Persuasion with Verifiable Information*

Maria Titova†

November 2023

Abstract

This paper studies a game in which an informed sender with state-independent preferences uses verifiable messages to convince a receiver to take an action out of a finite set. I find that every equilibrium is deterministic and is characterized by a partition of the state space into sets on which the receiver takes a particular action. My first result is that a partition of the state space is an equilibrium outcome if and only if each action set satisfies the sender's incentive-compatibility and the receiver's obedience constraints. My second result states that an ex-ante optimal commitment outcome is an equilibrium outcome of the game with verifiable information if and only if the sender receives at least his complete information payoff in every state, as long as the state space or the message space is sufficiently rich.

KEYWORDS: Persuasion, Evidence, Value of Commitment

JEL CLASSIFICATION: D72, D82, D83, M37

*I am very grateful to Renee Bowen and Joel Sobel for their guidance and support. I thank S. Nageeb Ali, Arjada Bardhi, Simone Galperti, Germán Gieczewski, Navin Kartik, Remy Levin, Aleksandr Levkun, Elliot Lipnowski, Denis Shishkin, Joel Watson, and the audience members at Amsterdam, Bern, BSE, CERGE-EI, CIDE, Princeton, Manchester, Vanderbilt, UCSD, SWET 2018, YES 2019, EGSC 2019, CSQIEP 2020 and SITE 2021 for their helpful comments. All errors are my own.

†Department of Economics, Vanderbilt University. E-mail: motitova@gmail.com.

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? On average, what is the best outcome that the sender can hope for?

Persuasion with verifiable information plays an essential role in courtrooms, electoral campaigns, product advertising, financial disclosure, job market signaling, and many other economic situations. 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 in order to obtain higher compensation. In a labor market, a job candidate lists specific certifications in order to make his application more attractive to an employer.

I consider the following model of persuasion with verifiable information. First, the sender (he) 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). Verifiability means that the messages contain the truth (the true state of the world), but not necessarily the whole truth (they may contain other states, as well). Upon observing the message, the receiver takes action j out of a finite set. The sender's preferences are state-independent and strictly increase in the receiver's action. The receiver's preferences depend on both her action and the state.

The first question is: what is the set of equilibrium outcomes of this game? To answer it, I consiider the mapping from the state space into the receiver's actions to analyze what decision the receiver makes in equilibrium depending on the realized state. This mapping must be deterministic, or the sender would have a profitable deviation toward a message that induces the higher action. Hence, every equilibrium forms a partition of the state space into action-j sets. My first result states that a collection of action-j sets is an equilibrium partition if and only if it satisfies the sender's incentive-compatibility (IC) and the receiver's obedience constraints. The sender's IC constraint states that the sender receives at least his complete information payoff in each state of the world. That is, the only deviations of the sender we need to rule out are those toward fully revealing messages. The receiver's obedience constraint is standard in the communication and information design literature and ensures that the receiver maximizes her expected payoff. Due to the particular assumption on the structure of messages, it is easy to implement any equilibrium partition by having the sender reveal which action-j set the state of the world belongs to. I rank equilibria by the sender's ex-ante utility. The sender-worst equilibrium is outcome equivalent to full disclosure, meaning that the receiver acts as if fully informed. The sender-preferred equilibrium outcome partition maximizes his ex-ante utility across all incentive-compatible and obedient partitions of the state space.

In my model, the sender does not have (ex-ante) commitment power: he learns the state then chooses a (verifiable) message that maximizes his expected payoff in that state. The second question I ask is: how much better off would the sender be if he had commitment power? To answer it, I compare the sender-preferred equilibrium outcome to the (optimal) commitment outcome of the Bayesian persuasion game (Kamenica and Gentzkow, 2011), in which the sender commits to which (non-verifiable) signals are sent from which states ahead of learning the state. Two differences arise.

The first difference between equilibrium and commitment outcomes is that the latter need not form a partition of the state space. Indeed, the commitment outcome in the leading example of Kamenica and Gentzkow, 2011 is not deterministic: the judge convicts a fraction $3/7 \in (0, 1)$ of innocent defendants. Clearly, a non-deterministic commitment outcome cannot be implemented in equilibrium. I argue, however, that this is not a substantive concern. When the state space is rich, a deterministic commitment outcome exists. If the state space is finite, I show that augmenting the game with a rich space of (verifiable) messages allows us to implement any (incentive-compatible but not necessarily deterministic) commitment outcome.

The second difference between equilibrium and communication outcomes is more substantive: the sender does not face incentive-compatibility concerns when he has commitment power. Therefore, my second result states that a (deterministic) commitment outcome is an equilibrium outcome if and only if the sender obtains at least his complete information payoff in every state. Additionally, when the receiver chooses between two actions, a (deterministic) commitment outcome is always an equilibrium outcome because the sender convinces the receiver as often as possible and thus never wants to deviate to full disclosure.

