta$ ,  $\sigma(\cdot \mid \theta)$  is supported on  $\arg \max_{\{m \in M \mid \theta \in m\}} \sum_{j \in J} v(j) \tau(j \mid m)$ ;
- (ii) for all  $m \in M$ ,  $\tau(\cdot \mid m)$  is supported on  $\arg \max_{j \in J} \int_{\Theta} u(j, \theta) dq(\theta \mid m)$ ;
- (iii)  $q$  is obtained from  $\mu_0$ , given  $\sigma$ , using Bayes' rule whenever possible;<sup>10</sup>
- (iv) for all  $m \in M$ ,  $q(\cdot \mid m) \in \Delta m$ .

In words, in equilibrium, (i) S chooses verifiable messages that maximize his expected utility in every state  $\theta \in \Theta$ ; (ii) R maximizes her expected utility given her posterior belief; (iii) R uses Bayes' rule to update her beliefs whenever possible; and (iv) R's posteriors are consistent with disclosure on and off the path.

To analyze the model, we use the following approach. Let  $\Psi$  be the set of all Borel measurable functions from  $\Theta$  to  $\Delta J$ . We refer to any  $\alpha \in \Psi$  as an *outcome*; it specifies, for each state  $\theta \in \Theta$ , the probability  $\alpha(j \mid \theta)$  that R takes action  $j \in J$ . Given a pair of strategies  $(\sigma, \tau)$  of S and R, we let  $M_j(\sigma, \tau) := \{m \in M \mid m \in \text{supp } \sigma(\cdot \mid \theta) \text{ for some } \theta \in \Theta \text{ and } \tau(j \mid m) > 0\}$  be the set of messages that convince R to take action  $j \in J$ , sent with a positive probability in some state  $\theta \in \Theta$ . We say that  $\alpha \in \Psi$  is an *equilibrium outcome* if there exists an equilibrium  $(\sigma, \tau, q)$  that *induces* it, meaning that  $\alpha(j \mid \theta) = \sum_{m \in M_j(\sigma, \tau)} \tau(j \mid m) \sigma(m \mid \theta)$  for all  $j \in J$  and  $\theta \in \Theta$ .

We say that an outcome  $\alpha \in \Psi$  is *deterministic* if  $\alpha(\cdot \mid \theta)$  is degenerate for each  $\theta \in \Theta$ . For a deterministic outcome  $\alpha$ , we refer to the collection of sets  $\{W_j\}_{j \in J}$ , where

<sup>9</sup>That is, we assume that S mixes between finitely many messages. This assumption imposes no restriction when  $\Theta$  is finite. When  $\Theta$  is rich, it guarantees that  $\sigma(\cdot \mid \theta)$  is well-defined, and the restriction does not affect the set of achievable equilibrium payoffs.

<sup>10</sup>That is,  $q$  is a regular conditional probability system.



$W_j := \{\theta \in \Theta \mid \alpha(j \mid \theta) = 1\}$ , as the *outcome partition* (into subsets  $W_j$  of the state space in which R takes action  $j \in J$  with probability one) of  $\alpha$ .

Given an outcome  $\alpha$ , we let  $v_\alpha(\theta) := \sum_{j \in J} v(j) \alpha(j \mid \theta)$  be S's interim (expected) payoff in state  $\theta \in \Theta$  and  $V_\alpha := \int_{\Theta} v_\alpha(\theta) d\mu_0(\theta)$  be S's ex-ante utility.

### 3. EQUILIBRIUM ANALYSIS

We begin by establishing the lower bound on S's payoff in an equilibrium outcome  $\alpha$ . One thing that S can do in state  $\theta$  is fully reveal it by sending message  $\{\theta\}$  with probability one. Upon receiving message  $\{\theta\}$ , R learns that the state is  $\theta$  and takes an action that is a best response under complete information. Thus, S's equilibrium payoff in state  $\theta$  is bounded below by  $\underline{v}(\theta) := \min_{j \in J \text{ s.t. } \theta \in A_j} v(j)$ . We refer to this condition as S's IC constraint:<sup>11</sup>

$$v_\alpha(\theta) \geq \underline{v}(\theta). \quad (\text{IC}_\theta)$$

DEFINITION 2. An outcome  $\alpha$  is incentive-compatible (IC) if it satisfies  $(\text{IC}_\theta)$  for each state  $\theta \in \Theta$ .

Next, in equilibrium, if R finds it optimal to play action  $j$  after several messages, that action must remain optimal even if R does not know which of these messages was sent. We can thus “bundle” all these messages into a single “recommendation,” giving rise to R's obedience constraint for action  $j$ :

$$\int_{\Theta} (u(j, \theta) - u(j', \theta)) \alpha(j \mid \theta) d\mu_0(\theta) \geq 0, \quad \text{for all } j' \in J. \quad (\text{obedience}_j)$$

DEFINITION 3. An outcome  $\alpha$  is obedient if it satisfies  $(\text{obedience}_j)$  for each action  $j \in J$ .

If  $\alpha$  is a deterministic outcome with partition  $\{W_j\}_{j \in J}$ , then  $(\text{IC}_\theta)$  becomes  $\theta \in W_j \implies v(j) \geq \underline{v}(\theta) \iff j \geq \min_{i \in J \text{ s.t. } \theta \in A_i} i$ , indicating that the action taken in state  $\theta$  must be no lower than R's worst best response under complete information. The obedience constraint for action  $j$  simplifies to  $\int_{W_j} (u(j, \theta) - u(j', \theta)) d\mu_0(\theta) \geq 0$  for all  $j' \in J$ .

---

<sup>11</sup>In fact,  $\underline{v}(\theta)$  is the lower bound on S's equilibrium payoff in state  $\theta$ , meaning there exists an equilibrium in which S's interim payoff is exactly  $\underline{v}(\theta)$  for each  $\theta \in \Theta$ . In this equilibrium, S fully reveals every state, R takes the lowest action that is a best response under complete information, and R's beliefs are skeptical off-path (we define R's skeptical beliefs in the proof of Theorem 1).



Our first result confirms that every equilibrium outcome is IC and obedient. For deterministic outcomes, these two properties are necessary and sufficient for equilibrium implementation.

THEOREM 1.

- (a) Every equilibrium outcome is IC and obedient.
- (b) If a deterministic outcome is IC and obedient, then it is an equilibrium outcome.

