Persuasion with Verifiable Information*

Maria Titova[†] Kun Zhang[‡]

March 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 to commitment outcomes from information design. We provide conditions for a commitment outcome to be an equilibrium outcome and identify environments in which the sender does not benefit from having commitment power. Our findings offer insights into the interchangeability of verifiability and commitment in applied settings.

*This paper combines material from an earlier paper with the same title by the first author and from Zhang (2022). We thank the editor, the associate editor, and two anonymous referees for their excellent editorial comments. We are indebted to Renee Bowen, Andreas Kleiner, and Joel Sobel for their guidance and support. We are also grateful to Nageeb Ali, Arjada Bardhi, Simone Galperti, Germán Gieczewski, Navin Kartik, Remy Levin, Aleksandr Levkun, Elliot Lipnowski, Denis Shishkin, Joel Watson, Mark Whitmeyer, and various seminar and conference audiences for their helpful comments and suggestions. All errors are our own.

[†]Department of Economics, Vanderbilt University. Email: motitova@gmail.com.

[‡]School of Economics, University of Queensland. Email: kun@kunzhang.org.

1. Introduction

Suppose that a sender is privately informed about the state of the world and would like to convince a receiver to take a particular action. The sender does not have commitment power, but his messages are verifiable statements about the state of the world. What are the possible outcomes of this interaction?

Persuasion with verifiable information plays an essential role in many economic settings, including courtrooms, electoral campaigns, product advertising, financial disclosure, and job market signaling. For example, 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 in order to win over voters. In advertising, a firm convinces consumers to purchase its product by highlighting only specific product characteristics. In finance, a CEO divulges only certain financial statements and indicators to board members to obtain higher compensation. In a labor market, a job candidate lists specific certifications in order to make her application more attractive to an employer.

We consider the following model of persuasion with verifiable information. First, the sender (S, 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 (R, she/her). Verifiability means that any feasible message must contain the truth (the true state of the world), but not necessarily the whole truth (it may contain other states). Upon observing the message, R takes an action from a finite set. S's preferences are state-independent and are strictly increasing in R's action, while R's preferences depend on both her action and the state. We focus on the *outcomes* of this interaction, which we define as mappings from the state space to a distribution over R's actions.

Our first set of results describes equilibrium outcomes, which are induced by the (perfect Bayesian) equilibria of the game. Our first main result (Theorem 1) states that every equilibrium outcome must be incentive-compatible (for S, IC for short) and obedient (for R). The IC constraint requires that S receives at least his complete information payoff in each state; otherwise, he would have a profitable deviation toward fully revealing the state. R's obedience constraint states that if she takes some action with a positive probability in some states, then that action 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 R takes some action with probability one in every state. Although not all equilibrium outcomes are deterministic, we show in Lemma 1 that all equilibria wherein R does not mix (e.g. if R uses a predetermined tie-breaking rule) induce deterministic outcomes.

In this 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 S can achieve the same payoff in equilibrium as he does in information design (e.g., Kamenica and Gentzkow, 2011). In information design, S commits to a disclosure strategy ahead of 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 deterministic and IC commitment outcome.

In general, commitment outcomes that are IC and non-deterministic cannot be induced by an equilibrium because R breaks ties in favor of S in a commitment outcome. This implies that a commitment outcome must be deterministic in order to be an equilibrium outcome. We argue; however, that this is not a substantive concern. When the state space is rich, a deterministic commitment outcome exists (Proposition 2). When the state space is finite, we show that modifying our game to one in which the set of available verifiable messages is determined stochastically allows us to implement any IC (but not necessarily deterministic) commitment outcome (see Section 5).

The second difference between equilibrium and commitment outcomes is that S does not face any incentive-compatibility concerns when he has commitment power. Therefore, our second main result essentially states that a (deterministic) commitment outcome is an equilibrium outcome if and only if S obtains at least his complete information payoff in every state. Additionally, we show that when R chooses between two actions, an IC commitment outcome always exists (Proposition 1). In those settings, which are common in applications, a (deterministic) commitment outcome is always an equilibrium outcome.[1]

Related Literature

The literature on persuasion with verifiable information is pioneered by Milgrom (1981) and Grossman (1981); this paper has the same mapping from state to available messages as Milgrom and Roberts (1986), except in Section 5.[2] In all these papers, unraveling is the unique equilibrium outcome, because S's preferences are strictly monotone in R's action (e.g. he is maximizing quantity sold) and R's action space is rich (e.g. she is choosing a perfectly divisible quantity to buy). The unraveling argument goes as

[1]For example, pharmaceutical companies persuade the Food and Drug Administration to approve their drug (Kolotilin, 2015); schools persuade employers to hire their graduates (Ostrovsky and Schwarz, 2010; Boleslavsky and Cotton, 2015); politicians persuade voters (Alonso and Câmara, 2016; Bardhi and Guo, 2018); and governments persuade citizens (Gehlbach and Sonin, 2014).

[2]The verifiable information literature is reviewed by Milgrom (2008) and Dranove and Jin (2010).

follows: the sender who is privately informed about the quality of his product always wants to separate himself from all lower-quality senders, as that convinces R to purchase a strictly higher quantity of the product. We argue that if R's action space is finite, then unraveling is not the only equilibrium outcome.Our argument is easiest to see when R's action space is binary, such as when she is choosing between buying and not buying. Then, the high-quality senders may not mind pooling with some lower-quality senders, as they are already getting the highest possible payoff.

A few recent papers also characterize the equilibrium set (or the set of equilibrium payoffs of the sender) and assess the value of commitment in various models of communication with verifiable information. In our companion paper, Zhang and Titova (2025), we impose considerably more structure on the receiver's preferences, which yields sharper results. Gieczewski and Titova (2024) consider a game with a generalized message mapping and focus on coalition-proof equilibria. In Ali, Kleiner, and Zhang (2024), the sender could have state-dependent preferences, but they focus on settings where IC is not a concern.