Related Literature

The literature on persuasion with verifiable information was pioneered by Milgrom (1981) and Grossman (1981), and reviewed by Milgrom (2008); this paper has the same assumption on verifiability of messages as Milgrom and Roberts (1986). In all these papers, unraveling is the unique equilibrium outcome, because the sender's preferences are monotone in receiver's action (e.g. he is maximizing quantity sold) and receiver's action space is rich (e.g. she is choosing an infinitely divisible quantity to buy). The unraveling 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 that convinces the receiver to purchase a (strictly) higher quantity of the product. I argue that if the receiver's action space is finite, then unraveling is not the only equilibrium outcome. This argument is easiest to see when the receiver'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.

While I establish the value of sender's commitment power, there is growing literature studying the value of receiver's commitment power and compares 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's preferences depend on the state of the world.

A recent strand of literature asks the same questions as me (What is the equilibrium set? What is the sender's value of commitment?) but for other communication protocols. Most closely related to mine are Chakraborty and Harbaugh (2010) and Lipnowski and Ravid (2020)' analyses of cheap talk games and Koessler and Skreta (2023) and Zapechelnyuk (2023)' analyses of informed information design. I compare my results to theirs in Section 5. In short, whether the sender prefers cheap talk, verifiable messages, or disclosure mechanisms depends on players' preferences.

There is an extensive applied Bayesian persuasion literature. In many settings, the receiver chooses between two actions. 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); matching platforms persuade sellers to match with buyers (Romanyuk and Smolin, 2019); politicians persuade voters (Alonso and Câmara, 2016; Bardhi and Guo, 2018); and governments persuade citizens (Gehlbach and Sonin, 2014). My contribution is to show that in these applications, one can replace the assumption that the sender has commitment power with the assumption that the sender's messages are verifiable.

2. Model

I study a game of persuasion with verifiable information between a sender (S, he/him) and a receiver (R, she/her). Below I describe the timing of the game along with the technical restrictions:[1]

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 ($\Theta = [0,1]$). The state of the world is drawn from a common prior $\mu_0 \in \Delta\Theta$ with $\operatorname{supp} \mu_0 = \Theta$. If the state space is rich, I further require that the prior is atomless.
2. S sends message $m \in M := \mathcal{B}(\Theta)$ to the receiver.
   Each message is a subset of the state space and we can interpret it as a statement about the state of the world. Sender's messages are verifiable:[2]

   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 $j \in [J] := \{1, \ldots, J\}$, where $J \geq 2$.
4. Game ends, payoffs are realized.
   S's payoff $u_S : [J] \to \mathbb{R}$ depends only on R's action. I assume that S prefers higher actions of R, i.e. $u_S(j)$ is strictly increasing in $j \in [J]$.
   R's preferences are described by an integrable utility function $u : [J] \times \Theta \to \mathbb{R}$. I define R's *complete information action-j set* as $A_j := \big\{\theta \in \Theta \mid u(j, \theta) \geq u(j', \theta)$ for all $j' \in [J] \setminus \{j\}\big\}$ to include all the states of the world in which she prefers to take action j under complete information. I assume that $\{A_j\}_{j \in [J]}$ is a partition of the state space, meaning that R has a unique best action under complete information. I denote the set of all partitions of a set Y into at most K non-empty sets by $\mathcal{P}_K(Y)$.[3]

I consider perfect Bayesian equilibria of this game. Firstly, S's strategy $\sigma : \Theta \times M \to [0,1]$ is a regular conditional probability, such that *1)* for every $\theta \in \Theta$, $\sigma(\cdot \mid \theta)$ is a probability measure over M, and *2)* for every $m \in M$, $\sigma(m \mid \cdot)$ is measurable. Secondly, R's strategy is a function $a : M \to [J]$. Finally, R's posterior is a regular conditional probability $q : M \times \Theta \to [0,1]$, such that *1)* for every $m \in M$, $q(\cdot \mid m)$ is a probability measure over Θ, and *2)* for every $\theta \in \Theta$, $q(\theta \mid \cdot)$ is measurable.

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

(i) $\forall \theta \in \Theta$, $\sigma(\cdot \mid \theta)$ is supported on $\arg\max_{m \in M} u_s\big(a(m)\big)$, subject to $\theta \in m$;

---

[1] For a compact metrizable space Y, I let ΔY denote the set of all Borel probability measures over Y, endowed with the weak* topology. For $\gamma \in \Delta Y$, I let $\operatorname{supp} \gamma$ denote the support of γ.

[2] I borrow from Milgrom and Roberts (1986) the definition of a verifiable message as a subset of the state space that includes the true realization. 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.

[3] Formally, $\mathcal{P}_K(Y) := \Big\{\{T_k\}_{k=1}^K \subseteq Y^K \mid \bigcup_{k=1}^K T_k = Y,\ T_k \neq T_l \implies T_k \cap T_l = \varnothing\Big\}$.

