Derandomization of Persuasion Mechanisms

Yishu Zeng*

York University

Abstract

We consider a setting where one sender can communicate with several privately informed receivers through a persuasion mechanism before the receivers play a game. We show that for any potentially randomized persuasion mechanism, under certain conditions, there is an *effectively equivalent* deterministic persuasion mechanism, and these two mechanisms have the same set of equilibria. We exhibit the usefulness of our result in an information disclosure application, where our technique helps to derive the optimal persuasion mechanism. Overall, this paper provides a rationale for the fact that persuasion mechanisms are often deterministic in practice.

Keywords: Derandomization; Persuasion mechanism; Bayesian persuasion; Information design

JEL classification: C72, D82, D83

1 Introduction

In practice, persuasion mechanisms often have a simple and deterministic structure. Sometimes, types considered similar are assigned the same categorical identifier. For example, schools use letter grades to evaluate the performance of students, and bond rating agencies adopt coarse ratings to measure the creditworthiness of bonds. Sometimes, some types are fully revealed and others are pooled without introducing extra noise. For example, with bank stress tests, banks that pass the test are pooled, while those that fail may have their types revealed. A variety of recent papers have observed that the optimal persuasion mechanism is deterministic in several specific economic environments.[1] This paper uncovers a general underlying principle behind this phenomenon and provides tight conditions under which any persuasion mechanism has a deterministic counterpart. In fact, under these conditions, the sets of outcomes induced by deterministic and randomized persuasion mechanisms are equivalent.

*Department of Economics, York University, 4700 Keele St., Toronto, ON, M3J 1P3, Canada. E-mail: zengyish@yorku.ca.

[1]See, for example, Kolotilin [2018], Guo and Shmaya [2019], Wei and Green [2019], and Dworczak and Martini [2019] who show that the optimal persuasion mechanism takes a special deterministic structure under certain conditions. See the literature review for more discussion.

To make our statement concrete, consider a setting where a sender, who can generate information about the underlying state Ω, engages in two-way communication with multiple interacting receivers who have private types. For simplicity, we fix the message spaces for both parties: for each receiver, the set of feasible reports $\widehat{T}$ coincides with his type space; for the sender, there is a fixed finite message set $\hat{A}$ which specifies all feasible messages she can send. A *persuasion mechanism* π is a menu of information structures (an information structure is a mapping $\Omega \to \Delta(\hat{A})$) that collects receivers' reports and then selects an information structure. Therefore, π is a mapping $\widehat{T} \times \Omega \to \Delta(\hat{A})$. Say π is deterministic if, for any report profile $\hat{t}$, the selected information structure conditional on every state realization is a degenerate distribution over message space $\hat{A}$ (i.e., for any state $\omega \in \Omega$ and report $\hat{t}$, $\pi(\cdot|\hat{t}, \omega)$ is a Dirac measure on $\hat{A}$). Each receiver takes an action after receiving their message, and the payoffs of all parties depend on the receivers' actions. Our solution concept is an adapted version of the communication equilibrium in Myerson [1982] and Myerson [1986].

The main result provides conditions under which it is possible to derandomize all information structures in a persuasion mechanism in an *effectively equivalent* sense. Under those conditions, for any persuasion mechanism π, we can find a deterministic persuasion mechanism $\bar{\pi}$ such that, when holding others' strategy fixed, each receiver's highest achievable payoff is the same under $\bar{\pi}$ and π. In particular, $\bar{\pi}$ preserves the set of equilibria under π, as well as the equilibrium payoff of the sender and of each receiver type in any specific equilibrium. The key conditions for our main results are: (i) the underlying state space is atomless; (ii) receivers' actions and possible sender's messages are finite; (iii) the players' utilities are *pseudo-separable* (that is, separable in the underlying states and receivers' type profiles). These conditions are tight in the sense that, if any of them are violated, one can find a counterexample. We provide such counterexamples in Section 5. An important assumption is that the underlying state and receivers' private information are independent, which can be relaxed to allow some interdependence.

This problem contains two technical aspects: derandomizing every information structure in the given mechanism and making selections among the derandomized ones to form a new menu. Given their similar nature, it is tempting to relate this problem to the purification literature and the standard technical tool therein (the Dvoretzky–Wald–Wolfowitz theorem). In Section 3.2, we show that this approach applies when receivers have finitely many types. However, it is no longer the case when the type space is infinite as in our setting, since the foundation of the Dvoretzky–Wald–Wolfowitz theorem does not extend to such a setting mathematically.[2] To overcome this difficulty, we propose an approach that combines the insight of the measurable choice theorem of Mertens [2003] and a pseudo-separability condition on players' utility functions. To our knowledge, such an approach is new in the literature.

From a constructive angle, we demonstrate the usefulness of our result in an information disclosure application studied in the literature and add new insights with our derandomization technique. In particular, our main result allows us to focus on deterministic persuasion mechanisms in which each information structure partitions the states and discloses the partition element in which the realized state is located. Then we can transform this persuasion problem into a standard deterministic control problem. As such, our derandomization technique allows us to explicitly derive the solution to this example, which otherwise

[2]A detailed discussion can be found at the end of Section 3.2.

cannot be solved by the existing approaches in the literature.

1.1 Literature Review

This paper provides a systematic study of persuasion mechanism derandomization, allowing for multiple interacting privately informed receivers in the Bayesian persuasion and information design context (Kamenica and Gentzkow [2011] and Rayo and Segal [2010]). Especially, our set-up relates to two specific strands of this literature: the first is the strand that studies receivers with private information about their preferences, including Rayo and Segal [2010], Kolotilin et al. [2017], Kolotilin [2018], and Guo and Shmaya [2019]; the second relates to the literature that studies Bayesian persuasion problems with an uncountable underlying state space, such as Gentzkow and Kamenica [2016], Kolotilin et al. [2017], Dworczak and Martini [2019], Dworczak and Kolotilin [2022], Kleiner et al. [2021], and Arieli et al. [forthcoming]. In particular, Kolotilin [2018], Guo and Shmaya [2019], Wei and Green [2019], Dworczak and Martini [2019], Kleiner et al. [2021], and Arieli et al. [forthcoming] also have relevant results in a single receiver setting where the optimal mechanism takes a deterministic structure under certain conditions.[3] By allowing multiple privately informed receivers, our setting extends beyond their frameworks.

We can still draw interesting implications of our result in some of the existing works. Specifically, the settings in Kolotilin et al. [2017] and Wei and Green [2019] have the following common features: the distribution of underlying states is atomless, the receiver has a binary action and a linear utility, the sender's utility separates the receiver's type and the underlying states, and the receiver's private type is independent of the state. In these settings, our result implies that it is without loss of generality to work directly with deterministic mechanisms. Guo and Shmaya [2019] allow for nonseparable utility functions, but their result requires an increasing monotone likelihood ratio condition. Dworczak and Martini [2019] consider the setting with a receiver who has no private information, and whose best response depends only on the posterior mean of the state. They show that if utility function u is *regular* and affine-closed, then for any continuous and full-support prior, there exists an optimal signal that is a monotone partitional signal.[4] Our result adds new insights that deterministic persuasion mechanisms may still be optimal when the receiver's utility violates the regularity properties required therein.