*Part (a).* Consider an equilibrium  $(\sigma, \tau, q)$  with outcome  $\alpha \in \Psi$ . Observe that  $\alpha$  must be IC, or else there exists a state  $\theta$  in which S has a profitable deviation to fully revealing the state. Next, we show that  $\alpha$  is also obedient. Consider any action  $j \in J$ . By the equilibrium condition (ii), we have

$$\begin{aligned} \text{for all } m \in M_j(\sigma, \tau) \text{ and } j' \in J, \quad & \int_{\Theta} (u(j, \theta) - u(j', \theta)) \, dq(\theta \mid m) \geq 0 \\ \implies & \int_{\Theta} (u(j, \theta) - u(j', \theta)) \tau(j \mid m) \, dq(\theta \mid m) \geq 0, \end{aligned}$$

where the second inequality follows because  $\tau(j \mid m) > 0$  for all  $m \in M_j(\sigma, \tau)$ . Using Bayes' rule, the above inequality implies that

$$\begin{aligned} \text{for all } j' \in J, \quad & \int_{\Theta} (u(j, \theta) - u(j', \theta)) \sum_{m \in M_j(\sigma, \tau)} \tau(j \mid m) \sigma(m \mid \theta) \, d\mu_0(\theta) \geq 0, \\ \implies & \int_{\Theta} (u(j, \theta) - u(j', \theta)) \alpha(j \mid \theta) \, d\mu_0(\theta) \geq 0, \end{aligned}$$

where the last inequality is (obedience<sub>j</sub>). Since  $j$  was chosen arbitrarily,  $\alpha$  is obedient.

[Part (b)] Consider a deterministic outcome  $\alpha$  that is IC and obedient, and denote its outcome partition by  $\{W_j\}_{j \in J}$ . We construct an equilibrium  $(\sigma, \tau, q)$  that induces  $\alpha$ . Let S's strategy be  $\sigma(m \mid \theta) = \mathbb{1}(m = W_j \text{ and } \theta \in W_j)$ , which reveals the element of the outcome partition that the realized state belongs to. When R receives an on-path message  $W_j$ , she learns that  $\theta \in W_j$  and nothing else; by (obedience<sub>j</sub>), playing action  $j$  is a best response; thus, we let  $\tau(j \mid W_j) = 1$  for all  $j \in J$ . For off-path messages, assume R is "skeptical" and believes that any unexpected message comes from the state in which R prefers to take the lowest action under complete information. Formally, for all  $m \notin \{W_j\}_{j \in J}$ , let  $q(\cdot \mid m) \in \Delta(m \cap A_{\underline{j}})$ , where  $\underline{j} \in J$  is the lowest action  $i \in J$  such that the set  $m \cap A_i$  is nonempty. Then playing action  $\underline{j}$  with probability one is a best response to message  $m$ , so we let  $\tau(\underline{j} \mid m) = 1$ .

We now show that S has no profitable deviations using the fact that  $\{W_j\}_{j \in J}$  is a partition of the state space. Consider a state  $\theta \in \Theta$ , which is in  $W_j$  for some action  $j \in J$ . S cannot send any other on-path message because  $\theta \in W_j$  implies  $\theta \notin W_i$  for any  $i \neq j$ . Therefore  $W_i$  is not a verifiable message in state  $\theta$ . If S deviates to an off-path (verifiable) message  $m \notin \{W_j\}_{j \in J}$ , then S's payoff is  $v(\underline{j}) \leq \underline{v}(\theta)$ , and this deviation is unprofitable by (IC $_\theta$ ). Therefore,  $(\sigma, \tau, q)$  is an equilibrium that induces  $\alpha$ . ■

Part (b) of Theorem 1 characterizes the set of deterministic equilibrium outcomes, and its proof suggests a simple way of implementing these outcomes in a pure-strategy equilibrium with at most  $K$  on-path messages that essentially serve as action recommendations. Specifically, if  $\{W_j\}_{j \in J}$  is an outcome partition, then  $W_j$  serves as both the set of states in which R plays action  $j$  and the on-path message recommending action  $j$  in the constructed equilibrium inducing this outcome.

While Theorem 1 fully characterizes the set of deterministic equilibrium outcomes, it does not provide a full characterization of the entire set of equilibrium outcomes. In general, an IC and obedient nondeterministic outcome may or may not be an equilibrium outcome. Consider the seminal example from Kamenica and Gentzkow (2011).

EXAMPLE 1. Suppose S is a prosecutor and R is a judge. The state of the world is binary:  $\Theta = \{1, 2\} = \{\text{innocent}, \text{guilty}\}$ ; R's action space is binary:  $J = \{1, 2\} = \{\text{acquit}, \text{convict}\}$ ; and the prior is  $\mu_0(1) = 0.7$ . S's preferences are  $v(1) = 0$  and  $v(2) = 1$ , and R's objective is to "match the state":  $u(1, 1) = u(2, 2) = 1$ , and  $u(1, 2) = u(2, 1) = 0$ . Consider an outcome  $\alpha^*$  in which  $\alpha^*(2 | 2) = 1$  and  $\alpha^*(2 | 1) = 3/7$ . It is easy to verify that  $\alpha^*$  is both IC and obedient. However,  $\alpha^*$  is not an equilibrium outcome: when  $\theta = 1$ , R convicts with probability  $3/7$  and acquits with probability  $4/7$ . Since S strictly prefers conviction, he has a profitable deviation to sending the message after which R convicts when  $\theta = 1$ .

Example 1 illustrates that (IC and obedient) outcomes in which S receives different payoffs from different messages in the same state cannot be equilibrium outcomes. Once the state is realized, S's message space becomes fixed and known. Thus, if S mixes between multiple messages in the same state, he must receive the same payoff from each of these messages. Of course, if S does receive the same payoff in every state, then an IC, obedient, and nondeterministic outcome could be an equilibrium outcome. However, in any such equilibrium, R must play a mixed strategy:

LEMMA 1. *Suppose that  $\alpha$  is a nondeterministic outcome induced by an equilibrium  $(\sigma, \tau, q)$ . Then in each state  $\theta \in \Theta$  such that  $\alpha(\cdot | \theta)$  is nondegenerate, R is playing a mixed strategy (meaning  $\tau(\cdot | m)$  is nondegenerate) for some  $m \in \text{supp } \sigma(\cdot | \theta)$ .*

*Proof.* Let  $\theta \in \Theta$  be a state such that  $\alpha(\cdot \mid \theta)$  is nondegenerate. By contradiction, suppose that  $\tau(\cdot \mid m)$  is degenerate for all  $m \in \text{supp } \sigma(\cdot \mid \theta)$ . By equilibrium condition (i), for any pair of messages  $m, m' \in \text{supp } \sigma(\cdot \mid \theta)$ , we have  $\sum_{j \in J} v(j) \tau(j \mid m) = \sum_{j \in J} v(j) \tau(j \mid m')$ , implying that there exists an action  $j^* \in J$  such that  $\tau(j^* \mid m) = \tau(j^* \mid m') = 1$ . In other words, if R is not mixing, every message sent by S in state  $\theta$  leads R to take the same action. Therefore,  $\alpha(j^* \mid \theta) = \sum_{m \in M_j(\sigma, \tau)} \tau(j^* \mid m) \sigma(m \mid \theta) = 1$ , which is a contradiction. ■