(ii) $\forall m \in M$, $a(m) \in \arg\max_{j\in[J]} \int u(j,\theta) dq(\theta \mid m)$;[4]

(iii) q *is obtained from* μ_0*, given* σ*, using Bayes rule;*[5]

(iv) $\forall m \in M$, $\operatorname{supp} q(\cdot \mid m) \subseteq m$.

In words, the equilibrium conditions state that *(i)* S chooses verifiable messages that maximize his 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; *(iv)* R places zero probability on the states in which the observed message is not verifiable.

To analyze the model, I use the following state-based approach. I let an outcome of the game specify the probability that the receiver takes action j if the realized state of the world is θ. Let Ψ be the set of all measurable maps from Θ to $\Delta[J]$.

Definition 3.

- *An* outcome $\alpha \in \Psi$ *specifies, for each state* $\theta \in \Theta$*, the probability* $\alpha(j \mid \theta)$ *that R takes action* $j \in [J]$.
- *An outcome* α *is an* equilibrium outcome *if it corresponds to some equilibrium.*[6]

Some outcomes are deterministic, meaning that R takes a particular action with certainty in every state of the world. For such outcomes, we can partition the state space into the sets on which R takes the same action.

Definition 4.

- *An outcome* α *is* deterministic *if, for every* $\theta \in \Theta$, $\alpha(j \mid \theta) = 1$ *for some* $j \in [J]$.
- *The* outcome partition $\{W_j\}_{j\in[J]} \in \mathcal{P}_J(\Theta)$ *of a deterministic outcome* α *satisfies* $W_j = \{\theta \in \Theta \mid \alpha(j \mid \theta) = 1\}$ *for each* $j \in [J]$.

Given outcome α, S's ex-ante utility is $U_S(\alpha) := \int_\Theta \sum_{j\in[J]} u_S(j)\alpha(j \mid \theta) d\mu_0(\theta)$. If α is a deterministic outcome with partition $\{W_j\}_{j\in[J]}$, then $\alpha(j \mid \theta) = \mathbb{1}(\theta \in W_j)$ and S's ex-ante utility simplifies to $U_S(\alpha) = \sum_{j\in[J]} u_S(j)\mu_0(W_j)$.

---

[4] I restrict attention to the pure strategies of R for ease of exposition as it does not influence the equilibrium range. The receiver is never indifferent in the sender-worst equilibrium, and breaks ties in favor of a higher action in the sender-preferred equilibrium.

[5] That is, $\int_{\widehat{\Theta}} \sigma(\widehat{M} \mid \theta) d\mu_0(\theta) = \int_\Theta \int_{\widehat{M}} q(\widehat{\Theta} \mid m) d\sigma(m \mid \theta) d\mu_0(\theta)$ for every Borel $\widehat{\Theta} \subseteq \Theta$ and $\widehat{M} \subseteq M$.

[6] Specifically, α is an equilibrium outcome if there exists an equilibrium (σ, a, q) such that $\forall \theta \in \Theta$, $\alpha(j \mid \theta) = \sigma(\mathcal{M}_j \mid \theta)$, where $\mathcal{M}_j := \{m \in M \mid a(m) = j\}$ is the set of messages that convince R to take action j.

3. Analysis

Direct Implementation

Consider an outcome α. Observe that in equilibrium, S must receive at least his complete information payoff in every state of the world. Indeed, if $\theta \in A_j$ for some action $j \in [J]$, then S can send a fully informative message $\{\theta\}$, reveal that the state is θ, thus convincing R to take action j. Therefore, we must require that in every equilibrium, if in state θ the receiver takes action j, then θ must belong to a complete-information action set of at most action j. This observation gives rise to S's incentive compatibility constraint for action $j \in [J]$:

$$\alpha(j \mid \theta) > 0 \implies \theta \in A_1 \cup \ldots \cup A_j. \tag{IC$_j$}$$

Definition 5. *An outcome α is* incentive-compatible *(for S) if* (IC$_j$) *constraint holds for every action $j \in [J]$.*

Next, if R takes action j in state θ with positive probability, then that action must maximize her expected utility. Thus we obtain R's obedience constraint for action $j \in [J]$:

$$\int_{\Theta} \big(u(j,\theta) - u(j',\theta)\big)\alpha(j \mid \theta) d\mu_0(\theta) \geq 0, \quad \forall j' \in [J] \setminus \{j\} \tag{obedience$_j$}$$

Definition 6. *An outcome α is* obedient *(for R) if* (obedience$_j$) *constraint holds for every action $j \in [J]$.*

If α is a deterministic outcome with partition $\{W_j\}_{j\in[J]} \in \mathcal{P}_J(\Theta)$, then (IC$_j$) becomes $\theta \in W_j \implies \theta \in A_1 \cup \ldots \cup A_j$ or simply $W_j \subseteq A_1 \cup \ldots \cup A_j$. Similarly, (obedience$_j$) simplifies to $\int_{W_j} \big(u(j,\theta) - u(j',\theta)\big) d\mu_0(\theta) \geq 0$ for all $j' \in [J] \setminus \{j\}$.