In the same setting as Dworczak and Martini, Kleiner et al. [2021] and Arieli et al. [forthcoming] characterize the extreme points among the distribution of posterior means that can be induced by information structures. In particular, they show that these extreme points can be implemented by deterministic information structures. Different from their considerations, our result focuses on derandomizing any

[3] To be exact, Kolotilin [2018] and Dworczak and Martini [2019] point out that, under certain curvatures of the sender's payoff function, the optimal mechanism may be deterministic within specific regions in the posterior mean (state) space. Guo and Shmaya [2019] show that under the increasing monotone likelihood ratio condition the optimal mechanism takes the form of nested intervals. Wei and Green [2019] show that under a monotone hazard rate condition, the optimal persuasion mechanism has a deterministic cutoff structure.

[4] In Dworczak and Martini [2019], a utility function u is *regular* provided that (i) it is upper semicontinuous with at most finitely many one-sided jump discontinuities at interior points $y_1, ..., y_k \in (0,1)$ and has bounded slope (i.e., Lipschitz continuous) in each (y_i, y_{i+1}), with $y_0 = 0$ and $y_{k+1} = 1$; (ii) there exists a finite partition of $[0,1]$ into intervals such that u is either strictly convex, strictly concave, or affine on each interval in that partition.

persuasion mechanism without changing the set of equilibria, not only those that induce extreme-point distributions. Nevertheless, the aforementioned works do not intend to provide a general study of persuasion mechanism derandomization, so many focus on specific environments, such as binary action and certain monotonicity conditions. Our main result allows arbitrarily finite actions and does not rely on any monotonicity condition.

The mechanism design literature that studies mechanism equivalence is also related, especially the strand that debates whether randomization creates extra benefits for the designer, such as McAfee and McMillan [1988], Strausz [2006], Manelli and Vincent [2006], Manelli and Vincent [2007], Hart and Reny [2015], and Chen et al. [2019]. In particular, Chen et al. [2019] establish the equivalence of stochastic and deterministic mechanisms in a general environment using a mutual purification technique. However, their technique is not applicable here, since we allow nonseparability among receivers' type profiles.[5] Moreover, there is a fundamental difference between persuasion mechanism derandomization and similar topics in the mechanism design literature, as a standard mechanism can assign the allocation but a persuasion mechanism cannot.[6]

From a methodological viewpoint, our result is technically related to the purification literature that focuses on purifying mixed strategy equilibria in games of incomplete information; See, for example, Harsanyi [1973], Radner and Rosenthal [1982], Milgrom and Weber [1985], Khan and Sun [1995], Khan et al. [2006], Podczeck [2009], Khan and Zhang [2014], and He and Sun [2019]. It is known that when each player does not care about others' types (or private information) and the types are independently distributed among players, the purification problem can be quite tractable.[7] Compared to the purification methods, derandomizing persuasion mechanisms impose additional mathematical challenges. This is because any receiver can misreport his type and affect others' actions through the mechanism. Thus players' actions can be endogenously correlated under a mechanism even with independent types. As such, our derandomization problem requires the preservation of all relevant correlations between the sender's messages and the receivers' messages to avoid altering the set of equilibria.

Organization: The rest of the paper is organized as follows: Section 2 introduces our model. Section 3 presents our main result, and Section 4 applies our result to an application. Section 5 provides all counterexamples. Section 6 concludes. The Appendix collects all proofs.

[5]The separability condition in Chen et al. [2019] requires each receiver's payoff to be separable in his own type t_i and the types of the other receivers t_{-i} (see Definition 4 therein).

[6]On mechanism comparison, the literature mainly focuses on the standard mechanisms (such as auctions) that directly assign the allocation. For a few exceptions, note that the coordination mechanism in Myerson [1982] also recommends actions. Nonetheless, the notion of the coordination mechanism is still different from that of the persuasion mechanism, as the sender in a persuasion setting has extra information not available to the players. This is the new feature of the information design problem added to the problem of communication in games (see Bergemann and Morris [2019]).

[7]For instance, Milgrom and Weber [1985] consider a general Bayesian game with finite players and actions and the following key assumptions: (i) players' types are independent (conditional on a public state); (ii) each player's payoff depends only on his own type, the (finite) public state, and action profile (iii) payoffs are equicontinuous. Even with infinite types, they can use the result of Dvoretzky et al. [1950] to purify mixed-strategy equilibria. However, Dvoretzky et al. [1950] is not sufficient to solve our problem with infinite type space, as we will illustrate in Section 3.2.

2 Model

The underlying state space is a probability space $(\Omega, \mathcal{F}, \mu^\Omega)$, where $\mathcal{F}$ is a countably generated σ-algebra, and μ^Ω is an atomless probability measure.[8] Note that Ω cannot be finite.

There is a sender (Player 0) and a finite set of receivers, denoted as $\mathcal{I}$. Receivers have private information: each receiver i has a private type t_i from a Polish space T_i with $\mathcal{T}_i$ the associated Borel σ-algebra. Let $T := \prod_{i\in\mathcal{I}} T_i$ be the product space of types with $\mathcal{T} := \otimes_{i\in\mathcal{I}}\mathcal{T}_i$ the associated product σ-algebra, and let μ^T be a probability distribution over T. For expositional purposes, in the text we assume that the players have a common prior $\mu^\Omega \times \mu^T$ such that type profiles and underlying states are distributed independently (although receiver types may be correlated with each other). As noted in Remark 2 below, the proof of our main theorem is provided under a more general case that allows some interdependence between ω and t. In particular, such interdependence requires extra assumptions that are immediately satisfied when ω and t are independent.

Each receiver i chooses an action from a finite set A_i; let $A := \prod_{i\in\mathcal{I}} A_i$ be the product of action sets. Each player is a von Neumann-Morgenstern utility maximizer with a bounded measurable utility function $u_i : T \times \Omega \times A \to \mathbb{R}$ for $i \in \mathcal{I} \cup \{0\}$.

Players communicate through a *persuasion mechanism*. We fix the sender's message space to be a finite set $\hat{A} := \prod_{i\in\mathcal{I}} \hat{A}_i$. Following Kamenica and Gentzkow [2011], an *information structure* is a mapping $\Omega \to \Delta(\hat{A})$. Once committed by the sender, the persuasion mechanism asks for each receiver i's report, based on which it selects an information structure. Receivers submit reports about their types; each receiver i's report space $(\widehat{T}_i, \widehat{\mathcal{T}}_i)$ is identical to his type space.[9] A persuasion mechanism π is a $\widehat{\mathcal{T}} \otimes \mathcal{F}$-measurable mapping $\pi : \widehat{T} \times \Omega \to \Delta(\hat{A})$, where $\widehat{T} := \prod_{i\in\mathcal{I}} \widehat{T}_i$ is the product of report spaces with the associated product σ-algebra $\widehat{\mathcal{T}} := \otimes_{i\in\mathcal{I}}\widehat{\mathcal{T}}_i$. We further assume that $|\hat{A}_i| \geq |A_i|$ for all $i \in \mathcal{I}$, so information structures that recommend actions are always feasible for the sender.[10]

The game proceeds as follows: (i) the sender commits to a persuasion mechanism; (ii) each receiver submits a report based on their type realization; (iii) given the joint report, the persuasion mechanism selects an information structure; (iv) nature picks a message profile given the chosen information structure; (v) each receiver privately observes their individual message, and then takes an action.