The contrapositive of Lemma 1 also tells us that if R is not mixing in an equilibrium (e.g., if she uses an exogenously given tie-breaking rule like in IID), then an obedient outcome is an equilibrium outcome *if and only if* it is IC and deterministic. Theorem 1 and Lemma 1 together highlight the difference in equilibrium sets between our verifiable disclosure game and IID (KS and Z). KS’s characterization (Proposition 2) states that an outcome is interim optimal (IO) if and only if it is obedient and IOC, where IOC essentially requires that for every set of states  $Q$ , and for every state in  $Q$ , S does not strictly prefer R having a belief supported on  $Q$ .<sup>12</sup> Naturally, the first difference—IOC in KS’s setting versus IC in ours—arises from the difference in equilibrium selection, as they impose a stronger restriction on off-path beliefs than we do. The second difference is that IOC and obedience are necessary and sufficient for an outcome to be IO, whereas for us IC and obedience are not sufficient. Since in our model S chooses messages, an additional restriction applies: S can mix between different messages only if each message yields the same expected payoff—a constraint absent in IID. For this reason, some nondeterministic IO, and thus IC, outcomes are not equilibrium outcomes in our game (e.g., one from Example 1).

Theorem 1 characterizes all pure-strategy equilibria of the game, as these equilibria are deterministic. Koessler and Renault (2012) find that IC and obedience are necessary and sufficient for a pure-strategy outcome to be an equilibrium outcome in a setting in which S has state-independent preferences, sends verifiable messages, and sets a price and R chooses between two actions. Theorem 1 highlights that this result (1) extends to cases in which R has more than two actions and (2) is not driven by S’s additional choice variable (price).<sup>13</sup>

---

<sup>12</sup>In contrast, IC only requires that in any given state, S does not strictly prefer inducing the degenerate belief at that state. Interestingly, KS also show that IOC and IC are equivalent if S’s value function is quasiconvex in R’s belief (Proposition 3) or if R chooses between two actions (Lemma B.2).

<sup>13</sup>However, as the authors point out, the price choice in their setting ensures that R plays a pure strategy in equilibrium.

## 4. VALUE OF COMMITMENT

In this section, we ask when a commitment outcome, a solution to the information design problem, is also an equilibrium outcome. In the information design problem, Stage 1 of the game (in which S learns the state) is removed, and Stage 2 of the game (in which S chooses a verifiable message) is replaced by S committing to an experiment that sends signals depending on state realizations.<sup>14</sup> Importantly, when S has commitment power, he no longer faces incentive-compatibility constraints, i.e., he does not need to maximize his utility state by state.

Following [Kamenica and Gentzkow \(2011\)](#), we focus on straightforward signals that R interprets as action recommendations. Therefore, an (optimal) *commitment outcome*  $\bar{\psi} \in \Psi$  solves

$$\begin{aligned} \max_{\psi \in \Psi} V_\psi \quad & \text{subject to, for each action } j \in J, \\ & \int_{\Theta} (u(j, \theta) - u(j', \theta)) \psi(j \mid \theta) d\mu_0(\theta) \geq 0 \quad \text{for all } j' \in J. \end{aligned} \quad (\text{CO})$$

Simply put, a commitment outcome is an obedient outcome that maximizes S's ex ante utility. We refer to the value of problem (CO) as the *commitment payoff*. Our second result shows that a commitment outcome must be deterministic and IC to be an equilibrium outcome.

**THEOREM 2.** *Consider a commitment outcome  $\bar{\psi} \in \Psi$ .*

- (a) *If  $\bar{\psi}$  is IC and deterministic, then it is an equilibrium outcome.*
- (b) *If  $\bar{\psi}$  is an equilibrium outcome, then it is IC and  $\mu_0$ -almost everywhere deterministic.*

*Proof.* [Part (a)] Recall that every commitment outcome is obedient. Therefore, if  $\bar{\psi}$  is IC and deterministic, Part (b) of [Theorem 1](#) implies that it is an equilibrium outcome.

[Part (b)] Suppose a commitment outcome  $\bar{\psi}$  is an equilibrium outcome, meaning that there exists an equilibrium  $(\sigma, \tau, q)$  that induces it. By [Theorem 1](#),  $\bar{\psi}$  is IC. We will now show that  $\bar{\psi}$  is deterministic  $\mu_0$ -almost everywhere. Define  $T := \{\theta \in \Theta \mid \bar{\psi}(\cdot \mid \theta) \text{ is nondegenerate}\}$  as the set of states in which R plays multiple actions, and suppose, by contradiction, that  $\mu_0(T) > 0$ . By [Lemma 1](#), for each  $\theta \in T$ , there

---

<sup>14</sup>An experiment  $(\mathcal{S}, \chi)$  consists of a compact metrizable space  $\mathcal{S}$  of signals and a Borel measurable function  $\chi : \Theta \rightarrow \Delta \mathcal{S}$ . R observes the choice of the experiment and a signal realization  $s \in \mathcal{S}$  drawn from  $\chi(\cdot \mid \theta)$ , where  $\theta$  is the realized state.

exists a message  $m \in \text{supp } \sigma(\cdot | \theta)$  such that  $\tau(\cdot | m)$  is nondegenerate. Let  $\widetilde{M} := \{m \in M \mid \tau(\cdot | m) \text{ is nondegenerate}\}$  be the set of messages after which R plays a mixed strategy. Define  $\widetilde{\tau}(j^* | m) := \mathbb{1}(j^* = \max_{j \in \text{supp } \tau(\cdot | m)} j)$  for all  $m \in M$  as R's strategy that breaks all ties in  $\tau$  in favor of S. Denote the outcome from the strategy profile  $(\sigma, \widetilde{\tau})$  by  $\widetilde{\psi}$ .