The first result of this paper allows us to characterize all equilibrium outcomes in terms of partitions of the state space into action sets, each of which satisfies these two constraints.

Theorem 1. *Every equilibrium outcome is deterministic. Moreover, $\{W_j\}_{j\in[J]} \in \mathcal{P}_J(\Theta)$ is an equilibrium outcome partition if and only if, for every action $j \in [J]$,*

$$W_j \subseteq A_1 \cup \ldots \cup A_j,$$
$$\int_{W_j} \big(u(j,\theta) - u(j',\theta)\big) d\mu_0(\theta) \geq 0, \quad \forall j' \in [J] \setminus \{j\}.$$

Proof. In this proof, (σ, a, q) is an equilibrium and α is the corresponding equilibrium outcome. For each action $j \in [J]$, let $\mathcal{M}_j := \{m \in M \mid a(m) = j\}$ be the set of messages that convince R to take action j; then, $\alpha(j \mid \theta) = \sigma(\mathcal{M}_j \mid \theta)$.

First, observe that every equilibrium outcome is deterministic. To verify this, suppose instead that in some state $\theta \in \Theta$, the receiver takes actions $j < j' \in [J]$ with positive probability, i.e. $\alpha(j \mid \theta), \alpha(j' \mid \theta) > 0$. Now, since S prefers higher actions of R, $u_S(j) < u_S(j')$. Consequently, S is not maximizing his expected utility by sending messages from $\mathcal{M}_j$ in state θ, a contradiction.

Next, consider an equilibrium (σ, a, q) with an outcome partition $\{W_j\}_{j\in[J]} \in \mathcal{P}_J(\Theta)$. For each action $j \in [J]$, the set W_j must satisfy S's IC constraint, or else S can deviate to full disclosure. It remains to show that each W_j also satisfies R's obedience constraint.

Notice that since α is a deterministic outcome, $\alpha(j \mid \theta) = \mathbb{1}(\theta \in W_j) = \sigma(\mathcal{M}_j \mid \theta)$. In words, if $\theta \in W_j$, S convinces R to take action $j \in [J]$ with certainty, so he must only be sending messages from $\mathcal{M}_j$. Now, since every message $m \in \mathcal{M}_j$ convinces R to take action $j \in [J]$, from equilibrium condition *(ii)* we have

$$\forall m \in \mathcal{M}_j, \quad \int\limits_{\Theta} \big(u(j,\theta) - u(j',\theta)\big) dq(\theta \mid m) \geq 0, \quad \forall j' \in [J] \setminus \{j\}.$$

Using Bayes rule, we conclude that W_j satisfies the obedience constraint:

$$\begin{aligned}
\forall m \in \mathcal{M}_j, \quad &\int\limits_{\Theta} \big(u(j,\theta) - u(j',\theta)\big) \sigma(m \mid \theta) d\mu_0(\theta) \geq 0 \\
\implies &\int\limits_{\Theta} \big(u(j,\theta) - u(j',\theta)\big) \underbrace{\sigma(\mathcal{M}_j \mid \theta)}_{=\mathbb{1}(\theta\in W_j)} d\mu_0(\theta) \geq 0 \\
&\implies \int\limits_{W_j} \big(u(j,\theta) - u(j',\theta)\big) d\mu_0(\theta) \geq 0, \quad \forall j' \in [J] \setminus \{j\}.
\end{aligned}$$

Finally, consider a partition $W = \{W_j\}_{j\in[J]} \in \mathcal{P}_J(\Theta)$ such that, for every $j \in [J]$, the set W_j satisfies the two constraints. Then we can construct an equilibrium (σ, a, q) that directly implements W as an outcome partition. Formally, let S's strategy be $\sigma(m \mid \theta) = \mathbb{1}(m = W_j \text{ and } \theta \in W_j)$. On the path, when R receives message W_j, she learns that $\theta \in W_j$ and nothing else, and takes action j by the obedience constraint. 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, $\forall m \notin W$, let supp $q(\cdot \mid m) \subseteq m \cap A_i$, where $i \in [J]$ is the lowest action such that the set $m \cap A_i$ is non-empty.

Now, S has no profitable deviations. To see this, consider a state $\theta \in W_j$ for some $j \in [J]$. Clearly, S cannot send any other on path message $W_i \in W$ because W is a partition: $\theta \in W_j$ implies $\theta \notin W_i$ so that another W_i is not a verifiable message in state θ. At the same time, if S deviates to an off-path message $m \notin W$, then the skeptical receiver will take at most action j: by (IC_j), if $\theta \in W_j$, then θ must be in the complete-information action-i set for some $i \leq j$. Thus, we have an equilibrium with an outcome partition W. □

Theorem 1 states that a deterministic obedient outcome is an equilibrium outcome if and only if the sender's payoff in every state exceeds his full disclosure payoff. This formulation complements Koessler and Renault (2012) who obtain a similar result (Proposition 2) in a game wherein R has private information which may be correlated with the state of the world, and S chooses a verifiable message along with the price of his product.