While we study the conditions under which the sender does not benefit from commitment power, there is growing literature studying how much the receiver gains from commitment power by comparing equilibrium outcomes to the optimal mechanism outcomes for 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 to reach the optimal mechanism outcome. Hart, Kremer, and Perry (2017) establish the conditions for the equivalence of the equilibrium and optimal mechanism outcomes. Ben-Porath, Dekel, and Lipman (2019) confirm that the receiver does not need commitment even if the sender has state-dependent preferences.

Some other papers also compare equilibrium outcomes in one-shot communication games with commitment outcomes, but for different communication protocols. Most closely related ones are Chakraborty and Harbaugh (2010), Lipnowski and Ravid (2020), and Lipnowski (2020)'s analyses of cheap talk games, and a few papers in the informed information design literature. In cheap talk games, the sender's messages need not be verifiable: in every state, the sender has access to the same (sufficiently rich) set of messages. The verifiability requirement faced by our sender can significantly impact equilibrium outcomes.[3]

[3] It may both help and hurt the sender, depending on the preferences of the players. In fact, the entire equilibrium set of the verifiable information game may be ex-ante better for the sender than the entire set of cheap talk equilibria, and vice versa.

Our paper is also related to the informed information design (IID henceforth) literature, pioneered by Perez-Richet (2014). There, the sender chooses an *arbitrary* Blackwell experiment like in the information design literature, except he observes the state of the world before making the choice. As Koessler and Skreta (2023) (KS henceforth) point out, in certain informed information design settings, "PBE has very little predictive power ... because off-path beliefs can be chosen in a way that completely cancels out the information revealed by off-path experiments." To address this issue, KS focus on interim optimal (IO) outcomes that place a restriction on the receiver's off-path beliefs: any such belief must assign positive probability only to states in which the sender strictly benefits from the deviation. As pointed out by Perez-Richet (2014), our persuasion game with verifiable information can be interpreted as a *constrained* IID problem: S observes the state of the world, then chooses a verifiable message (rather than an arbitrary experiment). Our approach is thus complementary to KS: our constraint on the set of feasible experiments already restricts the set of feasible off-path beliefs (when R hears a message, she knows it cannot have come from a state where that message is not available), and we choose not to restrict it any further.[4] Our characterization of equilibrium outcomes happens to be directly comparable to KS's characterization IO outcomes. We discuss the differences and similarities throughout the paper.

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:[5]

1. S observes the state of the world $\theta \in \Theta$.
   The state space Θ is either finite ($\Theta = \{1, \ldots, N\}$, $N \geq 2$) or rich (Θ 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 supp $\mu_0 = \Theta$. If the state space is rich, we further assume that the prior is atomless.
2. S sends message $m \in M$ to the receiver, where M is the collection of nonempty

[4] Zapechelnyuk (2023), like us, does not restrict off-path beliefs; instead, he assumes that the sender's experiments are never fully revealing, and finds that every Pareto undominated commitment outcome can be implemented by an informed designer.

[5] For a topological space Y, let $\Delta(Y)$ denote the set of Borel probability measures on Y. For $\gamma \in \Delta Y$, let supp γ denote the support of γ. We say that $\gamma \in \Delta Y$ is degenerate if supp γ is a singleton, and non-degenerate otherwise.

Borel subsets of Θ. Each message is a subset of the state space and we can interpret it as a statement about the state of the world. The sender's messages are verifiable:[6]

Definition 1. *Message* $m \in M$ *is* verifiable in state $\theta \in \Theta$ *if* $\theta \in m$.

3. R observes the message (but not the state) and takes action from a finite set J with $|J| \geq 2$.
4. Game ends, payoffs are realized.

   S's payoff $v : J \to \mathbb{R}$ depends only on R's action. We assume that S prefers higher actions, i.e. $v(j)$ is strictly increasing in $j \in J$.

   R's preferences are described by a bounded utility function $u : J \times \Theta \to \mathbb{R}$. We define R's *complete information action-j set* as $A_j := \{\theta \in \Theta \mid u(j, \theta) \geq u(j', \theta) \text{ for all } j' \in J \setminus \{j\}\}$ 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. Firstly, S's strategy is a function $\sigma : \Theta \to \Delta_0 M$, where $\Delta_0 M$ is the set of probability measures on M with a finite support.[7] Secondly, R's strategy is a function $\tau : M \to \Delta J$. Finally, R's belief system is $q : M \to \Delta\Theta$ describes R's beliefs about the state after any observed message.

Definition 2. *A triple* (σ, τ, q) *is an* equilibrium *if*

(i) For all $\theta \in \Theta$, $\sigma(\cdot \mid \theta)$ *is supported on* $\underset{\{m \in M \mid \theta \in m\}}{\arg\max} \sum_{j \in J} v(j)\, \tau(j \mid m)$*;*

(ii) For all $m \in M$, $\tau(\cdot \mid m)$ *is supported on* $\underset{j \in J}{\arg\max} \int_\Theta u(j, \theta)\, \mathrm{d}q(\theta \mid m)$*;*

(iii) q *is obtained from* μ_0*, given* σ*, using Bayes rule;*[8]

(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

[6]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.

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

[8]That is, q is a regular conditional probability system.

posterior belief; *(iii)* R uses Bayes' rule to update her beliefs whenever possible; and *(iv)* R takes into account that messages are verifiable on and off the path.

To analyze the model, we use the following approach. Let Ψ be the set of all Borel measurable functions from Θ to ΔJ. 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 (σ, τ) of S and R, we let $M_j(\sigma, \tau) := \{m \in M \mid m \in \operatorname{supp} \sigma(\cdot \mid \theta) \text{ for some } \theta \in \Theta \text{ and } \tau(j \mid m) > 0\}$ be the set of messages sent with a positive probability after which R takes action $j \in J$ with a positive probability. We say that $\alpha \in \Psi$ is an *equilibrium outcome* if there exists an equilibrium (σ, τ, 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 $\theta \in \Theta$.