We derive a contradiction by showing that  $\widetilde{\psi}$  is an obedient outcome with  $V_{\widetilde{\psi}} > V_{\overline{\psi}}$ , which implies that  $\overline{\psi}$  is not a commitment outcome. Indeed, we have  $v_{\widetilde{\psi}}(\theta) > v_{\overline{\psi}}(\theta)$  for all  $\theta \in T$  (since there exists an  $m \in \widetilde{M}$  with  $\sigma(m | \theta) > 0$ ), while  $v_{\widetilde{\psi}}(\theta) = v_{\overline{\psi}}(\theta)$  for all  $\theta \notin T$ . Therefore,  $V_{\widetilde{\psi}} - V_{\overline{\psi}} = \int_T (v_{\widetilde{\psi}}(\theta) - v_{\overline{\psi}}(\theta)) d\mu_0(\theta) > 0$  since  $\mu_0(T) > 0$ . To prove that  $\widetilde{\psi}$  is obedient, we apply equilibrium condition (ii) to the equilibrium  $(\sigma, \tau, q)$  and follow the steps in the proof of [Theorem 1\(a\)](#), replacing  $\tau$  with  $\widetilde{\tau}$  and noting that  $M_j(\sigma, \widetilde{\tau}) \subseteq M_j(\sigma, \tau)$ . ■

The nontrivial part of [Theorem 2](#) involves proving that if  $\overline{\psi}$  is both an equilibrium outcome and a commitment outcome, then it is deterministic almost everywhere. This is equivalent to showing that a nondeterministic equilibrium outcome cannot be a commitment outcome. Indeed, by [Lemma 1](#), in the equilibrium that induces  $\overline{\psi}$ , R must play a mixed strategy following some on-path messages from a positive measure of states. However, breaking those ties in favor of the S-preferred action strictly increases S's ex-ante utility, which implies that  $\overline{\psi}$  is not a commitment outcome.

In many relevant settings, R chooses between two actions. In this case, the analysis vastly simplifies. From R's perspective, there are “bad” states  $\theta \in A_1$ , in which R prefers the low action 1, and “good” states  $\theta \notin A_1$ , in which she prefers the high action 2. The highest payoff that S can achieve is  $v(2)$  (when R takes action 2 with probability one), and the lowest is  $v(1)$ . To state that an outcome  $\psi \in \Psi$  is IC, it suffices to show that  $\theta \notin A_1$  implies that  $v_\psi(\theta) = v(2)\psi(2 | \theta) + v(1)\psi(1 | \theta) \geq v(2)$ , which is equivalent to  $\psi(2 | \theta) = 1$ . The IC condition for  $\theta \in A_1$  is not relevant because  $v(1)$  is already the lowest payoff in the game. In words, an outcome is IC if and only if R plays action 2 with probability one in all states in which action 2 is the unique best response under complete information. The following result establishes the existence of an IC commitment outcome when R chooses between two actions.

**PROPOSITION 1.** *If  $|J| = 2$ , then there exists an IC commitment outcome.*

*Proof.* Since  $\Theta$  is a compact subset of  $\mathbb{R}^n$ , a commitment outcome exists by Proposition 3 in the online appendix of [Kamenica and Gentzkow \(2011\)](#) and Theorem 1 in [Terstiege and Wasser \(2023\)](#). Let  $\overline{\psi} \in \Psi$  be a commitment outcome and let  $\widetilde{\psi} \in \Psi$  be an outcome such that  $\widetilde{\psi}(\cdot | \theta) = \overline{\psi}(\cdot | \theta)$  for all  $\theta \notin A_2$  and  $\widetilde{\psi}(2 | \theta) = 1$  for all  $\theta \in A_2$ . By

construction,  $\tilde{\psi}$  is IC and weakly increases S's ex-ante utility over  $\bar{\psi}$ . Define  $\delta(\theta) := u(2, \theta) - u(1, \theta)$  and observe that

$$\int_{\Theta} \delta(\theta) \tilde{\psi}(2 | \theta) d\mu_0(\theta) = \int_{\Theta} \delta(\theta) \bar{\psi}(2 | \theta) d\mu_0(\theta) + \int_{A_2} \delta(\theta) (1 - \bar{\psi}(2 | \theta)) d\mu_0(\theta),$$

where the last term is nonnegative because  $\delta(\theta) \geq 0$  for all  $\theta \in A_2$ . Consequently, obedience of  $\bar{\psi}$  (for both actions) implies obedience of  $\tilde{\psi}$ . Hence,  $\tilde{\psi}$  is also a commitment outcome.  $\blacksquare$

The existing literature provides additional insights into commitment outcomes when  $|J| = 2$  and  $\Theta$  is finite. [Alonso and Câmara \(2016\)](#) show that every commitment outcome is characterized by a cutoff state  $\theta^*$ , with all states satisfying  $\delta(\theta) > \delta(\theta^*)$  pooled together to recommend action 2. In particular, in all good states  $\theta \notin A_1$ , S recommends action 2, which implies that every commitment outcome is IC (see also Lemma B.2 in [Koessler and Skreta, 2023](#)). Our [Proposition 1](#) also addresses the case in which  $\Theta$  is rich. In this case, some commitment outcomes are not IC (although they are IC  $\mu_0$ -almost everywhere), and its proof outlines how to make an existing commitment outcome incentive-compatible.

Returning to the more general case in which  $J \geq 2$ , [Theorem 2](#) is useful for verifying whether an existing commitment outcome  $\bar{\psi}$  is an equilibrium outcome. The answer is affirmative if and only if  $\bar{\psi}$  is deterministic  $\mu_0$ -a.e. and IC. Although verifying incentive compatibility may be straightforward, a deterministic commitment outcome is not guaranteed to exist. In the remainder of this section, we consider the cases in which  $\Theta$  is rich and  $\Theta$  is finite separately. We show that when  $\Theta$  is rich, a deterministic commitment outcome always exists. Furthermore, if  $|J| = 2$ , the commitment payoff is always attained in equilibrium. When  $\Theta$  is finite, we derive an approximation result.

#### 4.1. RICH STATE SPACE

When the state space  $\Theta$  is rich (a convex and compact subset of  $\mathbb{R}^n$ ) and the prior  $\mu_0$  is atomless, the existence of a deterministic commitment outcome is guaranteed.

**PROPOSITION 2.** *If  $\Theta$  is rich, then a deterministic commitment outcome exists. Furthermore, a deterministic commitment outcome is an equilibrium outcome if and only if it is IC.*

*Proof.* The existence of a commitment outcome  $\bar{\psi}$  follows from the same argument as that used in the proof of [Proposition 1](#). Furthermore,  $\bar{\psi}(j | \cdot) : \Theta \rightarrow [0, 1]$  is Borel

measurable for every  $j \in J$  and  $\sum_{j \in J} \bar{\psi}(j \mid \theta) = 1$  for all  $\theta \in \Theta$ . Let  $\mu_j$  be such that  $d\mu_j := u(j, \cdot) d\mu_0$  for each  $j \in J$ .