Equilibrium Range

For the purposes of characterizing equilibrium outcomes, Theorem 1 allows us to restrict attention to partitions $\{W_j\}_{j\in[J]}$ of the state space Θ such that each action set W_j satisfies the IC and obedience constraints. I rank equilibria by the sender's ex-ante utility.

The sender-worst equilibrium outcome minimizes S's ex-ante utility across all equilibria. Therefore, the sender-worst equilibrium partition is $\{A_j\}_{j\in[J]}$, meaning that S simply reveals which complete information action-j set the state of the world is in. Note that this equilibrium is outcome-equivalent to *full disclosure* (or *full unraveling*), which is salient in the verifiable-information literature.

The sender-preferred equilibrium outcome $\overline{\alpha}$ maximizes S's utility across all equilibria and solves

$$
\begin{aligned}
&\max_{\alpha\in\Psi} \int_{\Theta} \sum_{j\in[J]} u_S(j)\alpha(j \mid \theta)d\mu_0(\theta) \\
&\text{subject to } \{W_j\}_{j\in[J]} \in \mathcal{P}_J(\Theta) \text{ and, for each action } j \in [J], \\
&\qquad \alpha(j \mid \theta) = \mathbb{1}(\theta \in W_j), \qquad\qquad (1)\\
&\qquad \alpha(j \mid \theta) > 0 \implies \theta \in A_1 \cup \ldots \cup A_j, \\
&\qquad \int_{\Theta} \big(u(j,\theta) - u(j',\theta)\big)\alpha(j \mid \theta)d\mu_0(\theta) \geq 0, \quad \forall j' \in [J] \setminus \{j\}.
\end{aligned}
$$

Here, the constraints ensure that the outcome is deterministic, and that each element of the outcome partition satisfies the IC and obedience constraints.

4. Value of Commitment

To establish the value of commitment, let us compare S's maximal ex-ante utility in the game of persuasion with verifiable information (the model so far) and the the game of Bayesian persuasion (Kamenica and Gentzkow, 2011). In Bayesian persuasion, 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 with S committing to an experiment. Specifically, S can now choose a set of signals $\mathcal{S}$ and commit to a measurable map from Θ to $\Delta\mathcal{S}$. A triple consisting of S's experiment, R's strategy, and R's belief system is a *commitment protocol* if it satisfies equilibrium conditions *(ii)* and *(iii)*, replacing the message space M with the signal space $\mathcal{S}$. Importantly, S who has commitment power no longer faces an incentive compatibility constraint: he need not maximize his utility state-by-state, and his messages need not be verifiable. Like most of the literature, I focus on the sender-optimal commitment protocol, in which his ex-ante utility is the highest. For brevity, I drop the mention of sender optimality when discussing commitment protocols.

To find a commitment outcome, I employ the revelation principle and focus on straightforward signals that are interpreted by R as action recommendations (Kamenica and Gentzkow, 2011). A *commitment outcome* $\overline{\psi} \in \Psi$ solves

$$\begin{aligned} \max_{\psi\in\Psi} & \int\limits_{\Theta} \sum_{j\in[J]} u_S(j)\psi(j \mid \theta)d\mu_0(\theta) \\ & \text{subject to, for each action } j \in [J], \\ & \int\limits_{\Theta} \big(u(j,\theta) - u(j',\theta)\big)\psi(j \mid \theta)d\mu_0(\theta) \geq 0, \quad \forall j' \in [J] \setminus \{j\}. \end{aligned} \tag{2}$$

Now, comparing a sender-preferred equilibrium outcome $\overline{\alpha}$ to a commitment outcome $\overline{\psi}$ boils down to comparing solutions to problems (1) and (2). Upon close inspection of the two problems, we can conclude that problem (1) is problem (2) with two additional constraints. Specifically, commitment outcomes need not be deterministic; even deterministic commitment outcomes need not be incentive-compatible for S.

Rich State Space

When the state space is rich, there always exists a deterministic commitment outcome. Consequently, a commitment outcome is an equilibrium outcome of the verifiable information game if and only it is incentive-compatible for S.

THEOREM 2. *If $\Theta = [0,1]$, then a deterministic commitment outcome exists. Furthermore, a deterministic commitment outcome $\overline{\psi}$ is an equilibrium outcome if and only if it is incentive-compatible for S, or*

$$\overline{\psi}(j \mid \theta) > 0 \implies \theta \in A_1 \cup \ldots \cup A_j.$$

Proof. To see why there exists a deterministic commitment outcome, consider Problem (2). It is a linear problem with a finite number of linear constraints, so the solution is an extreme point of the constraint set. Since the state space is $[0,1]$ and μ_0 is atomless, it is well-known that the extreme points are deterministic.[7] The second part of the theorem follows from Theorem 1. □

Theorem 2 is useful for checking whether an already-found commitment outcome can be implemented with verifiable messages and without commitment. The answer is affirmative if and only if S receives at least his complete information payoff in every state of the world. Consider the following example from Gentzkow and Kamenica (2016).