Of particular interest are outcomes in which R takes a particular action in every state. 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 α, we refer to the collection of sets $\{W_j\}_{j \in J}$, where $W_j := \{\theta \in \Theta \mid \alpha(j \mid \theta) = 1\}$, as the *outcome partition* (into subsets W_j of the state space in which R takes action $j \in J$ with probability one) of α.

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

3. Equilibrium Analysis

We begin by establishing the lower bound on S's payoff in an equilibrium outcome α. One thing that S can do in state θ is fully reveal the state by sending message $\{\theta\}$ with probability one. When R hears message $\{\theta\}$, she learns that the state is θ and takes an action that is a best response under complete information. Therefore, S's equilibrium payoff in state θ is bonded 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 incentive-compatibility constraint.[9]

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

Definition 3. *An outcome α is* incentive-compatible *(IC) if it satisfies* (IC$_\theta$) *for each state $\theta \in \Theta$.*

---

[9] In fact, $\underline{v}(\theta)$ is the lower bound on S's equilibrium payoff in state θ, meaning that there exists an equilibrium in which S's equilibrium payoff is exactly $\underline{v}(\theta)$ for each $\theta \in \Theta$. In that 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).

In information design, $\alpha(j \mid \theta)$ is interpreted as the probability that S recommends action j in state θ; R's best response is to follow that recommendation if

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

Definition 4. *An outcome α is* obedient *if it satisfies* (obedience$_j$) *for each action $j \in J$.*

Of course, our setting does not allow for action recommendations as S's message space is not J. However, we will see shortly that all equilibrium outcomes must be obedient.

If α is a deterministic outcome with partition $\{W_j\}_{j\in J}$, then (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 θ must be no lower than R's worst best response under complete information. The obedience constraint for action j simplifies to $\int_{W_j} (u(j,\theta) - u(j',\theta))\,\mathrm{d}\mu_0(\theta) \geq 0$ for all $j' \in J \setminus \{j\}$.

Our first result confirms that every equilibrium outcome is incentive-compatible and obedient. If an outcome is deterministic, then these two properties are necessary and sufficient for it to be an equilibrium outcome.

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.

Proof. [Part *(a)*] Consider an equilibrium (σ, τ, q) with outcome $\alpha \in \Psi$. Observe that α must be incentive-compatible, or else there exists a state θ in which S has a profitable deviation to fully revealing the state. Next, we show that α is also obedient. Consider any action $j \in J$. By the equilibrium condition *(ii)*, we have

$$\text{for all } m \in M_j(\sigma,\tau) \text{ and } j' \in J \setminus \{j\}, \quad \int_{\Theta} (u(j,\theta) - u(j',\theta))\ \mathrm{d}q(\theta \mid m) \geq 0$$
$$\implies \int_{\Theta} (u(j,\theta) - u(j',\theta))\tau(j \mid m)\,\mathrm{d}q(\theta \mid m) \geq 0,$$

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

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

where the last inequality is (obedience$_j$). Since the choice of j is arbitrary, α is obedient.

[Part *(b)*] Consider a deterministic outcome α that is IC and obedient and denote its outcome partition by $\{W_j\}_{j\in J}$. We construct an equilibrium (σ, τ, q) that induces α. Let S's strategy be $\sigma(m \mid \theta) = \mathbb{1}(m = W_j \text{ and } \theta \in W_j)$ be to reveal which element of the outcome partition 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$_j$), playing action j is a best response, so we let $\tau(j \mid W_j) = 1$ for all $j \in J$. For off-path messages, 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 $q(\cdot \mid m) \in \Delta(m \cap A_{\underline{j}})$, where $\underline{j} \in J$ is the lowest action $i \in J$ such that the set $m \cap A_i$ is non-empty. 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 $i \neq j$, so W_i is not a verifiable message in state θ. 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, (σ, τ, q) is an equilibrium that induces α. ∎

Part *(a)* of Theorem 1 confirms that every equilibrium outcome must be obedient. To see why, consider all on-path messages after which R plays j with a positive probability. If the same action is a best response after all these messages, then she should choose the same action without knowing which of these messages was sent. Since R chooses the same optimal action after all these messages, we can "bundle" them into a single "recommendation" to take action j. Part *(b)* of Theorem 1 characterizes the set of deterministic equilibrium outcomes, and its proof suggests a simple way of implementing them in a pure-strategy equilibrium with at most J on-path messages that essentially serve as action recommendations. Specifically, if $\{W_j\}_{j\in J}$ is an outcome partition, then W_j is both the set of states in which R plays action j and the on-path message that "recommends" that R plays j in the equilibrium that induces this outcome.

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

Example 1. 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 $\mu_0(\text{innocent}) = 0.7$. S's preferences are $v(\text{acquit}) = 0$ and $v(\text{convict}) = 1$, while R wants to "match the state": $u(\text{acquit}, \text{innocent}) = u(\text{convict}, \text{guilty}) = 1$, and $u(\text{acquit}, \text{guilty}) = u(\text{convict}, \text{innocent}) = 0$. Consider an outcome α^* such that $\alpha^*(\text{convict} \mid \text{guilty}) = 1$ and $\alpha^*(\text{convict} \mid \text{innocent}) = 3/7$. It is straightforward to verify that α^* is both IC and obedient. However, α^* is not an equilibrium outcome: when $\theta = \text{innocent}$, 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 $\theta = \text{innocent}$.