Since  $\mu_0$  is a finite and atomless positive measure and  $u$  is bounded,  $\mu_j$  is a finite and atomless signed measure for each  $j \in J$ . By Theorem 2.1 in [Dvoretzky, Wald, and Wolfowitz \(1951\)](#), since  $J$  is finite, there exist Borel measurable functions  $\tilde{\psi}(j \mid \cdot) : \Theta \rightarrow \{0, 1\}$  for all  $j \in J$ , with  $\sum_{j \in J} \tilde{\psi}(j \mid \cdot) = 1$ , such that (I)  $\int_{\Theta} \tilde{\psi}(j \mid \theta) d\mu_0 = \int_{\Theta} \bar{\psi}(j \mid \theta) d\mu_0$  and (II)  $\int_{\Theta} \tilde{\psi}(j \mid \theta) d\mu_j = \int_{\Theta} \bar{\psi}(j \mid \theta) d\mu_j$  for all  $j \in J$ . Condition (I) implies that

$$V_{\tilde{\psi}} = \int_{\Theta} \sum_{j \in J} v(j) \tilde{\psi}(j \mid \theta) d\mu_0 = \int_{\Theta} \sum_{j \in J} v(j) \bar{\psi}(j \mid \theta) d\mu_0 = V_{\bar{\psi}}.$$

Condition (II) implies that  $\tilde{\psi}$  is obedient, as  $\bar{\psi}$  is. Hence,  $\tilde{\psi}$  is a deterministic commitment outcome. The second part follows from [Theorem 2](#). ■

Verifying whether a deterministic commitment outcome with partition  $\{W_j\}_{j \in J}$  is IC (and therefore an equilibrium outcome) is straightforward. It requires determining whether  $\theta \in W_j$  implies  $v(j) \geq \underline{v}(\theta)$  for all  $\theta \in \Theta$ . Consider the following example from [Gentzkow and Kamenica \(2016\)](#).

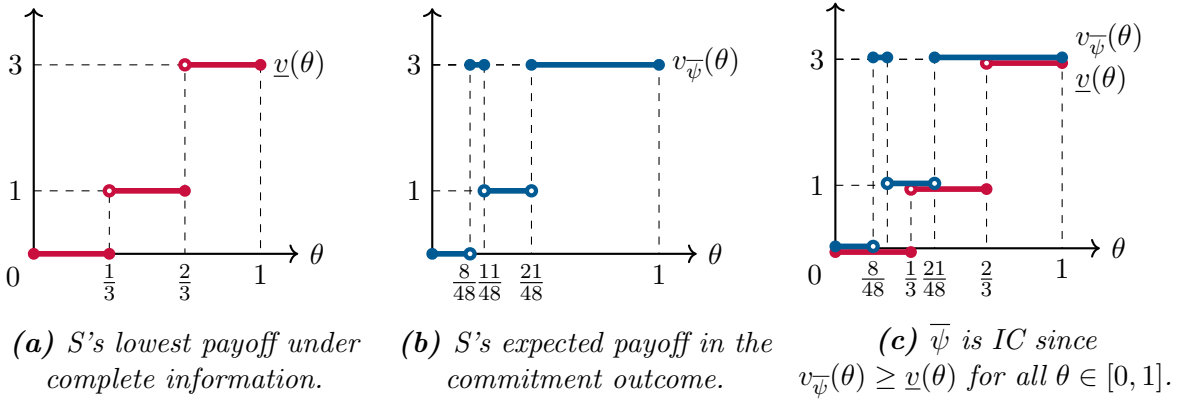
**EXAMPLE 2.** Suppose  $R$  has three actions,  $J = \{1, 2, 3\}$ , and the prior is uniform on  $\Theta = [0, 1]$ .  $S$ 's payoffs are given by  $v(1) = 0$ ,  $v(2) = 1$ , and  $v(3) = 3$ .  $R$ 's preferences depend only on the posterior mean. Given belief  $\mu \in \Delta\Theta$ , action 1 is optimal if and only if  $\mathbb{E}_{\mu}[\theta] \leq 1/3$ ; action 2 is optimal if and only if  $\mathbb{E}_{\mu}[\theta] \in [1/3, 2/3]$ ; and action 3 is optimal if and only if  $\mathbb{E}_{\mu}[\theta] \geq 2/3$ . Therefore,  $R$ 's complete-information action sets are  $A_1 = [0, 1/3]$ ,  $A_2 = [1/3, 2/3]$ , and  $A_3 = [2/3, 1]$ . [Gentzkow and Kamenica \(2016\)](#) identify a deterministic commitment outcome  $\bar{\psi}$  with an outcome partition  $\bar{W}_1 = [0, 8/48)$ ,  $\bar{W}_2 = (11/48, 21/48)$ , and  $\bar{W}_3 = [8/48, 11/48] \cup [21/48, 1]$ . This outcome is IC, which we illustrate in [Figure 1](#). Since  $\bar{\psi}$  is a deterministic and IC commitment outcome, it is an equilibrium outcome by [Proposition 2](#).

When  $R$  chooses between two actions,  $S$  always attains his commitment payoff in equilibrium.

**PROPOSITION 3.** *If  $\Theta$  is rich and  $|J| = 2$ , then there exists a commitment outcome that is an equilibrium outcome.*

*Proof.* By [Proposition 2](#), there exists a deterministic commitment outcome  $\bar{\psi}$ . Using the same argument as that used in the proof of [Proposition 1](#), we construct a deterministic commitment outcome  $\tilde{\psi}$  that is IC. By [Proposition 2](#),  $\tilde{\psi}$  is an equilibrium outcome. ■





**Figure 1.** Commitment outcome  $\bar{\psi}$  is IC since  $S$  receives at least his complete-information payoff in every state of the world.

## 4.2. FINITE STATE SPACE

When the state space is finite, i.e.,  $\Theta = \{1, \dots, N\}$ , a deterministic commitment outcome may not exist. For instance, in [Example 1](#), the *unique* commitment outcome is not deterministic. As a result,  $S$  may not be able to achieve the commitment payoff in equilibrium.

However, here we show that when the state space is sufficiently rich (in the sense that  $\mu_0(\theta)$  is sufficiently small for each  $\theta \in \Theta$ ), then  $S$ 's equilibrium payoff approaches his commitment payoff. For a concise argument, we adopt the assumptions of [Alonso and C  mara \(2016\)](#):  $R$  has a binary action and

$$\theta' \neq \theta'' \implies \delta(\theta') \neq \delta(\theta''), \quad (\text{RU})$$

where  $\delta(\theta) = u(2, \theta) - u(1, \theta)$  for all  $\theta \in \Theta$ .