Definition 1. We say a persuasion mechanism π is *deterministic* if for each t, ω there is a message profile $\hat{a}(t,\omega) \in \hat{A}$ such that $\pi(\cdot|t,\omega) = \delta_{\hat{a}}(\cdot)$.

In other words, a deterministic persuasion mechanism, conditional on the realized state and report profile, is a Dirac measure on $\hat{A}$. We now consider each receiver's strategy, which includes two components: a reporting strategy that files a report based on their realized type, and an action strategy that responds to the observed message based on their updated belief and knowledge. Formally:

[8] For example, the Borel σ-algebra of the n-dimensional Euclidean space $\mathbb{R}^n$ is countably generated.

[9] Even though the report space is equal to the type space, we will use different notations to clearly distinguish between true types and reported types.

[10] The persuasion mechanism is similar to the "coordination mechanism" in Myerson [1982]. The only conceptual difference between our setting and Myerson's is that there is a state ω that is outside receivers' private information. Hence, the mediator in our setting cannot elicit this information from receivers.

Definition 2. A reporting strategy of receiver i is a mapping $r_i : T_i \to \Delta(\widehat{T}_i)$ such that (i) for any t_i, $r_i(\cdot \mid t_i)$ is a probability measure on the report set $\widehat{T}_i$; (ii) for any measurable set $E_i \in \widehat{\mathcal{T}}_i$, $r_i(E_i|\cdot)$ is $\mathcal{T}_i$-measurable. An action strategy of receiver i is a measurable mapping $\sigma_i : T_i \times \widehat{T}_i \times \hat{A}_i \to \Delta(A_i)$. Let R_i and Σ_i denote the sets of reporting strategies and action strategies for receiver $i \in \mathcal{I}$, respectively.

Given a persuasion mechanism π and a strategy profile $(r, \sigma) := (r_j, \sigma_j)_{j\in\mathcal{I}}$, the expected payoff of any receiver i conditional on his type t_i and others' types t_{-i} is

$$U_i^\pi(r, \sigma|t_i, t_{-i}) := \int_\Omega \int_{\widehat{T}} \sum_{\hat{a}\in\hat{A}} \sum_{a\in A} u_i(t, \omega, a)\sigma(a \mid t, \hat{t}, \hat{a})\pi(\hat{a} \mid \hat{t}, \omega)r(\mathrm{d}\hat{t} \mid t)\mu^\Omega(\mathrm{d}\omega), \tag{1}$$

where

$$\sigma(a \mid t, \hat{t}, \hat{a}) := \prod_{j\in\mathcal{I}} \sigma_j(a_j \mid t_j, \hat{t}_j, \hat{a}_j) \text{ and } r(\mathrm{d}\hat{t} \mid t) := \prod_{j\in\mathcal{I}} r_j(\mathrm{d}\hat{t}_j \mid t_j).$$

Moreover, given his type t_i, receiver i forms a belief over the state and others' types conditional on t_i, which we denote as $\mu^T_{i|t_i} \in \Delta(T_{-i})$. Thus, the expected payoff of receiver i conditional on his type t_i is the following:

$$U_i^\pi(r, \sigma|t_i) := \int_{T_{-i}} U_i^\pi(r, \sigma|t_i, t_{-i})\mu^T_{i|t_i}(\mathrm{d}t_{-i}). \tag{2}$$

Our solution concept for the receivers' game is adapted from the solution concept of "communication equilibrium" in Myerson [1982] and Myerson [1986]. The only difference is that we require every receiver type to choose a strategy that maximizes his expected payoff, which turns out to be slightly stronger than Myerson's original definition that allows a μ^T-null set of receiver types to behave irrationally.

Definition 3 (equilibrium)**.** A strategy profile (r, σ) is an equilibrium under a persuasion mechanism π if and only if for each receiver $i \in \mathcal{I}$ and type $t_i \in T_i$,

$$U_i^\pi(r, \sigma|t_i) \geq \sup_{\hat{\sigma}_i\in\Sigma_i, \hat{r}_i\in R_i} U_i^\pi\left((\hat{r}_i, \hat{\sigma}_i), (r_j, \sigma_j)_{j\neq i}|t_i\right). \tag{3}$$

Under the given persuasion mechanism, for the action strategy σ_i to be a best response to (r_{-i}, σ_{-i}), it must be the case that, conditional on any type t_i, report $\hat{t}_i$, and message realization $\hat{a}_i$, receiver i weakly prefers the action specified in this strategy $\sigma_i(\cdot \mid t_i, \hat{t}_i, \hat{a}_i)$ to any other action. This statement is captured by inequality (3) that each receiver i chooses an action strategy that maximizes his interim utility, as discussed in Bergemann and Morris [2016]. Moreover, as this paper is meant to derandomize persuasion mechanisms in which there exists at least one such equilibrium, the question of whether an equilibrium exists is not relevant here.

One may wonder how our result relies on the choice of solution concepts; in particular, the Perfect Bayesian equilibrium (PBE) can be another appropriate alternative. Here, using PBE as the solution concept does not fundamentally change our result.[11] The advantage of our current choice is that it

[11] Our main result will be slightly different if the solution concept is PBEs instead. Specifically, for any given persuasion

can explain our key idea clearly without invoking extra complications such as belief systems. To briefly compare these two concepts: the definition of the equilibrium employed here imposes sequential rationality and (implicitly) consistency of beliefs with Bayes' rule on the equilibrium path, as well as off the equilibrium path after a receiver deviates from his equilibrium reporting strategy, whereas PBE does not require a receiver's belief to be consistent with his own deviation. Moreover, when some message arises that is supposed to be absent unless other receivers deviate, Bayes' rule does not impose any restriction on receivers' beliefs. Upon the appearance of such a message, our equilibrium concept allows receivers to play arbitrarily, while PBE requires receivers to maximize their payoff based on some arbitrarily assigned beliefs.

Before closing this section, we define the comparison of persuasion mechanisms in this environment as follows:

Definition 4. For two persuasion mechanisms, π_1 and π_2, we say π_1 is *more effective* than π_2 for receiver $i \in \mathcal{I}$ (denoted as $\pi_1 \succcurlyeq_i \pi_2$) under prior μ if there exist bijection mappings $(\psi_j : \hat{A}_j \to \hat{A}_j)_{j \neq i}$ such that, for any strategy $(r_i, \sigma_i) \in R_i \times \Sigma_i$ under persuasion mechanism π_2, there exists a strategy $(r'_i, \sigma'_i) \in R_i \times \Sigma_i$ under persuasion mechanism π_1 such that

$$U_i^{\pi_2}(r_i, r_{-i}, \sigma_i, \sigma_{-i}|t_i, t_{-i}) \leq U_i^{\pi_1}(r'_i, r_{-i}, \sigma'_i, (\sigma_j(\cdot, \cdot, \psi_j(\cdot))_{j \neq i}|t_i, t_{-i}),$$

for any realized type t_i, any others' realized type profile $t_{-i} \in T_{-i}$, and strategy profile $r_{-i} \in R_{-i}, \sigma_{-i} \in \Sigma_{-i}$. Two persuasion mechanisms π_1 and π_2 are *effectively equivalent* if $\pi_1 \succcurlyeq_i \pi_2$ and $\pi_1 \preccurlyeq_i \pi_2$ for each receiver $i \in \mathcal{I}$.