EXAMPLE 1. The receiver has three actions, $[J] = \{1,2,3\}$ and the prior is uniform on $\Theta = [0,1]$. The sender's utility function is $u_S(1) = 0$, $u_S(2) = 1$, $u_S(2) = 3$. The receiver's preferences only depend on the posterior mean and, given belief $\mu \in \Delta\Theta$, she prefers to take action 1 if $\mathbb{E}_\mu[\theta] < 1/3$; action 2 if $\mathbb{E}_\mu[\theta] \in [1/3, 2/3)$; action 3 if $\mathbb{E}_\mu[\theta] \geq 2/3$. Therefore, R's complete-information actions sets are $A_1 = [0, 1/3)$, $A_2 = [1/3, 2/3)$, $A_3 = [2/3, 1]$.

The authors find the following commitment outcome partition: $\overline{W}_1 = [0, 8/48)$; $\overline{W}_2 = (11/48, 21/48)$; $\overline{W}_3 = [8/48, 11/48] \cup [21/48, 1]$. This outcome is IC for S: if $\theta \in A_1$, R takes actions 1, 2, or 3. If $\theta \in A_2$, R takes actions 2 or 3. If $\theta \in A_3$, R takes action 3. Consequently, S does not have a profitable deviation toward a fully informative message in any state $\theta \in [0,1]$. We conclude that this commitment outcome is also an equilibrium outcome.

When R chooses between two actions, S always reaches his commitment payoff in his most preferred equilibrium.

COROLLARY 1. *If $\Theta = [0,1]$ and $[J] = \{1,2\}$, then the sender-preferred outcome is a commitment outcome.*

Proof. Consider a deterministic commitment outcome $\overline{\psi}$ with partition $\{\overline{W}_1, \overline{W}_2\} \in \mathcal{P}_J(\Theta)$. There are two IC constraints: firstly, $\overline{W}_2 \subseteq A_1 \cup A_2 = \Theta$, which is trivially

[7] I thank the associate editor for pointing this out.

satisfied. Secondly, $\overline{W}_1 \subseteq A_1 \iff \Theta \setminus \overline{W}_1 \supseteq \Theta \setminus A_1 \iff A_2 \subseteq \overline{W}_2$. Clearly, $\overline{W}_2$ satisfies $A_2 \subseteq \overline{W}_2$, or else the partition $\{\overline{W}_2 \setminus A_2, \overline{W}_2 \cup A_2\}$ increases S's ex-ante utility while still satisfying the obedience constraints. Therefore, $\{\overline{W}_1, \overline{W}_2\}$ is also a sender-preferred outcome partition. □

Intuitively, when R chooses between two actions, and S strictly prefers one of them, an ex-ante optimal obedient outcome is also incentive-compatible, essentially because the sender convinces the receiver to take the higher action as often, as possible. Indeed, Koessler and Skreta (2023) obtain a similar result (Proposition 4) when the designer chooses disclosure mechanisms, instead of messages. This result remains true in binary action environments with multiple receivers under certain conditions on the interdependence between the receivers' actions and payoffs, which are satisfied in much of information design literature (see Koessler and Skreta, 2023 and references therein).

Finite State Space

Suppose that $\Theta = \{1, \dots, N\}$ is finite. Consider a non-deterministic commitment outcome $\overline{\psi}$ for which there exists a state θ and two actions $j < j' \in [J]$ such that $\overline{\psi}(j \mid \theta), \overline{\psi}(j' \mid \theta) \in (0,1)$. Clearly, $\overline{\psi}$ cannot be an equilibrium outcome , because S has a profitable deviation to only sending messages that induce action j' in state θ. To prevent the sender from such deviations, we can augment the (original) game with a rich message space that provides the sender with (infinitely) many ways of verifying his statements about the state of the world.

Augmented Message Space

Suppose that at the first stage of the game, the sender observes $\theta \in \Theta = \{1, \dots, N\}$ and $x \sim U[X^\theta]$, where $X^\theta := [t_{\theta-1}, t_\theta]$, $t_0 := 0$, and $t_\theta := \sum_{\theta'=1}^{\theta} \mu_0(\theta')$. At the second stage, he chooses a verifiable message $m \in \mathcal{B}([0,1])$ such that $x \in m$. Then, the receiver forms a posterior belief about θ and acts. I refer to the game with an augmented message space as the *augmented game*.