**PROPOSITION 4.** *Suppose that  $\Theta$  is finite,  $|J| = 2$ , (RU) holds, and  $S$ 's payoffs are normalized to  $v(2) = 1$  and  $v(1) = 0$ .<sup>15</sup> Let  $V^*$  be  $S$ 's commitment payoff. For every  $\varepsilon > 0$ , there is  $\gamma > 0$  such that if  $\mu_0(\theta) < \gamma$  for all  $\theta \in \Theta$ , then there exists an equilibrium outcome  $\alpha$  with  $|V^* - V_\alpha| < \varepsilon$ .*

*Proof.* If  $A_2 = \Theta$ , then let  $\alpha(2 \mid \theta) = 1$  for all  $\theta \in \Theta$  so that  $V_\alpha = V^*$ . Thus, we assume for the remainder of the proof that  $A_2$  is a proper subset of  $\Theta$ . Since (RU) holds, we can use Proposition 2 in [Alonso and C  mara \(2016\)](#) to find a cutoff state  $\theta^* \in \Theta$  such that  $\delta(\theta^*) < 0$  and, for every commitment outcome  $\psi$ , we have  $\psi(2 \mid \theta) = 1$  ( $\psi(1 \mid \theta) = 1$ ) for

<sup>15</sup>Normalizing  $S$ 's payoffs is without loss of generality; condition (RU) simplifies the proof, but the result remains true without it.

all  $\theta \in \Theta$  such that  $\delta(\theta) > \delta(\theta^*)$  ( $\delta(\theta) < \delta(\theta^*)$ ). Now consider a deterministic outcome  $\alpha$  with partition  $\{W_1, W_2\}$  such that  $W_2 = \{\theta \in \Theta \mid \psi(2 \mid \theta) = 1\}$  and  $W_1 = \Theta \setminus W_2$ . It is easy to see that  $\alpha$  is IC and obedient, and therefore it is an equilibrium outcome by [Theorem 1](#). If  $\psi(2 \mid \theta^*) = 1$ , the difference in S's ex-ante payoffs is zero; otherwise, we have  $V^* - V_\alpha = \psi(2 \mid \theta^*)\mu_0(\theta^*) < \mu_0(\theta^*) < \gamma := \varepsilon$ . ■

Thus, when R chooses between two actions, S can attain a payoff arbitrarily close to his commitment payoff in equilibrium as long as the prior probability of each state is sufficiently small.

## 5. A MODEL WITH A STOCHASTIC MESSAGE MAPPING

In the main model, IC and obedience are not sufficient for an outcome to be an equilibrium outcome; there exist nondeterministic but IC and obedient outcomes in which S effectively recommends multiple actions, leading to different expected payoffs in the same state. This violates equilibrium condition [\(i\)](#). The reason why [\(i\)](#) is violated is that the mapping  $E : \Theta \rightrightarrows M$ , a correspondence that determines the set of messages available in state  $\theta$ , is deterministic. This assumption is standard in the literature on verifiable disclosure and cheap talk.<sup>16</sup> In some cases, however, it is reasonable to assume that the mapping  $E(\theta)$  is stochastic: for example, there may be different labels for the same state, and S can make statements about the label rather than the state. In this section, we introduce a verifiable disclosure game with a stochastic message mapping (henceforth, the SMM game) and show that IC and obedience are sufficient for an outcome to be an equilibrium outcome of this game.

The SMM game has the same timeline and player objectives as our main model, with the only modification occurring in Stage 2, in which S communicates with R. Specifically, we assume that along with the state of the world  $\theta \in \Theta$ , where the state space  $\Theta = \{1, \dots, N\}$  is finite, S also observes a *label*  $x \in [0, 1]$ , which is payoff-irrelevant to both S and R. The label  $x$  is drawn from the uniform distribution on  $X^\theta$ , where  $\{X^\theta\}_{\theta \in \Theta}$  forms a partition of the unit interval such that  $\lambda(X^\theta) = \mu_0(\theta)$ , where

---

<sup>16</sup>For example, in [Grossman \(1981\)](#), [Milgrom \(1981\)](#), and [Milgrom and Roberts \(1986\)](#),  $E(\theta)$  includes subsets of  $\Theta$  that contain  $\theta$ . In [Dye \(1985\)](#),  $E(\theta)$  is binary; S can reveal  $\theta$  or say nothing. In [Hart, Kremer, and Perry \(2017\)](#) and [Ben-Porath, Dekel, and Lipman \(2019\)](#),  $E(\theta)$  is a partial order on  $\Theta$ . In cheap talk,  $E(\theta)$  is the same for all  $\theta \in \Theta$ .

$\lambda$  is the Lebesgue measure.<sup>17</sup> Having observed  $\theta$  and  $x$ , S sends message  $m \in \widehat{M}$  such that  $x \in m$ , where  $\widehat{M}$  is the collection of nonempty Borel subsets of  $[0, 1]$ . Thus, the set of messages available to S in state  $\theta$  is now determined stochastically (through  $x$ ). The *equilibrium of the SMM game*  $(\widehat{\sigma}, \widehat{\tau}, \widehat{q})$  is defined analogously to that of the main model, except S's strategy also depends on  $x$ .

DEFINITION 4. A triple  $(\widehat{\sigma}, \widehat{\tau}, \widehat{q})$ , where  $\widehat{\sigma} : \Theta \times [0, 1] \rightarrow \Delta_0 \widehat{M}$  is S's strategy,  $\widehat{\tau} : \widehat{M} \rightarrow \Delta J$  is R's strategy, and  $\widehat{q} : \widehat{M} \rightarrow \Delta \Theta$  is R's belief system, is an equilibrium of the SMM game if

- (i) for all  $\theta \in \Theta$  and  $x \in [0, 1]$ ,  $\widehat{\sigma}(\cdot \mid \theta, x)$  is supported on  $\arg \max_{\{m \in \widehat{M} \mid x \in m\}} \sum_{j \in J} v(j) \widehat{\tau}(j \mid m)$ ;
- (ii) for all  $m \in \widehat{M}$ ,  $\widehat{\tau}(\cdot \mid m)$  is supported on  $\arg \max_{j \in J} \int_{\Theta} u(j, \theta) d\widehat{q}(\theta \mid m)$ ;
- (iii)  $\widehat{q}$  is obtained from  $\mu_0$ , given  $\widehat{\sigma}$ , using Bayes' rule;
- (iv) for all  $m \in \widehat{M}$ ,  $\widehat{q}(\cdot \mid m) \in \Delta\{\theta \in \Theta \mid X^\theta \cap m \neq \emptyset\}$ .

Since  $x$  is payoff-irrelevant, an outcome  $\alpha$  of the SMM game is an element of  $\Psi$ . An outcome  $\alpha$  is an *equilibrium outcome* of the SMM game if an equilibrium  $(\widehat{\sigma}, \widehat{\tau}, \widehat{q})$  exists that induces it, i.e.,  $\alpha(j \mid \theta) = \int_{X^\theta} \sum_{m \in \text{supp } \widehat{\sigma}(\cdot \mid \theta, x)} \widehat{\sigma}(m \mid \theta, x) \widehat{\tau}(j \mid m) dx / \mu_0(\theta)$ .

We derive a sharp characterization of equilibrium outcomes in the SMM game.

THEOREM 3. Let  $\Theta$  be finite. Then  $\alpha \in \Psi$  is an equilibrium outcome of the SMM game  $\iff \alpha$  is IC and obedient.