Intuitively, a persuasion mechanism π_1 is more effective than another persuasion mechanism π_2 for a receiver i if, for any reporting strategy r_i in π_2, he is able to find a strategy r'_i in π_1 such that the information structure given by r'_i under π_1 generates a weakly higher payoff (or is "more effective") than that given by r_i under π_2 for all possible decision problems he may face. Here, the class of decision problems is parametrized by his own type realization t_i, others' type realizations t_{-i}, and the strategy profile (r_{-i}, σ_{-i}). Moreover, for persuasion mechanisms, the "more effective" orders defined above should be robust to relabeling the messages; and the introduction of $(\psi_j)_{j \neq i}$ allows us to accommodate such relabelling. In particular, the interpretation of incorporating the bijection ψ into (other) receivers' action strategies is that all the other players agree to relabel messages as ψ described so they behave accordingly.

3 Derandomizing persuasion mechanisms

To begin with, let us present the intuition behind the proof of finding effectively equivalent persuasion mechanisms.

mechanism π, we can find a deterministic mechanism $\overline{\pi}$ such that any strategy profile that constitutes a PBE with some belief system under π is still a PBE with a potentially different belief system under $\overline{\pi}$ (as our derandomization does not preserve beliefs at each information set but, instead, preserves beliefs over states in aggregate), and vice versa. However, invoking this version will require us to explicitly define the belief system which incurs unnecessary complications and distracts from the main focus of this paper.

3.1 The intuitive idea

Fix an arbitrary persuasion mechanism π and any receivers' strategy profile (r, σ). Note that each receiver i's expected payoff conditional on their message $\hat{a}_i$, report $\hat{t}_i$, type profile t, and players' strategies (r, σ) takes the following explicit form:

$$\frac{\sum_{\hat{a}_{-i}} \int_{\widehat{T}_{-i}} \sum_{a \in A} \overbrace{\int_{\Omega} u_i(t, \omega, a)\pi(\hat{a} \mid \hat{t}, \omega)\mu^{\Omega}(\mathrm{d}\omega)}^{\text{Term } A} \, \sigma(a \mid t, \hat{t}, \hat{a}) \cdot r_{-i}(\mathrm{d}\hat{t}_{-i} \mid t_{-i})}{\sum_{\hat{a}_{-i} \in \hat{A}_{-i}} \int_{\widehat{T}_{-i}} \underbrace{\int_{\Omega} \pi(\hat{a}_{-i}, \hat{a}_i \mid \hat{t}_{-i}, \hat{t}_i, \omega)\mu^{\Omega}(\mathrm{d}\omega)}_{\text{Term } B} \, r_{-i}(\mathrm{d}\hat{t}_{-i} \mid t_{-i})},$$

where the denominator captures the probability that receiver i receives message $\hat{a}_i$ by reporting $\hat{t}_i$ under the above strategies and the given type profile. Suppose we can find a deterministic persuasion mechanism $\overline{\pi}$ under which, conditional on any messages $\hat{a}$, types t, and reports $\hat{t}$, the values of Terms A and B for each receiver i's payoff is the same as under π, i.e.,

$$(\text{Term } A): \int_{\Omega} u_i(t, \omega, a)\pi(\hat{a} \mid \hat{t}, \omega)\mu^{\Omega}(\mathrm{d}\omega) = \int_{\Omega} u_i(t, \omega, a)\overline{\pi}(\hat{a} \mid \hat{t}, \omega)\mu^{\Omega}(\mathrm{d}\omega) \text{ for all } a, t, \hat{t}, \hat{a} \quad (4)$$

$$(\text{Term } B): \int_{\Omega} \pi(\hat{a} \mid \hat{t}, \omega)\mu^{\Omega}(\mathrm{d}\omega) = \int_{\Omega} \overline{\pi}(\hat{a} \mid \hat{t}, \omega)\mu^{\Omega}(\mathrm{d}\omega) \text{ for all } \hat{t}, \hat{a}. \quad (5)$$

Then, we can verify that $\overline{\pi}$ and π are effectively equivalent. As a result, any equilibrium under π is still an equilibrium under $\overline{\pi}$, and vice versa.

It is worth mentioning that our construction does not rely on carefully picking off-equilibrium path beliefs to preserve all the equilibria between π and $\overline{\pi}$. Suppose that under the original persuasion mechanism π, receiver i receives a message m_i that indicates some other receivers have deviated from their equilibrium strategies. Suppose further that, as a result, receiver i holds a particular belief about other receivers' deviations which gives rise to this message with strictly positive probability. Then, equation (5) implies that, if receiver i holds the same beliefs under the derandomized mechanism $\bar{\pi}$, the message m_i will arise with strictly positive probability under $\bar{\pi}$ provided that the other receivers play the same deviation. The converse also holds, starting with deviations in $\bar{\pi}$. Thus, for any equilibrium under $\overline{\pi}$, the receiver who observes an off-equilibrium message can derive the same belief about others' deviations, as in its original equilibrium under π. In fact, for these deviations, (4) and (5) guarantee the pointwise preservation of each receiver's expected utility conditional on the realized types, reports, and message realizations for this off-equilibrium path under π and $\overline{\pi}$. Such preservation goes through regardless of whether the specified belief is sensible given the off-equilibrium event.

To further proceed along this line, the main task is to find a deterministic persuasion mechanism $\overline{\pi}$ that satisfies (4) and (5). Let us start with the case in which receivers have finite types.

3.2 Finite type space and the Dvoretzky–Wald–Wolfowitz approach

If receivers have finitely many types, then we can apply the Dvoretzky–Wald–Wolfowitz theorem to find a deterministic persuasion mechanism $\overline{\pi}$ that satisfies equations (4) and (5). For readers' convenience, we include the formal statement of the Dvoretzky–Wald–Wolfowitz (henceforth DWW) theorem below:

Theorem 1. *(Theorem 2.1, Dvoretzky et al. [1951]): Let $(Y, \mathcal{Y})$ be a measurable space in which $\{\nu_k\}$, $k = 1, ..., n$, are finite, atomless, signed measures. Suppose $h_j : Y \to \mathbb{R}$, $j = 1, ..., m$, are $\mathcal{Y}$-measurable nonnegative functions, such that $\sum_{j=1}^{m} h_j(y) \equiv 1$ for all $y \in Y$. Then, there exist $\mathcal{Y}$-measurable nonnegative functions $h_j^* : Y \to \{0, 1\}$, $j = 1, ..., m$, such that $\sum_{j=1}^{m} h_j^*(y) \equiv 1$ for all $y \in Y$ and the following holds*

$$\int_Y h_j(y)\nu_k(\mathrm{d}y) = \int_Y h_j^*(y)\nu_k(\mathrm{d}y), \tag{6}$$

for any $j = 1, ..., m$ and $k = 1, ..., n$.