Example 1 illustrates that (IC and obedient) outcomes in which S receives different payoffs in the same state cannot be equilibrium outcomes. Once the state is realized, S's message space is fixed and known. Thus, if S mixes between multiple messages in the same state, he must be receiving the same payoff following these messages. In Section 5, we show that once S's set of available messages is rich and determined stochastically, IC and obedience become necessary and sufficient for equilibrium implementation, because requirement that S's payoff is constant in state θ is lifted.

Of course, if S does receive the same payoff in state θ, then an IC, obedient and non-deterministic outcome could be an equilibrium outcome. However, in any such equilibrium, R must be playing a mixed strategy in every such equilibrium:

Lemma 1. *Suppose that α is a non-deterministic outcome induced by an equilibrium (σ, τ, q). Then in each $\theta \in \Theta$ such that $\alpha(\cdot \mid \theta)$ is non-degenerate, R is playing a mixed strategy (that is, $\tau(\cdot \mid m)$ is non-degenerate) for some $m \in \operatorname{supp} \sigma(\cdot \mid \theta)$.*

Proof. Consider a state $\theta \in \Theta$ such that $\alpha(\cdot \mid \theta)$ is non-degenerate. By contradiction, suppose that $\tau(\cdot \mid m)$ is degenerate for all $m \in \operatorname{supp} \sigma(\cdot \mid \theta)$. By the equilibrium condition *(i)*, for any pair of messages $m, m' \in \operatorname{supp} \sigma(\cdot \mid \theta)$, we have $\sum\limits_{j \in J} v(j)\tau(j \mid m) = \sum\limits_{j \in J} v(j)\tau(j \mid m')$, implying that there exists an action $j^* \in J$ such that $\tau(j^* \mid m) = \tau(j^* \mid m') = 1$. In words, if R is not mixing, then every message sent by S in state θ convinces R to take the same action. Therefore, $\alpha(j^* \mid \theta) = \sum\limits_{m \in M_j(\sigma,\tau)} \tau(j^* \mid m)\sigma(m \mid \theta) = 1$, a contradiction. ∎

The converse of Lemma 1 also tells us that if R is not mixing in an equilibrium (e.g. if she uses a predetermined tie-breaking rule), then the outcome is deterministic. Lemma 1 will play an important role later in establishing our next main result.

Theorem 1 characterizes all pure-strategy equilibria of the game, since those 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 where S has state-independent preferences, sends verifiable messages and sets a price, and R chooses between two actions. Theorem 1 highlights that their result (1) extends to more than two actions of R and (2) is not driven by S's extra choice variable (price).[10]

Theorem 1 also highlights the differences between our approach and that of Koessler and Skreta (2023). Their characterization (Proposition 2) states that an outcome is IO if and only if it is obedient and IOC, where IOC essentially requires that S does not wish to deviate to any experiment (which is just an outcome) that pools any fraction of states. Naturally, the first difference (they have IOC and we have IC) is due to our different approaches: they restrict the set of off-path beliefs whereas we restrict the set of feasible experiments. Interestingly, KS also show that IC implies IOC if S's value function is quasiconvex in R's belief (Proposition 3) or when R chooses between two actions (Lemma B.2). The second difference is that IOC and obedience are necessary and sufficient for an outcome to be an IO outcome, while for us IC and obedience are not sufficient. Since in our model the sender only chooses messages, an additional restriction applies: S may mix between different messages only when they yield the same expected payoffs — a constraint absent in (unconstrained) informed information design. For this reason, some IO outcomes (e.g., one from Example 1) are not equilibrium outcomes in our game. This difference washes away once we replace our assumption that the set of available messages is fixed and known (which we borrowed from Milgrom and Roberts, 1986) with that the message space is determined stochastically (see Section 5).

4. Value of Commitment

In this section, we ask: when is a commitment outcome (a solution to the information design problem) also an equilibrium outcome? In the information design problem, stage 1 of the game (wherein the sender learns the state) is removed, and stage 2 of the game (wherein the sender chooses a verifiable message) is replaced by S committing to an

[10]Except, as the authors point out, the price choice ensures that R's best response is a pure strategy, which implies all equilibria are deterministic.

experiment that sends signals depending on state realizations.[11] Importantly, S who has commitment power no longer faces an incentive-compatibility constraint: he need not maximize his utility state-by-state, and his signals need not be verifiable.

We follow Kamenica and Gentzkow (2011) and focus on straightforward signals that R interprets as action recommendations. Therefore, an (optimal) *commitment outcome* $\overline{\psi} \in \Psi$ solves

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

Simply put, a commitment outcome is an outcome (i.e. an element of Ψ) that maximizes S's ex-ante utility and is obedient, since the constraints in problem (CO) are the obedience constraints. We call the value of problem (CO) the *commitment payoff*. Our second result establishes that a commitment outcome has to be deterministic and incentive-compatible in order to be an equilibrium outcome.

Theorem 2. *Consider a commitment outcome $\overline{\psi} \in \Psi$.*

(a) If $\overline{\psi}$ is IC and deterministic, then it is an equilibrium outcome.
(b) If $\overline{\psi}$ is an equilibrium outcome, then it is IC and deterministic μ_0-almost everywhere.

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

[Part *(b)*] Suppose a commitment outcome $\overline{\psi}$ is an equilibrium outcome, meaning that there exists an equilibrium (σ, τ, q) that induces it. By Theorem 1, $\overline{\psi}$ is incentive-compatible. In what follows, we show that $\overline{\psi}$ is deterministic μ_0-almost everywhere. Let $T := \{\theta \in \Theta \mid \overline{\psi}(\cdot \mid \theta) \text{ is non-degenerate}\}$ be 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 exists a message $m \in \operatorname{supp} \sigma(\cdot \mid \theta)$ such that $\tau(\cdot \mid m)$ is non-degenerate. Let $\widetilde{M} := \{m \in M \mid \tau(\cdot \mid m) \text{ is non-degenerate}\}$ be the set of messages after which R plays a mixed strategy. Also, let $\widetilde{\tau}(j^* \mid m) := \mathbb{1}(j^* = \max_{j \in \operatorname{supp} \tau(\cdot \mid m)} j)$ for all $m \in M$ be the strategy of R that breaks all ties in τ in favor of S. Define the outcome from the strategy profile $(\sigma, \widetilde{\tau})$ by $\widetilde{\psi}$.