*Proof.*  $(\implies)$  is proved exactly the same way as Theorem 1 (a). An equilibrium outcome must be IC or else S has a profitable deviation to fully revealing  $x$  (which also reveals  $\theta \in \Theta$  since  $x \in X^\theta$ ). An equilibrium outcome must be obedient by Bayes' rule.

$(\impliedby)$  Consider an IC and obedient outcome  $\alpha$ . For every  $\theta \in \Theta$ , let  $J^\theta := \text{supp } \alpha(\cdot \mid \theta)$  be the set of actions that R takes with a positive probability when the realized state is  $\theta$ . Next, partition  $X^\theta$  into a set of intervals  $\{X_j^\theta\}_{j \in J^\theta}$  such that  $\lambda(X_j^\theta) / \lambda(X^\theta) = \alpha(j \mid \theta)$ . Also, for each action  $j \in J$ , let  $W_j := \bigcup_{\theta \in \Theta} X_j^\theta$ ; by construction,  $\{W_j\}_{j \in J}$  is a partition of  $[0, 1]$ .

Now let S's strategy be  $\widehat{\sigma}(m \mid \theta, x) = \mathbb{1}(m = W_j \text{ and } x \in W_j)$ . Then R's posterior after an on-path message  $W_j$  is  $\widehat{q}(\theta \mid W_j) = \lambda(X_j^\theta) / \lambda(W_j)$ . Furthermore, since  $\alpha$  is

<sup>17</sup>For example, let  $t_0 := 0$ ,  $t_\theta := \sum_{\theta'=1}^{\theta} \mu_0(\theta')$  for all  $\theta \in \Theta$ ; also, let  $X^\theta = [t_{\theta-1}, t_\theta]$  for all  $\theta \in \{1, \dots, N-1\}$

and  $X^N = [t_{N-1}, 1]$ . Then  $\{X^\theta\}_{\theta \in \Theta}$  is a partition of  $[0, 1]$  and  $\lambda(X^\theta) = \sum_{\theta'=1}^{\theta} \mu_0(\theta') - \sum_{\theta'=1}^{\theta-1} \mu_0(\theta') = \mu_0(\theta)$ .

obedient, for every action  $j \in J$  such that  $\lambda(W_j) > 0$ , we have

$$\begin{aligned} \sum_{\theta \in \Theta} (u(j, \theta) - u(j', \theta)) \alpha(j | \theta) \mu_0(\theta) &\geq 0 \iff \\ \sum_{\theta \in \Theta} (u(j, \theta) - u(j', \theta)) \frac{\lambda(X_j^\theta)}{\lambda(W_j)} &\geq 0 \quad \text{for all } j' \in J, \end{aligned}$$

meaning that R prefers to take action  $j$  after message  $W_j$ , so we let  $\hat{\tau}(j | W_j) = 1$ . Off the path, let R be “skeptical” and assume that any unexpected message comes from the state in which S benefits from such deviation the most. Formally, for all  $m \notin \{W_j\}_{j \in J}$ , let  $\hat{q}(\cdot | m) \in \Delta A_{\underline{j}}$ , where  $\underline{j} \in J$  is the lowest action such that  $m \cap X^\theta \neq \emptyset$  and  $\theta \in A_i$ . Then playing action  $\underline{j}$  is a best response to message  $m$ , so we let  $\hat{\tau}(\underline{j} | m) = 1$ . Since  $\alpha$  is IC, S does not have profitable deviations by the same argument as in the proof of [Theorem 1](#). Deviations to on-path messages are not available because  $\{W_j\}_{j \in J}$  is a partition, and deviations to off-path messages are not profitable since the payoff from any deviation in state  $\theta$  is at most  $\underline{v}(\theta)$ , which is below  $v_\alpha(\theta)$  by the [\(IC \$\_\theta\$ \)](#) constraint. Hence,  $(\hat{\sigma}, \hat{\tau}, \hat{q})$  is an equilibrium of the SMM game.  $\blacksquare$

In contrast to [Theorem 1](#), IC and obedience are necessary and sufficient for an outcome to be an equilibrium outcome of the SMM game. Two properties of the SMM game ensure that every IC and obedient outcome is an equilibrium outcome. First, S’s message space depends on  $x$ , which means S may receive different equilibrium payoffs in some state  $\theta$  (but for different realizations of  $x$ ). Second, the message space is “rich,” meaning that for every vector  $p = (p_1, \dots, p_N) \in [0, 1]^N$ , there exists a message  $m$  that is available in state  $\theta \in \Theta$  with probability  $p_\theta$ . This richness allows us to “purify” any nondeterministic outcome: the equilibria that we construct to implement an IC and obedient outcome is in pure strategies of both S and R.

Using the sharp equilibrium characterization of the SMM game, we derive the following results.

**COROLLARY 1.** *Let  $\Theta$  be finite. Then a commitment outcome is an equilibrium outcome of the SMM game if and only if it is IC.*

**COROLLARY 2.** *If  $\Theta$  is finite and  $|J| = 2$ , then every commitment outcome is an equilibrium outcome of the SMM game.*

[Corollary 1](#) is a direct consequence of [Theorem 3](#). [Corollary 2](#) follows from [Theorem 3](#) and the fact that every commitment outcome is IC when R has two actions (see [Alonso and Câmara, 2016](#) and our discussion after [Proposition 1](#)).

The set of equilibrium outcomes in the SMM game coincides with the set of IO outcomes found in KS if S’s value function is quasiconvex in R’s belief (KS Proposition 3), or when R chooses between two actions (KS Proposition 4). Generally, the set of IO outcomes is a subset of the set of equilibrium outcomes in KS, because IO imposes a stronger restriction on off-path beliefs than our equilibrium concept.

## 6. CONCLUSION

This paper examined a persuasion game with verifiable information in which a sender with transparent motives chooses which verifiable messages to send to a receiver in order to convince her to take a particular action from a finite set. We showed that every equilibrium outcome must be incentive-compatible for the sender and obedient for the receiver. If an outcome is deterministic, then these conditions are both necessary and sufficient for it to be an equilibrium outcome. We also identified sufficient conditions under which the ex-ante commitment assumption in Bayesian persuasion can be replaced by communication with verifiable information. We showed that if the state space is rich, then a deterministic commitment outcome always exists; this commitment outcome is an equilibrium outcome if and only if the sender receives at least his complete information payoff in every state. If the receiver chooses between two actions, this condition is automatically satisfied. We hope these results prove useful in applied settings.