With the joint actions A, messages $\hat{A}$, reports $\widehat{T}$, and types T all being finite, equations (4) and (5) follow from applying the DWW theorem separately for each $\hat{t} \in \widehat{T}$, for the measures $\{\mu_{t,a} | t \in T, a \in A\} \cup \{\mu^\Omega\}$, where $\mathrm{d}\mu_{t,a} = u(t, \cdot, a)d\mu^\Omega$, and the functions $\{\pi(\hat{a}|\hat{t}, \cdot)|\hat{a} \in \hat{A}\}$. As $\widehat{T}$ is finite, the resulting deterministic mechanism is measurable, and hence a well-defined persuasion mechanism. Thus, we have the following corollary.

Corollary 1. *Suppose that $|T_i| < \infty$ for all $i \in \mathcal{I}$. Then, for any persuasion mechanism π, there exists a deterministic persuasion mechanism $\overline{\pi}$ that satisfies equations (4) and (5) for any $\hat{a}$, t, and $\hat{t}$.*

Note that applying the DWW theorem in this way requires A and T to be finite (as the set of measures must be finite), as well as $\hat{A}$ to be finite (so that the set of functions $\{\pi(\hat{a}|\hat{t}, \cdot)|\hat{a} \in \hat{A}\}$ is finite). As $\widehat{T} = T$, $\widehat{T}$ is also finite, so that measurability is not an issue. With some extension, however, we can allow the message space $\hat{A}$ to have countably many messages.

Remark 1. Khan and Rath [2009] extend the Dvoretzky, Wald and Wolfowitz theorem (Theorem 1) to allow the cardinality of $\{h_j\}_{1 \le j \le m}$ to be countably infinite. Hence, the conclusion in Corollary 1 still holds if we relax $\hat{A}$ to have countably many elements (while keeping A, T, and $\widehat{T}$ finite).

As is well known in the literature, the DWW theorem is based on the Lyapunov convexity theorem which, however, does not extend to an infinite set of atomless measures (in other words, the statement of Theorem 1 does not hold when $n = \infty$).[12] Because of this limitation, we cannot generalize the above approach based on the DWW Theorem; more importantly, such breakdown implies that (4) and (5) cannot hold for any bounded measurable function u_i, $i \in \mathcal{I} \cup \{0\}$, and an arbitrary Polish space T with $\widehat{T} = T$ (while keeping actions and messages finite). This is because, if such a general version of (4) and (5) were true, then a more general DWW theorem that allows $n = \infty$ will be an immediate implication, which is known to be false.

[12] The Lyapunov convexity theorem states that the range of a nonatomic vector measure with values in a finite-dimensional space is compact and convex. However, the range of a countably infinite set of vector measures can fail to be convex. See Example 2, p. 261, Diestel and Uhl [1977] for such a counterexample.

3.3 Main result: Infinite type space

To overcome the limitation of the DWW approach in settings with infinite receiver types, we develop a new approach based on a selection lemma we establish that borrows the insight of Mertens [2003]. But this comes with a price: our approach requires players' preferences to satisfy a pseudo-separability condition introduced below.

Definition 5 (pseudo-separable utility)**.** A player i's utility function is *pseudo-separable* if his utility function can be written as $u_i(t,\omega,a)=\sum_{n=1}^{N} f_{i,n}(\omega,a)\cdot g_{i,n}(t,a)$, where N is a positive integer.

The above condition requires the player's utility to be the sum of finitely many additive terms; each depends on the type profile through multiplicity and can be taken outside the integrals when integrating over the state ω. Imposing such a condition on every player allows us to focus on finitely many integrands in our equations (4) and (5) of interest, and therefore bypasses the difficulty from infinite types mentioned in the finite-type case (though there will be other issues that we will soon talk about). Furthermore, if such a condition is absent, we do not expect a similar result to hold; see Example 2 for a counterexample.

Pseudo-separable utility functions are common in the literature. In fact, several well-known works, including Rayo and Segal [2010], Kolotilin et al. [2017], and Kolotilin [2018] (Assumption 2), consider a privately informed receiver with a binary-action set $\{0,1\}$ and utility $a\cdot(\omega-t)$. Thus, the receiver prefers action 1 if and only if the state is above his private type (or reservation value). There are other special forms of pseudo-separable utility functions. For another example, the sender's utility function in Bloedel and Segal [2019] is pseudo-separable, taking the form of a type-dependent weighted average $ta+(1-t)u_R(a,\omega)$ between biasing towards the receiver taking a high action and aligning with the receiver's interests.[13] A third example, adapted from Kolotilin [2018] (Assumption 2') is a receiver with utility $-(a-\omega)^2-c(a,t)$, who would like to take an action that matches the state conditional on his information while suffering a type-dependent cost.[14] More broadly, any continuous function defined on any compact set in $\mathbb{R}^n$ can be uniformly approximated by polynomials to any degree of accuracy (Stone-Weierstrass theorem), and all polynomial functions with variables t,ω,a are pseudo-separable.

Hereafter, our use of "all players" or "each player" shall indicate that the statement also applies to the sender. The following theorem shows that, if all players' payoffs satisfy such a condition, any persuasion mechanism can be derandomized in an effectively equivalent way.

Theorem 2. *Suppose that μ^Ω is atomless, the message space $\hat{A}$ is finite, and all players have pseudo-separable utility functions. Then, for any persuasion mechanism π, there exists an effectively equivalent deterministic persuasion mechanism $\overline{\pi}$.*

In particular, the set of equilibria under π and $\overline{\pi}$ coincide. For any such π and $\overline{\pi}$ under which (r^*,σ^*) is an equilibrium, the expected equilibrium payoffs of the sender and all receiver types are the same under π as they are under $\overline{\pi}$.

[13] For the linear model in Bloedel and Segal [2019], the underlying state is $[\underline{\omega},\overline{\omega}]$ with $\underline{\omega}<0<\overline{\omega}$. The receiver has binary actions $\{0,1\}$ with the utility function $a\omega$. Hence, the receiver prefers action 1 if and only if the underlying state $\omega>0$.

[14] In the original equivalent model in Kolotilin [2018], the receiver is uninformed and, by taking action a, gets a utility of $-(a-\omega)^2$. Thus, the receiver would like to match the state as much as possible given his information.

As mentioned, the basic outline of the proof is to construct a deterministic persuasion mechanism $\overline{\pi}$ based on an arbitrarily given mechanism π that satisfies (4) and (5) for all reports, actions, type realizations, and message realizations. Observe from this pseudo-separability condition that it is sufficient to show the following to establish (4): for any $i \in \mathcal{I} \cup \{0\}$ and $n = 1, ..., N$ with N being the maximum index for each player's utility (this is w.l.o.g., as some additive terms may be zero),

$$\int_{\Omega} f_{i,n}(\omega, a)\pi(\hat{a} \mid \hat{t}, \omega)\mu^{\Omega}(\mathrm{d}\omega) = \int_{\Omega} f_{i,n}(\omega, a)\overline{\pi}(\hat{a} \mid \hat{t}, \omega)\mu^{\Omega}(\mathrm{d}\omega). \tag{7}$$

From the above, one can infer that for any given $\hat{t}$, there are only finitely many measures under which the two information structures $\pi(\cdot \mid \hat{t}, \cdot)$ and $\overline{\pi}(\cdot \mid \hat{t}, \cdot)$ must yield the same integrals. As such, we avoid the known issue that the DWW theorem generally fails when such to-be-purified measures are infinitely many. Thus, for any fixed $\hat{t}$, one can apply the DWW theorem to find a deterministic information structure such that (4) and (5) are satisfied for this particular $\hat{t}$. However, if T, and hence $\widehat{T}$, is uncountable, a new issue may emerge, that is, a deterministic mechanism resulting from "arbitrary" pointwise applications of the DWW theorem may still fail to be measurable (in which case it is not a persuasion mechanism). We resolve this issue as follows.