Proposition 1. *Let Θ be finite. Then, a commitment outcome $\overline{\psi}$ is an equilibrium outcome of the augmented game if and only if $\overline{\psi}$ is incentive-compatible for S, or*

$$\overline{\psi}(j \mid \theta) > 0 \implies \theta \in A_1 \cup \ldots \cup A_j.$$

Proof. ($\Longleftarrow$) Consider an incentive-compatible commitment outcome $\overline{\psi}$. For every $\theta \in \Theta$, let $J^\theta := \operatorname{supp} \overline{\psi}(\cdot \mid \theta)$ be the set of actions that R takes with positive probability in state θ. Then, partition X^θ into a set of intervals $\{X_j^\theta\}_{j\in J^\theta}$ such that $\frac{\lambda(X_j^\theta)}{\lambda(X^\theta)} = \overline{\psi}(j \mid \theta)$, where $\lambda(X)$ is the Lebesgue measure of $X \subseteq [0,1]$. Next, for each action $j \in [J]$, let $W_j := \bigcup_{j\in[J]} X_j^\theta$. Clearly $\{W_j\}_{j\in[J]}$ is a partition of X.

Next, consider the following strategy of S in the augmented game, $\sigma(m \mid \theta, x) = \mathbb{1}(m = W_j \text{ and } x \in W_j)$. In words, suppose that S reveals to R which W_j the observed x belongs to. Then, R's posterior after message W_j is $q(\theta \mid W_j) = \frac{\lambda(X_j^\theta)}{\lambda(X^\theta)}$. Furthermore, since $\overline{\psi}$ is obedient, for every action $j \in [J]$ such that $\lambda(W_j) > 0$,

$$\sum_{\theta\in\Theta} \big(u(j,\theta) - u(j',\theta)\big) \overbrace{\overline{\psi}(j \mid \theta)\mu_0(\theta)}^{=\lambda(X_j^\theta)} \geq 0 \iff$$

$$\sum_{\theta\in\Theta} \big(u(j,\theta) - u(j',\theta)\big) \frac{\lambda(X_j^\theta)}{\lambda(W_j)} \geq 0 \quad \forall j' \in [J] \setminus \{j\},$$

meaning that R prefers to take action j after message W_j. 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 $\operatorname{supp} q(\cdot \mid m) \subseteq A_i$, where $i \in [J]$ is the lowest action such that $m \cap X^\theta \neq \varnothing$ and $\theta \in A_i$. Now, S does not have profitable deviations. Indeed, he receives at least his complete information payoff in every state because $\lambda(X_j^\theta) > 0 \implies \theta \in A_1 \cup \ldots \cup A_j$ from the incentive-compatibility of $\overline{\psi}$. S would also not benefit from deviating to an off path message in any state, because then S would get at most his complete information payoff. Hence, the described strategies and beliefs constitute an equilibrium of the augmented game, and this equilibrium outcome is the same as the commitment outcome, since $\Pr(j \mid \theta) = \frac{\lambda(X_j^\theta)}{\lambda(X^\theta)} = \overline{\psi}(j \mid \theta)$.

($\Longrightarrow$) Suppose that $\overline{\psi}$ is not IC, meaning that there exist an action $j \in [J]$ and a state $\theta \in \Theta$ such that $\overline{\psi}(j \mid \theta) > 0$ and $\theta \in A_{j+1} \cup \ldots \cup A_J$. Clearly, $\overline{\psi}$ is not an equilibrium outcome, because S has a profitable deviation toward revealing x in state θ. Thus, if $\overline{\psi}$ is not IC, then $\overline{\psi}$ is not an equilibrium outcome.

□

5. Discussion and Conclusion

This paper considered a game in which the sender learns the state of the world and then chooses which verifiable message to send to the receiver in order to convince her to take a particular action out of a finite set. I characterized the full set of equilibrium outcomes and ranked them by the sender's ex-ante utility. The sender-worst equilibrium outcome is equivalent to full disclosure. The sender reaches his commitment payoff in his most preferred equilibrium whenever he does not have profitable deviations toward a fully informative strategy when he has commitment power. Below I compare my results to the recent findings in informed information design and the cheap talk literature. I focus my comparison on setups wherein there is one receiver and the sender has state-independent preferences, although some of the papers I mention below are more general.

Comparison to Informed Information Design. Koessler and Skreta (2023) and Zapechelnyuk (2023) consider a game in which the sender learns the state of the world and then chooses a set of signals $\mathcal{S}$ and a disclosure mechanism $\Theta \to \Delta\mathcal{S}$ to convince the receiver to take a particular action. The key difference in gameplay is that their sender (who I refer to as the designer) has *interim* commitment power, while mine does not. Specifically, once the designer chooses a mechanism, he cannot deviate from that mechanism. My sender simply chooses a message.

Zapechelnyuk (2023) imposes a substantive assumption that the designer's mechanisms are never fully revealing, and finds that every Pareto undominated commitment outcome can be implemented by an informed designer. In contrast, Koessler and Skreta (2023), like me, make no such assumption. Their result (Proposition 2) states that a commitment outcome is implementable in equilibrium if and only if the designer does not have profitable deviations to mechanisms that pool any subset of states in any proportion. As such, the interim commitment power actually hurts the sender: without it, the only deviations we need to check are toward fully revealing the state. Notably, the authors also find (in Propositions 3 and 4, respectively) that when the designer's value function V is quasi-convex in R's belief or when R's choice set is binary, the most profitable deviations will be toward fully informative mechanisms. In these cases, the equilibrium sets of the informed designer and the verifiable information games are essentially the same.