## REFERENCES

- ALI, S. NAGEEB, ANDREAS KLEINER, and KUN ZHANG (2024), “From Design to Disclosure”, working paper. (p. 3.)
- ALONSO, RICARDO and ODILON CÂMARA (2016), “Persuading Voters”, *American Economic Review*, 106, 11 (Nov. 2016), pp. 3590-3605. (pp. 2, 13, 15, 18.)
- BARDHI, ARJADA and YINGNI GUO (2018), “Modes of Persuasion toward Unanimous Consent”, *Theoretical Economics*, 13, 3, pp. 1111-1149. (p. 2.)
- BEN-PORATH, ELCHANAN, EDDIE DEKEL, and BARTON L. LIPMAN (2019), “Mechanisms With Evidence: Commitment and Robustness”, *Econometrica*, 87, 2, pp. 529-566. (pp. 4, 16.)
- BOLES LAVSKY, RAPHAEL and CHRISTOPHER COTTON (2015), “Grading Standards and Education Quality”, *American Economic Journal: Microeconomics*, 7, 2 (May 2015), pp. 248-279. (p. 2.)
- BULL, JESSE and JOEL WATSON (2007), “Hard Evidence and Mechanism Design”, *Games and Economic Behavior*, 58, 1, pp. 75-93. (p. 5.)

- CHAKRABORTY, ARCHISHMAN and RICK HARBAUGH (2010), “Persuasion by Cheap Talk”, *American Economic Review*, 100, 5, pp. 2361-2382. (p. 4.)
- DRANOVE, DAVID and GINGER ZHE JIN (2010), “Quality Disclosure and Certification: Theory and Practice”, *Journal of Economic Literature*, 48, 4, pp. 935-963. (p. 3.)
- DVORETZKY, A., A. WALD, and J. WOLFOWITZ (1951), “Elimination of Randomization in Certain Statistical Decision Procedures and Zero-Sum Two-Person Games”, *The Annals of Mathematical Statistics*, 22 (1 Mar. 1951), pp. 1-21. (p. 14.)
- DYE, RONALD A (1985), “Disclosure of Nonproprietary Information”, *Journal of Accounting Research*, pp. 123-145. (p. 16.)
- GEHLBACH, SCOTT and KONSTANTIN SONIN (2014), “Government Control of the Media”, *Journal of Public Economics*, 118 (Oct. 2014), pp. 163-171. (p. 2.)
- GENTZKOW, MATTHEW and EMIR KAMENICA (2016), “A Rothschild-Stiglitz Approach to Bayesian Persuasion”, *American Economic Review*, 106, 5 (May 2016), pp. 597-601. (p. 14.)
- GIECZEWSKI, GERMÁN and MARIA TITOVA (2024), “Coalition-Proof Disclosure”, working paper. (p. 3.)
- GLAZER, JACOB and ARIEL RUBINSTEIN (2004), “On Optimal Rules of Persuasion”, *Econometrica*, 72 (6 Nov. 2004), pp. 1715-1736. (p. 4.)
- (2006), “A Study in the Pragmatics of Persuasion: a Game Theoretical Approach”, *Theoretical Economics*, 1, pp. 395-410. (p. 4.)
- GROSSMAN, SANFORD J. (1981), “The Informational Role of Warranties and Private Disclosure about Product Quality”, *The Journal of Law and Economics*, 24, 3 (Dec. 1981), pp. 461-483. (pp. 1, 3, 16.)
- GROSSMAN, SANFORD J. and OLIVER D. HART (1980), “Disclosure Laws and Takeover Bids”, *The Journal of Finance*, 35, 2, pp. 323-334. (p. 3.)
- HART, SERGIU, ILAN KREMER, and MOTTY PERRY (2017), “Evidence Games: Truth and Commitment”, *American Economic Review*, 107 (3 Mar. 2017), pp. 690-713. (pp. 4, 16.)
- KAMENICA, EMIR and MATTHEW GENTZKOW (2011), “Bayesian Persuasion”, *American Economic Review*, 101, 6 (Oct. 2011), pp. 2590-2615. (pp. 2, 9, 11, 12.)
- KAMENICA, EMIR and XIAO LIN (2024), “Commitment and Randomization in Communication”, *arXiv preprint arXiv:2410.17503*. (p. 5.)
- KOESSLER, FRÉDÉRIC and RÉGIS RENAULT (2012), “When Does a Firm Disclose Product Information?”, *The RAND Journal of Economics*, 43, 4, pp. 630-649. (p. 10.)
- KOESSLER, FRÉDÉRIC and VASILIKI SKRETA (2023), “Informed Information Design”, *Journal of Political Economy* (Mar. 28, 2023). (pp. 3, 13.)
- KOLOTILIN, ANTON (2015), “Experimental Design to Persuade”, *Games and Economic Behavior*, 90 (Mar. 2015), pp. 215-226. (p. 2.)
- LIPNOWSKI, ELLIOT (2020), “Equivalence of Cheap Talk and Bayesian Persuasion in a Finite Continuous Model”, working paper. (p. 4.)

- LIPNOWSKI, ELLIOT and DORON RAVID (2020), “Cheap Talk With Transparent Motives”, *Econometrica*, 88, 4, pp. 1631-1660. (p. 4.)
- MILGROM, PAUL R. (1981), “Good News and Bad News: Representation Theorems and Applications”, *The Bell Journal of Economics*, 12, 2, p. 380. (pp. 1, 3, 16.)
- (2008), “What the Seller Won’t Tell You: Persuasion and Disclosure in Markets”, *Journal of Economic Perspectives*, 22, 2 (Mar. 2008), pp. 115-131. (p. 3.)
- MILGROM, PAUL R. and JOHN ROBERTS (1986), “Relying on the Information of Interested Parties”, *The RAND Journal of Economics*, 17, 1, p. 18. (pp. 3, 5, 16.)
- OSTROVSKY, MICHAEL and MICHAEL SCHWARZ (2010), “Information Disclosure and Unraveling in Matching Markets”, *American Economic Journal: Microeconomics*, 2, 2 (May 2010), pp. 34-63. (p. 2.)
- PEREZ-RICHET, EDUARDO (2014), “Interim Bayesian Persuasion: First Steps”, *American Economic Review*, 104, 5, pp. 469-474. (p. 3.)
- SHER, ITAI (2011), “Credibility and Determinism in a Game of Persuasion”, *Games and Economic Behavior*, 71 (2 Mar. 2011), pp. 409-419. (p. 4.)
- TERSTIEGE, STEFAN and CÉDRIC WASSER (2023), “Experiments Versus Distributions of Posteriors”, *Mathematical Social Sciences*, 125, pp. 58-60. (p. 12.)
- ZAPECHELNYUK, ANDRIY (2023), “On the Equivalence of Information Design by Uninformed and Informed Principals”, *Economic Theory*, 76, 4 (Nov. 2023), pp. 1051-1067. (p. 3.)
- ZHANG, KUN (2022), “Withholding Verifiable Information”, working paper. (p. 3.)