Our construction relates to the idea in Gentzkow and Kamenica [2017] which propose representing a potentially randomized information structure as a deterministic function defined on the product space $\Omega \times X$ that enlarges the states Ω with the range X of the independent randomization device chosen in the model. Specifically, our construction involves the following two steps: Firstly, we will rewrite persuasion mechanisms into deterministic functions on an enlarged space $\widehat{T} \times \Omega \times X \to \{\delta_{\hat{a}} \mid \hat{a} \in \hat{A}\}$ that selects messages based on the reports, state, and the third parameter x which represents the realized value of the randomization device. Secondly, we will collapse the above function into a deterministic mechanism defined only on $\widehat{T} \times \Omega$ by only selecting elements in X on a report-state basis. In particular, the collapse is done in a way that the resulting mechanism satisfies (5) and (7). Note that we will only be explicit about the range X of randomization devices but not their structures, since the collapse would make those structures irrelevant eventually.

For the first step, it is quite natural to choose the space X in a "direct" way such that each realization x of the randomization device directly reveals the relevant values from integrating the given mechanism in (5) and (7). To formalize this idea, we introduce a mapping H that maps any information structure $\nu : \Omega \to \Delta(\hat{A}) \subseteq \mathbb{R}_+^{|\hat{A}|}$ into the following vector:

$$H(\nu) := \left(\int_{\Omega} \nu(\omega) \,\mathrm{d}\mu^{\Omega}, \int_{\Omega} f_{0,1}(\omega, a)\nu(\omega) \,\mathrm{d}\mu^{\Omega}, ..., \int_{\Omega} f_{|\mathcal{I}|,N}(\omega, a)\nu(\omega) \,\mathrm{d}\mu^{\Omega} \right). \tag{8}$$

Given the mechanism π and for any fixed $\hat{t}$, $H(\pi(\cdot|\hat{t}, \cdot))$ collects all the left-hand side expressions of equation (5) and (7) when evaluated at the corresponding information structure $\pi(\cdot|\hat{t}, \cdot)$. To directly capture the relevant integration values in (5) and (7), we want to choose X and find a deterministic function $h : \widehat{T} \times \Omega \times X \to \{\delta_{\hat{a}} \mid \hat{a} \in \hat{A}\}$ such that the following is satisfied for any $x \in X$ and $\hat{t} \in \widehat{T}$:

$$x = \underbrace{\left(\int_{\Omega} h(\hat{t}, \omega, x) \, \mathrm{d}\mu^{\Omega}, \int_{\Omega} f_{0,1}(\omega, a) h(\hat{t}, \omega, x) \, \mathrm{d}\mu^{\Omega}, ..., \int_{\Omega} f_{|\mathcal{I}|,N}(\omega, a) h(\hat{t}, \omega, x) \, \mathrm{d}\mu^{\Omega} \right)}_{H(h(\hat{t},\cdot,x))}. \tag{9}$$

As such, the above (9) implies that X must include all possible values of applying H to deterministic persuasion mechanisms. Thus, we define X to be the union of the value of H evaluated at any deterministic persuasion mechanism across all possible report profiles $\hat{t}$, that is,

$$X := \bigcup_{\hat{t} \in \widehat{T}} \Big\{ H(k(\hat{t}, \omega)) \Big| k : \widehat{T} \times \Omega \to \{\delta_{\hat{a}} \mid \hat{a} \in \hat{A}\}; \text{ that is, } k \text{ is a deterministic persuasion mechanism} \Big\}. \tag{10}$$

Given such a choice of X, Lemma 3 in the Appendix, borrowing insight from Mertens [2003], shows that there exists a measurable function $h : \widehat{T} \times \Omega \times X \to \{\delta_{\hat{a}} \mid \hat{a} \in \hat{A}\}$ that satisfies (9).[15] This function h, when combined with a randomization device that takes values from X, will give rise to a persuasion mechanism, though it is not necessarily deterministic when projected onto $\widehat{T} \times \Omega$.

We now turn to the second step which, based on the above h and X, focuses on selecting an appropriate x for each report and state so that (5) and (7) are satisfied. From our construction, observe that choosing $x = H(\pi(\cdot|\hat{t}, \omega))$ under h for each report profile and state helps to establish (5) and (7); in fact, such a choice of x depends only on the report $\hat{t}$. To show that such a selection is always feasible, we rely on the fact that the state distribution μ^{Ω} is atomless, which allows us to utilize another insight from Mertens [2003] to show that any such $H(\pi(\cdot|\hat{t}, \omega))$ can always be attained by some deterministic mechanism, and thus is within the domain X for each report profile.

Combining these two steps, we can now use the selection for each $(\hat{t}, \omega)$ in the second step, $H(\pi(\cdot|\hat{t}, \omega))$, as the third argument to the function $h(\hat{t}, \omega, \cdot)$ in the first step. The result is a deterministic persuasion mechanism that satisfies equations (5) and (7). The following equation, derived from equation (9), summarizes this procedure:

$$\underbrace{H(\pi(\cdot|\hat{t}, \omega'))}_{(\int \pi(\hat{t},\omega') \, \mathrm{d}\mu^{\Omega}, ..., \int f_{i,n}(\omega',a)\pi(\hat{t},\omega') \, \mathrm{d}\mu^{\Omega}, ...)} \overset{(9)}{=} \Big(\int_{\Omega} \overbrace{h(\hat{t}, \omega', H(\pi(\cdot|\hat{t}, \omega')))}^{\pi(\hat{t},\omega')} \, \mathrm{d}\mu^{\Omega}, ... \; ..., \int_{\Omega} f_{i,n}(\omega', a) \underbrace{h(\hat{t}, \omega', H(\pi(\cdot|\hat{t}, \omega')))}_{\pi(\hat{t},\omega')} \, \mathrm{d}\mu^{\Omega}, ... \Big).$$

[15] As far as we know, the application of Mertens' result to information design, in combination with the separation condition in the players' utility, is new in the literature. The major application of Mertens' result in the previous literature is to study the existence of different classes of equilibria (e.g., stationary correlated equilibria and (stationary almost) Markov perfect equilibria) in stochastic games. Also, this result has been applied in the dynamic mechanism design literature to establish the measurability of the transfers (see Liu [2018]). Given that these problems have a completely different nature from ours, none of them require the pseudo-separability condition for their results to go through.

The proof of Theorem 2′ in the Appendix describes the technical details of the construction.

Remark 2. We actually prove the theorem for the more general situation where players have subjective priors absolutely continuous with respect to some underlying measure $\mu^{\Omega} \times \mu^{T}$, and some interdependence between the underlying state ω and the type profile t is allowed.[16] The set-up in the main text is therefore a special case. In the general case, the proof requires some additional conditions (including separability of the corresponding Radon–Nikodym densities) that are satisfied given the assumptions in the main text. We place the more general proof of Theorem 2 in the Appendix A.1.