Comparison to Cheap Talk. Cheap talk (Crawford and Sobel, 1982, Chakraborty and Harbaugh, 2010, Lipnowski and Ravid, 2020) is a different communication protocol in which there is no verifiability requirement on which messages S can send in which states. Consequently, a cheap-talk sender faces different incentive-compatibility constraints than a verifiable-information sender, simply because he also considers deviations toward messages that are not verifiable.

As it turns out, the verifiability of his messages can 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. To see this, suppose that the state space is binary, the receiver's preferred action increases in her belief, and u_S is still increasing in R's action. Then, as the number of receiver's actions tends to infinity, all equilibria of the verifiable information game converge to full disclosure (due to the IC constraint), while all equilibria of the cheap talk game approach babbling/no disclosure. Now, if $u_S(j)$ is strictly concave, the commitment protocol approaches no disclosure, meaning that cheap talk dominates. On the other hand, if $u_S(j)$ is strictly convex, then the commitment protocol approaches full disclosure, meaning that verifiable information dominates. For further details, see the lobbying example (2) of Koessler and Skreta (2023).

When the receiver chooses between two actions, the sender always reaches the commitment payoff in the sender-preferred equilibrium of the verifiable information game. With cheap talk, he is often unable to do so.[8]

References

Alonso, Ricardo and Odilon Câmara (2016), "Persuading Voters", *American Economic Review*, 106, 11 (Nov. 2016), pp. 3590-3605. (p. 4.)

Bardhi, Arjada and Yingni Guo (2018), "Modes of Persuasion toward Unanimous Consent", *Theoretical Economics*, 13, 3, pp. 1111-1149. (p. 4.)

Ben-Porath, Elchanan, Eddie Dekel, and Barton L. Lipman (2019), "Mechanisms With Evidence: Commitment and Robustness", *Econometrica*, 87, 2, pp. 529-566. (p. 4.)

Boleslavsky, Raphael and Christopher Cotton (2015), "Grading Standards and Education Quality", *American Economic Journal: Microeconomics*, 7, 2 (May 2015), pp. 248-279. (p. 4.)

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. (pp. 4, 14.)

Crawford, Vincent P. and Joel Sobel (1982), "Strategic Information Transmission", *Econometrica*, 50, 6 (Nov. 1982), p. 1431. (p. 14.)

Gehlbach, Scott and Konstantin Sonin (2014), "Government Control of the Media", *Journal of Public Economics*, 118 (Oct. 2014), pp. 163-171. (p. 4.)

Gentzkow, Matthew and Emir Kamenica (2016), "A Rothschild-Stiglitz Approach to Bayesian Persuasion", *American Economic Review*, 106, 5 (May 2016), pp. 597-601. (p. 11.)

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.)

[8] Lipnowski (2020) notes that S reaches the commitment outcome with cheap talk if his value function is continuous in R's posterior belief. When R is choosing between two actions and S's preferences are state-independent, S's value function can be continuous only if it is constant, meaning that S provides no information under cheap talk, verifiable information, and Bayesian persuasion. I thank Elliot Lipnowski for this insight.

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. (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. (p. 4.)

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian Persuasion", *American Economic Review*, 101, 6 (Oct. 2011), pp. 2590-2615. (pp. 3, 10.)

Koessler, Frédéric and Régis Renault (2012), "When Does a Firm Disclose Product Information?", *The RAND Journal of Economics*, 43, 4, pp. 630-649. (p. 9.)

Koessler, Frédéric and Vasiliki Skreta (2023), "Informed Information Design", *Journal of Political Economy* (Mar. 28, 2023). (pp. 4, 12, 14, 15.)

Kolotilin, Anton (2015), "Experimental Design to Persuade", *Games and Economic Behavior*, 90 (Mar. 2015), pp. 215-226. (p. 4.)

Lipnowski, Elliot (2020), "Equivalence of Cheap Talk and Bayesian Persuasion in a Finite Continuous Model", *note*. (p. 15.)

Lipnowski, Elliot and Doron Ravid (2020), "Cheap Talk With Transparent Motives", *Econometrica*, 88, 4, pp. 1631-1660. (pp. 4, 14.)

Milgrom, Paul R. (1981), "Good News and Bad News: Representation Theorems and Applications", *The Bell Journal of Economics*, 12, 2, p. 380. (p. 3.)

— (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.)

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. 4.)

Romanyuk, Gleb and Alex Smolin (2019), "Cream Skimming and Information Design in Matching Markets", *American Economic Journal: Microeconomics*, 11, 2 (May 2019), pp. 250-276. (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. 4.)

Zapechelnyuk, Andriy (2023), "On the Equivalence of Information Design by Uninformed and Informed Principals", *Economic Theory*, 76, 4 (Nov. 2023), pp. 1051-1067. (pp. 4, 14.)