[11] Specifically, an experiment (S, χ) consists of a compact metrizable space S of signals and a Borel measurable function $\chi : \Theta \to \Delta S$. When the realized state is $\theta \in \Theta$, R sees a signal $s \in S$ drawn from $\chi(\cdot \mid \theta)$.

We now arrive at a contradiction by showing that $\widetilde{\psi}$ is an obedient outcome such that $V_{\widetilde{\psi}} > V_{\overline{\psi}}$ (which means 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 $\sigma(m \mid \theta) > 0$ for some $m \in \widetilde{M}$), while $v_{\widetilde{\psi}}(\theta) = v_{\overline{\psi}}(\theta)$ for all $\theta \notin T$. Therefore, $V_{\widetilde{\psi}} - V_{\overline{\psi}} = \int_T (v_{\widetilde{\psi}}(\theta) - v_{\overline{\psi}}(\theta))\, \mathrm{d}\mu_0(\theta) > 0$ since $\mu_0(T) > 0$. To show that $\widetilde{\psi}$ is obedient, we can write down equilibrium condition (ii) for the equilibrium (σ, τ, q) and follow the steps of the proof of Theorem 1*(a)*, replacing τ with $\widetilde{\tau}$ and noting that $M_j(\sigma, \widetilde{\tau}) \subseteq M_j(\sigma, \tau)$. ∎

The non-trivial part of Theorem 2 involves proving that if $\overline{\psi}$ is an equilibrium and a commitment outcome, then it is deterministic almost everywhere, which is equivalent to showing that an equilibrium outcome $\overline{\psi}$ that is not deterministic a.e. cannot be a commitment outcome. Indeed, by Lemma 1, in the equilibrium that induces $\overline{\psi}$, R must be playing a mixed strategy after some on-path messages coming 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 means 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 point of view, there are "bad" states $\theta \in A_1$, where R prefers to take the low action 1, and "good" states $\theta \notin A_1$, where R prefers to take 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 incentive-compatible, it suffices to show that $\theta \notin A_1$ implies that $v_\psi(\theta) = v(2)\psi(2 \mid \theta) + v(1)\psi(1 \mid \theta) \geq v(2)$, which is equivalent to $\psi(2 \mid \theta) = 1$; there are no IC conditions for $\theta \in A_1$ since $v(1)$ 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 where action 2 is the unique best response under complete information. The following result establishes existence of an incentive-compatible commitment outcome when R chooses between two actions.

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

Proof. Since Θ is a compact subset of $\mathbb{R}^n$, the existence of a commitment outcome follows from 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 $\widetilde{\psi}(\cdot \mid \theta) = \overline{\psi}(\cdot \mid \theta)$ for all $\theta \notin A_2$ and $\widetilde{\psi}(2 \mid \theta) = 1$ for all $\theta \in A_2$. By construction, $\widetilde{\psi}$ is incentive-compatible and weakly increases S's ex-ante utility over $\overline{\psi}$. Next, let $\delta(\theta) := u(2,\theta) - u(1,\theta)$ and observe that

$$\int\limits_{\Theta} \delta(\theta)\widetilde{\psi}(2 \mid \theta)\,\mathrm{d}\mu_0(\theta) = \int\limits_{\Theta} \delta(\theta)\overline{\psi}(2 \mid \theta)\,\mathrm{d}\mu_0(\theta) + \int\limits_{A_2} (1 - \overline{\psi}(2 \mid \theta))\,\mathrm{d}\mu_0(\theta),$$

where the last term is non-negative, which means that obedience of $\overline{\psi}$ implies obedience of $\widetilde{\psi}$. Therefore, $\widetilde{\psi}$ is also a commitment outcome. ■

From the existing literature, we know even more about commitment outcomes when $|J| = 2$ and Θ is finite. Alonso and Câmara (2016) show that every commitment outcome is characterized by a cutoff state θ^*, and all states such that $\delta(\theta) > \delta(\theta^*)$ are pooled together to recommend action 2. In particular, all the good states $\theta \notin A_1$ are recommending action 2, which implies that every commitment outcome is incentive-compatible (see also Lemma B.2 in Koessler and Skreta, 2023). Our Proposition 1 also deals with the case when Θ is rich, in which case there exist commitment outcomes that are not IC (although they are IC μ_0-almost everywhere), and its proof describes how to make an existing commitment outcome incentive-compatible.

Theorem 2 is useful for checking if an already-found commitment outcome $\overline{\psi}$ is an equilibrium outcome. The answer is affirmative if and only if $\overline{\psi}$ is deterministic a.e. and incentive-compatible. While checking for incentive-compatibility can be straightforward, there is generally no guarantee that a deterministic commitment outcome exists. In the remainder of the section, we consider the cases when Θ is rich and finite separately. We show that when Θ is rich, there always exists a deterministic commitment outcome; if, additionally, R chooses between two actions, the (IC_θ) constraints are satisfied automatically, and therefore the commitment payoff is always achieved in an equilibrium. When Θ is finite, we obtain an approximation result.

4.1. Rich State Space

When the state space Θ is rich, i.e., Θ is a convex and compact subset of $\mathbb{R}^n$ with the prior μ_0 being atomless, the existence of a deterministic commitment outcome is guaranteed.