Based on the above insight, we may provide a simplification for the definition of *signals with rich structures* proposed by Gentzkow and Kamenica [2017]. Under their definition, a signal (or an information structure in our setting) is a finite partition of $\Omega \times [0, 1]$, where every partition element is a signal realization.[17] Thus, when the players share an atomless prior, Theorem 2 allows us to simplify this definition such that a signal with rich structure is a finite partition on Ω with each element a signal realization. For example, suppose that the underlying state space is $\Omega = [0, 1]$ and the common prior is the Lebesgue measure. Then one can define a signal as a finite partition on $\Omega = [0, 1]$ instead of a finite partition on the product set $\Omega \times [0, 1]$.

4 Application: Discriminatory disclosure problem

To explain our results constructively, we apply them to investigate a discriminatory information disclosure application. Our result allows us to focus on deterministic persuasion mechanisms. Such a focus allows us to transform the persuasion problem into a deterministic control problem, which can make the problem quite tractable. As far as we know, this example has not been covered by the literature.

Consider the following persuasion game: There is an underlying state, a random variable uniformly distributed in $[0, 1]$. The game has two players: a sender and a single receiver. The receiver has a private type t in $[0, 1]$ distributed according to a distribution with the density function $g(t) = 20t^3(1-t)$. The state and receiver's private type are distributed independently. The receiver chooses binary action $a \in \{0, 1\}$ and his utility function is $(\omega - t) \cdot a$. The sender's utility function is $\omega \cdot a$. For persuasion mechanisms, we fix the sender's messages $\hat{A}$ to be action recommendations. Whenever indifferent, assume that the receiver always takes action 1. Given that the message space is binary, write a persuasion mechanism $\pi : \widehat{T} \times \Omega \to [0, 1]$ a mapping from the state and the receiver's report to the probability of recommending the receiver $a = 1$.

There are two closely related works Kolotilin et al. [2017] and Guo and Shmaya [2019] exploring similar environments. Their assumptions are violated in this example. Specifically, the sender's utility

[16]Nevertheless, we cannot allow full generality in the interdependence between ω and t. Later, we provide two counterexamples (Examples 3 and 4) showing that such a derandomization result may fail if ω and t are correlated.

[17]Though Ω is a finite state space in their original definition; it is natural to extend it slightly to allow a general underlying state space $(\Omega, \mathcal{F}, \mu^{\Omega})$ as in our setting.

function in our example does not satisfy the specific form in Kolotilin et al. [2017].[18] Guo and Shmaya [2019] provide a set of assumptions under which the optimal disclosure mechanism takes a nested interval form. Our example violates their Assumption 3 which requires the ratio of any two receiver types' acceptance utility to be monotonic in states.[19] Relatedly, Dworczak and Martini [2019] develop a useful technique applicable to derive optimal (potentially stochastic) information structure when the sender cannot solicit reports from the receiver in this example. But it is unclear whether this restriction is without loss of generality since persuasion mechanisms allow the sender to solicit reports from the receiver.

We now introduce a new approach based on our main result that allows us to explicitly compute the optimal persuasion mechanism. As in Kolotilin et al. [2017], say a mechanism π implements a function pair $(q(\cdot), p(\cdot))$ if $q(\hat{t}) = \int_0^1 \pi(\omega, \hat{t})\,\mathrm{d}\omega$ and $p(\hat{t}) = \int_0^1 \omega\pi(\omega, \hat{t})\,\mathrm{d}\omega$ for every $\hat{t} \in [0, 1]$. To interpret, given the report $\hat{t}$, $q(\hat{t})$ is the overall probability of $\pi(\cdot, \hat{t})$ recommending $a = 1$, so $q(\hat{t})t$ measures the cost of the receiver type t from obeying the recommendation. $p(\hat{t})$ captures the receiver's type-independent revenue from obeying $\pi(\cdot, \hat{t})$ as well as the sender's payoff. For exposition elegance, we suppress the dependence of such function pairs on their underlying mechanism throughout this section.

Our approach would transform this persuasion problem into a standard deterministic control problem with the control variables $(q(\cdot), p(\cdot))$. To proceed, we first need to answer when the pair $(q(\cdot), p(\cdot))$ is implementable by a persuasion mechanism. Note that, as players' payoffs are pseudo-separable and the message is binary, our main result is applicable. By Theorem 2, it is without loss of generality to focus on deterministic persuasion mechanisms in this setting. This simplification comes in handy in resolving the implementation question. Recall that a deterministic persuasion mechanism is a menu of deterministic information structures. In this setting, any deterministic information structure can be thought of as partitioning the state space $[0, 1]$ into two subsets and disclosing which subset the realized state is in. This means, the pair (q, p) is implementable by a persuasion mechanism if and only if, for each report $\hat{t}$, one can find a subset $X(\hat{t}) \subseteq [0, 1]$ that satisfies $q(\hat{t}) = \int_{X(\hat{t})} \mathrm{d}F$ and $p(\hat{t}) = \int_{X(\hat{t})} \omega\,\mathrm{d}F$ simultaneously. The following lemma completely characterizes when a function pair is implementable by deterministic persuasion mechanisms.

Lemma 1. *The pair (q, p) is implementable by a deterministic persuasion mechanism if and only if the following two inequalities are satisfied for any $\hat{t} \in [0, 1]$:*

$$\frac{q(\hat{t})^2}{2} \leq p(\hat{t}) \leq q(\hat{t})(1 - \frac{q(\hat{t})}{2}) \text{ and } 0 \leq q(\hat{t}) \leq 1. \tag{11}$$

Note that our approach is constructive. The proof of Lemma 1 in Appendix A.2 provides a way to construct a deterministic persuasion mechanism that implements a qualified pair. Suppose we already know a pair (q, p) is implementable by a deterministic persuasion mechanism π, the ensuing question is under what additional conditions is the underlying mechanism π incentive compatible (that is, the receiver

[18] That is, the sender's utility $\omega \cdot a$ cannot be written in the form $a + f(t) \cdot (\omega - t)a$, where $f(t)$ is a bounded measurable function.

[19] In our context, their Assumption 3 requires that for any t, t' with $t' < t$, $\frac{g(t)(\omega-t)}{g(t')(\omega-t')}$ weakly increases in ω. However, for any such t, t' with $t' < t$, $\frac{g(t)(\omega-t)}{g(t')(\omega-t')}$ is non-montone aound $\omega = t'$.

is willing to report their true type and follow the recommendations)? Since the receiver's utility in our setting is the same as Kolotilin et al. [2017], we can borrow a result therein that answers this question.