Proposition 2. *If Θ 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 $\overline{\psi}$ follows from the same argument as in the proof of Proposition 1. For every $j \in J$, $\overline{\psi}(j \mid \cdot) : \Theta \to [0,1]$ is Borel measurable, and $\sum_{j\in J}\overline{\psi}(j \mid \theta)=1$ for all $\theta\in\Theta$. For each $j\in J$, define μ_j by $\mathrm{d}\mu_j := u(j,\cdot)\,\mathrm{d}\mu_0$. Because μ_0 is a finite and atomless positive measure, and u is bounded, for all $j\in J$, μ_j is a finite and atomless signed measure. Since J is finite, by Theorem 2.1 in Dvoretzky, Wald, and Wolfowitz (1951), there exist Borel measurable functions $\widetilde{\psi}(j \mid \cdot):\Theta\to\{0,1\}$, $j\in J$ with $\sum_{j\in J}\widetilde{\psi}(j\mid\cdot)=1$, such that $\int_\Theta \widetilde{\psi}(j\mid\theta)\,\mathrm{d}\mu_0=\int_\Theta\overline{\psi}(j\mid\theta)\,\mathrm{d}\mu_0$, and $\int_\Theta\widetilde{\psi}(j\mid\theta)\,\mathrm{d}\mu_j=\int_\Theta\overline{\psi}(j\mid\theta)\,\mathrm{d}\mu_j$ for all $j\in J$. The second set of equalities implies that $\widetilde{\psi}$ is obedient since $\overline{\psi}$ is, and the first equality implies

$$V_{\widetilde{\psi}}=\int_\Theta\sum_{j\in J}v(j)\widetilde{\psi}(j\mid\theta)\,\mathrm{d}\mu_0=\int_\Theta\sum_{j\in J}v(j)\overline{\psi}(j\mid\theta)\,\mathrm{d}\mu_0=V_{\overline{\psi}}.$$

Therefore, $\widetilde{\psi}$ is a deterministic commitment outcome. The second part follows from Theorem 2. ■

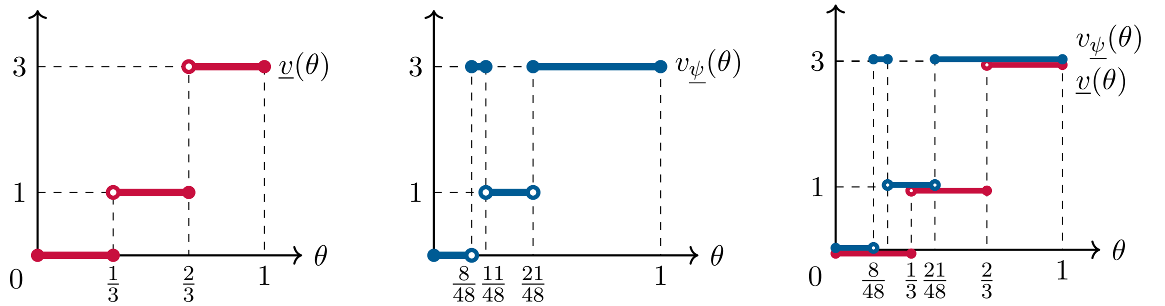
Checking if a deterministic commitment outcome with partition $\{W_j\}_{j\in J}$ is incentive-compatible (and is thus an equilibrium outcome) is straightforward and requires checking that $\theta\in W_j$ implies that $v(j)\geq\underline{v}(\theta)$ for all $\theta\in\Theta$. Consider the following example from Gentzkow and Kamenica (2016).

Example 2. Suppose that 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 and, given belief $\mu\in\Delta\Theta$, action 1 is optimal if and only if $\mathbb{E}_\mu[\theta]\geq 1/3$, action 2 — if and only if $\mathbb{E}_\mu[\theta]\in[1/3,2/3]$, action 3 — if and only if $\mathbb{E}_\mu[\theta]\geq 2/3$. Therefore, R's complete-information actions 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 $\overline{\psi}$ with an outcome partition $\overline{W}_1=[0,8/48)$, $\overline{W}_2=(11/48,21/48)$, and $\overline{W}_3=[8/48,11/48]\cup[21/48,1]$. This outcome is incentive-compatible, which we illustrate in Figure 1. Since $\overline{\psi}$ is a deterministic and IC commitment outcome, it is an equilibrium outcome by Proposition 2.

In the special case where R chooses between two actions, S always achieves his commitment payoff in an equilibrium.

Proposition 3. *If Θ 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 $\overline{\psi}$. We can then use the same argument as in the proof of Proposition 1 to find a deterministic commitment outcome $\widetilde{\psi}$ that is incentive-compatible. By Theorem 2, $\widetilde{\psi}$ is an equilibrium outcome. ■



(a) *S's lowest payoff under complete information.* ***(b)*** *S's expected payoff in the commitment outcome.* ***(c)*** *$\overline{\psi}$ is IC since $v_{\overline{\psi}}(\theta) \geq \underline{v}(\theta)$ for all $\theta \in [0,1]$.*

Figure 1. *Commitment outcome $\overline{\psi}$ is incentive-compatible 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., $\Theta = \{1, \ldots, N\}$, a deterministic commitment outcome often does not exist. For example, in Example 1, the *unique* commitment outcome is not deterministic. Therefore, S may not be able to achieve the commitment payoff in equilibrium.

However, we can show that when the state space is sufficiently rich (in the sense that $\mu_0(\theta)$ is sufficiently small for all $\theta \in \Theta$), then S's equilibrium payoff can get arbitrarily close to his commitment payoff. To make a clean argument, we adopt the assumptions of Alonso and Câmara (2016): R has a binary action and

$$\theta' \neq \theta'' \Longrightarrow \delta(\theta') \neq \delta(\theta''), \tag{RU}$$

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

Proposition 4. *Suppose that Θ is finite, $|J| = 2$, (RU) holds, and S's payoffs are normalized to $v(2) = 1$ and $v(1) = 0$.*[12] *Let V^* be S's commitment payoff. Then, for every $\varepsilon > 0$, there exists $\gamma > 0$ such that $\mu_0(\theta) < \gamma$ for all $\theta \in \Theta$ implies that there exists an equilibrium outcome α such that $|V^* - V_\alpha| < \varepsilon$.*

Proof. If $A_2 = \Theta$ (which means $\delta(\theta) \geq 0$ for all $\theta \in \Theta$) , let $\alpha(2 \mid \theta) = 1$ for all $\theta \in \Theta$ so that $V_\alpha = V^*$. Therefore, we assume for the remainder of the proof that A_2 is a proper subset of Θ. Since (RU) holds, we can use Proposition 2 in Alonso and Câmara (2016) to find a cutoff state $\theta^* \in \Theta$ such that $\delta(\theta^*) < 0$ and, for every

[12]The normalization is adopted without loss of generality, and although Condition (RU) is assumed to simplify the proof, the result holds even without it.

commitment outcome ψ, we have $\psi(2 \mid \theta) = 1$ ($\psi(1 \mid \theta) = 1$) for all $\theta \in \Theta$ such that $\delta(\theta) > \delta(\theta^*)$ ($\delta(\theta) < \delta(\theta^*)$). Next, consider a deterministic outcome α 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 α is IC and obedient, and therefore it is an equilibrium outcome by Theorem 1. The difference in ex-ante payoffs of S is zero if $\psi(2 \mid \theta^*) = 1$ and otherwise is $V^* - V_\alpha = \psi(2 \mid \theta^*)\mu_0(\theta^*) < \mu_0(\theta^*) < \gamma := \varepsilon$. ∎

Therefore, assuming R has a binary action, if the finite state space Θ has sufficiently many elements and the prior probability of each state is sufficiently small, S can get arbitrarily close to his commitment payoff in an equilibrium.

5. A Model with Stochastic Messages

In the main model, IC and obedience were not sufficient for an outcome to be induced by an equilibrium: there exist non-deterministic but IC and obedient outcomes in which S effectively recommends multiple actions and thus expects different payoffs in the same state, which violates equilibrium condition *(i)*. The reason why *(i)* is violated is that the mapping $E : \Theta \to M$, which determines the set of messages available in state θ, is deterministic. This assumption is standard in verifiable disclosure and cheap talk literature.[13] In some cases; however, it is reasonable to assume that that mapping 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 game with a stochastic message mapping (SMM game henceforth) and show that IC and obedience are sufficient for an outcome to be an equilibrium outcome.

The SMM game has the same timeline and players' objectives as our main model; we modify stage 2, wherein S communicates with R, only. Specifically, we assume that along with the state of the world $\theta \in \Theta$ (which we assume here is finite), S also observes a *label* $x \in [0, 1]$, which is payoff-irrelevant to both S and R. That label x is drawn from $U(X^\theta)$, where $\{X^\theta\}_{\theta\in\Theta}$ is a partition of the unit interval such that $\lambda(X^\theta) = \mu_0(\theta)$, where λ is the Lebesgue measure.[14] Then, S sends message $m \in \mathcal{B}([0, 1]) =: \widehat{M}$ such that $x \in$

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

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

m. Thus, the set of messages available to S in state θ is now determined stochastically (through x). The *equilibrium of the SMM game* $(\widehat{\sigma}, \widehat{\tau}, \widehat{q})$ is defined analogously to the equilibrium of the main model, except S's strategy is now also a function of x.

In the SMM game, S's strategy is a function $\widehat{\sigma} : \Theta \times [0,1] \to \Delta_0 \widehat{M}$, R's strategy is $\widehat{\tau} : \widehat{M} \to \Delta J$ and R's belief system is $\widehat{q} : \widehat{M} \to \Delta\Theta$.

Definition 5. *A triple* $(\widehat{\sigma}, \widehat{\tau}, \widehat{q})$ *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* $\underset{\{m \in \widehat{M} \mid x \in m\}}{\arg\max} \sum_{j \in J} v(j)\, \widehat{\tau}(j \mid m)$;

(ii) For all $m \in \widehat{M}$, $\widehat{\tau}(\cdot \mid m)$ *is supported on* $\underset{j \in J}{\arg\max} \int_{\Theta} u(j, \theta)\, \mathrm{d}q(\theta \mid m)$;

(iii) $\widehat{q}$ *is obtained from* μ_0, *given* $\widehat{\sigma}$, *using Bayes rule.*

(iv) For all $m \in \widehat{M}$, $\widehat{q}(\cdot \mid m) \in \Delta\{\theta \in \Theta \mid X^\theta \cap m \neq \varnothing\}$.

Since x is payoff-irrelevant, an outcome $\alpha \in \Psi$ is defined as before, and $\alpha(j \mid \theta)$ is the probability that R takes action $j \in J$ when the state is $\theta \in \Theta$. We say that α is an *equilibrium outcome* of the SMM game if there exists an equilibrium $(\widehat{\sigma}, \widehat{\tau}, \widehat{q})$ that induces it, i.e., $\alpha(j \mid \theta) = \frac{1}{\mu_0(\theta)} \int_{X^\theta} \sum_{m \in \operatorname{supp} \widehat{\sigma}(\cdot \mid \theta, x)} \widehat{\sigma}(m \mid \theta, x) \widehat{\tau}(j \mid m)\, \mathrm{d}x$.

We obtain a sharp equilibrium characterization for the SMM game.

Theorem 3. *Let* Θ *be finite. Then,* $\alpha \in \Psi$ *is an equilibrium outcome of the SMM game* $\iff$ α *is IC and obedient.*

Proof. ($\Longrightarrow$) 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.

($\Longleftarrow$) Consider an IC and obedient outcome α. For every $\theta \in \Theta$, let $J^\theta :=$ $\operatorname{supp} \alpha(\cdot \mid \theta)$ be the set of actions that R takes with a positive probability when the realized state is θ. Next, partition X^θ into a set of intervals $\{X_j^\theta\}_{j \in J^\theta}$ such that $\frac{\lambda(X_j^\theta)}{\lambda(X^\theta)} =$ $\alpha(j \mid \theta)$. Also, for each action $j \in J$, let $W_j := \bigcup_{j \in [J]} X_j^\theta$; 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 $\widehat{q}(\theta \mid W_j) = \frac{\lambda(X_j^\theta)}{\lambda(W_j)}$. Furthermore, since α is obedient, for every action $j \in [J]$ such that $\lambda(W_j) > 0$, we have

$$\sum_{\theta \in \Theta} \big(u(j,\theta) - u(j',\theta)\big)\alpha(j \mid \theta)\mu_0(\theta) \geq 0 \iff$$
$$\sum_{\theta \in \Theta} \big(u(j,\theta) - u(j',\theta)\big)\frac{\lambda(X_j^\theta)}{\lambda(W_j)} \geq 0 \quad \forall j' \in J \smallsetminus \{j\},$$

meaning that R prefers to take action j after message W_j, so we let $\widehat{\tau}(j \mid W_j) = 1$. 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, $\forall m \notin \{W_1, \ldots, W_J\}$, let $\widehat{q}(\cdot \mid m) \in \Delta A_{\underline{j}}$, where $\underline{j} \in J$ is the lowest action such that $m \cap X^\theta \neq \varnothing$ and $\theta \in A_i$. Then, playing action $\underline{j}$ is a best response to message m, so we let $\widehat{\tau}(\underline{j} \mid m) = 1$. Since α 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, while deviations to off-path messages are not profitable since the payoff from any deviation in state θ is at most $\underline{v}(\theta)$, which is below $v_\alpha(\theta)$ by the (IC_θ) constraint. Hence, $(\widehat{\sigma}, \widehat{\tau}, \widehat{q})$ is an equilibrium of the SMM game. ■

In contrast to Theorem 1, incentive-compatibility and obedience are necessary and sufficient for an outcome to be an equilibrium outcome of the SMM game. Two properties of the SMM model 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 the same state θ (but for different realizations of x). Secondly, the message space is "rich," meaning that for every vector $p = (p_1, \ldots, p_N) \in [0,1]^N$ there exists a message m that is available in state $\theta \in \Theta$ with probability p_θ. This richness allows us to "purify" non-deterministic outcomes — the equilibria that we construct to implement an IC and obedient outcome is in pure strategies of both S and R.

Given the sharp equilibrium characterization of the SMM game, we immediately obtain the following results.

Corollary 1. *Let Θ be finite. Then, a commitment outcome is an equilibrium outcome of the SMM game if and only if it is IC.*

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

Corollary 1 follows directly from Theorem 3. Corollary 2 follows from Alonso and Câmara (2016) who show that every commitment outcome is incentive-compatible (see also our discussion after Proposition 1) and Theorem 3 that ensures that every such outcome is an equilibrium outcome.

Here we can see that the set of IO outcomes from Koessler and Skreta (2023) is a subset of the set of equilibrium outcomes of the SMM game. Furthermore, if S's value function is quasiconvex in R's belief, or when R chooses between two actions, then these two sets coincide.

6. Conclusion

This paper considers a persuasion game with verifiable information, in which a sender with transparent motives chooses which verifiable messages to send to a receiver to convince her to take a particular action from a finite set. We show that every equilibrium outcome must be incentive-compatible for the sender and obedient for the receiver. These conditions are also sufficient when we focus on equilibria in which the receiver does not mix or when the sender's message mapping is stochastic. We also identify sufficient conditions under which the ex-ante commitment assumption in Bayesian persuasion can be replaced by communication with verifiable information. Specifically, if the sender receives at least his complete information payoff in every state, then that commitment outcome is an equilibrium outcome of a verifiable disclosure game. If the receiver chooses between two actions, that 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. 3, 16.)

Boleslavsky, 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. 3.)

Dranove, David and Ginger Zhe Jin (2010), "Quality Disclosure and Certification: Theory and Practice", *Journal of Economic Literature*, 48, 4, pp. 935-963. (p. 2.)

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. 3.)

— (2006), "A Study in the Pragmatics of Persuasion: a Game Theoretical Approach", *Theoretical Economics*, 1, pp. 395-410. (p. 3.)

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. 2, 16.)

Hart, Sergiu, Ilan Kremer, and Motty Perry (2017), "Evidence Games: Truth and Commitment", *American Economic Review*, 107 (3 Mar. 2017), pp. 690-713. (pp. 3, 16.)

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian Persuasion", *American Economic Review*, 101, 6 (Oct. 2011), pp. 2590-2615. (pp. 2, 9, 11, 12.)

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. 4, 10, 13, 19.)

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. 3.)

Lipnowski, Elliot and Doron Ravid (2020), "Cheap Talk With Transparent Motives", *Econometrica*, 88, 4, pp. 1631-1660. (p. 3.)

Milgrom, Paul R. (1981), "Good News and Bad News: Representation Theorems and Applications", *The Bell Journal of Economics*, 12, 2, p. 380. (pp. 2, 16.)

Milgrom, Paul R. (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. 2.)

Milgrom, Paul R. and John Roberts (1986), “Relying on the Information of Interested Parties”, *The RAND Journal of Economics*, 17, 1, p. 18. (pp. 2, 5, 10, 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. 4.)

Sher, Itai (2011), “Credibility and Determinism in a Game of Persuasion”, *Games and Economic Behavior*, 71 (2 Mar. 2011), pp. 409-419. (p. 3.)

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. 4.)

Zhang, Kun (2022), “Withholding Verifiable Information”, working paper. (p. 1.)

Zhang, Kun and Maria Titova (2025), “Withholding Verifiable Information”, working paper. (p. 3.)