Lemma 2 (Lemma 1, Kolotilin et al. [2017])**.** *Suppose that a persuasion mechanism π implements (p, q). Let $U(t) := \int_0^1 (\omega - t)\pi(\omega, t)\, \mathrm{d}\omega$. Then π is incentive compatible if and only if the following three conditions hold simultaneously: (i) $q(\cdot)$ is nonincreasing; (ii) $U(t) = \int_t^1 q(s)\, \mathrm{d}s$; (iii) $U(0) = E[\omega]$.*

Note that $U(t)$ in the above lemma is the expected utility of type t when they truthfully report their type and follow the recommendation under π. If π implements (q, p), by definition, $U(t) = p(t) - q(t)t$. We can rewrite the above (ii) to its respective differential format: $U'(t) = -q(t)$ almost surely with $U(1) = 0$. Note that (iii) and (11) imply $q(0) = 1$. For convenience, we will use the control variables (q, U) instead. In summary, combining the results in Lemma 1 and 2, we can transform the persuasion problem into the following deterministic control problem:

$$\begin{aligned} \max_{q,U} \int_0^1 p(t)g(t)\, \mathrm{d}t &= \int_0^1 (U(t) + q(t)t)g(t)\, \mathrm{d}t \\ \text{subject to} \quad & q(\cdot) \text{ is nonincreasing with } q(0) = 1 \text{ and } q(1) = 0 \\ & U'(t) = -q(t), U(0) = 0.5, U(1) = 0 \\ & \frac{q(t)^2}{2} \leq U(t) + q(t)t \leq q(t)(1 - \frac{q(t)}{2}) \\ & 0 \leq q(t) \leq 1. \end{aligned} \tag{12}$$

To derive the optimal persuasion mechanism, one can first solve the above control problem to obtain the optimal function pair (p, q). Then the proof of Lemma 1 provides a constructive way to implement the optimal triple with a deterministic persuasion mechanism.

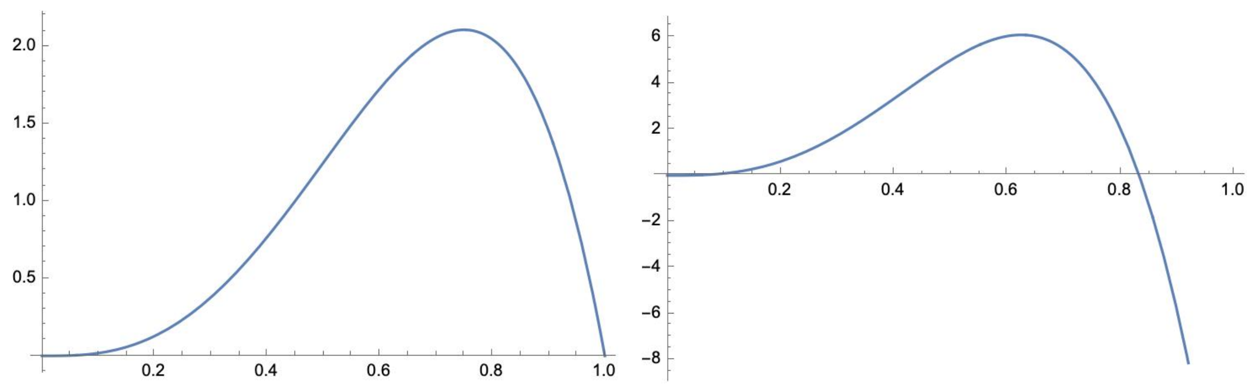
Consider two benchmark cases. First, under no revelation, because the expected state is 0.5, receiver type $t > 0.5$ takes action 0 and the sender's payoff is 0.09375. Second, under full revelation, receivers with type $t \leq \omega$ take action 1. The sender's expected payoff is then 0.261905.

We now solve this persuasion problem. For any pair (q, U) that satisfies $U'(t) = -q(t)$, it follows by an integration-by-parts argument that the sender's payoff function is

$$\int_0^1 (U(t) + q(t)t)g(t)\, \mathrm{d}t = \int_0^1 U(t)(2g(t) + g'(t)t)\, \mathrm{d}t. \tag{13}$$

We plot the density function and the induced weight function in the sender's payoff in Figure 1. From the plot, we can see that maximizing the value of $U(t)$ is optimal for the sender whenever $(2g(t) + g'(t)t) \geq 0$, while minimizing the value of $U(t)$ is optimal whenever $(2g(t) + g'(t)t) \leq 0$. We derive the solution in the appendix. From our analysis, the solution $(q^*(t), U^*(t))$ to (12) is the following:

$$q^*(t) = \begin{cases} 1 - t & 0 \leq t \leq 0.75; \\ 0.25 & 0.75 \leq t \leq 0.87; \\ 0 & t \geq 0.87; \end{cases} \quad U^*(t) = \begin{cases} \frac{(1-t)^2}{2} & 0 \leq t \leq 0.75; \\ 0.22 - 0.25t & 0.75 \leq t \leq 0.87; \\ 0 & t \geq 0.87. \end{cases}$$



(a) the density function $g(t)$ (b) the weight function $2g(t) + g'(t)t$

Figure 1: The plots of $g(x)$ and the induced weight function $2g(t) + g'(t)t$

We further plot the graphs of the solution q^* and U^* in the following figure:

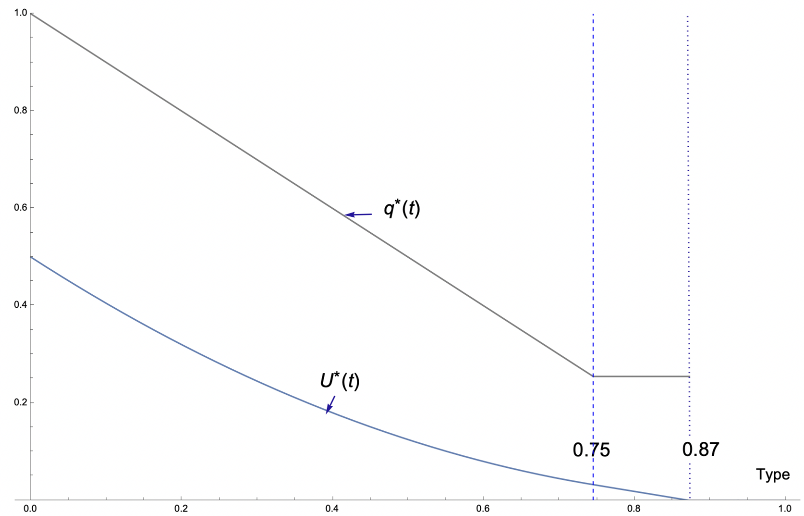


Figure 2: This figure presents the plot of the solution (q^*, U^*). The "regime-change" types 0.75 and 0.87 are marked with the two perpendicular dashed lines accordingly.

Using the insight of the proof of Lemma 1, the resulting optimal pair (q^*, p^*) with $p^*(t) = U^*(t) + q^*(t)t$ can be implemented by the following deterministic persuasion mechanism:

$$\pi^*(\omega, t) = \begin{cases} 1 & \text{when } 0 \leq t \leq 0.75, \omega \geq t; \\ 1 & \text{when } 0.75 \leq t \leq 0.87, \omega \geq 0.75; \\ 0 & \text{otherwise.} \end{cases}$$

We plot the optimal mechanism in the following figure:

The above optimal mechanism π^* implies that for any type $t \leq 0.75$, the optimal mechanism π^* recommends $a = 1$ if and only if $\omega \geq t$; for $0.75 \leq t \leq 0.87$, then π^* recommends $a = 1$ if and only if $\omega \geq 0.75$; and the mechanism always recommends $a = 0$ for any type $t > 0.87$. Under such π^*, each receiver type is willing to report his true type and follow the recommendation. Note that, although our example violates the assumption in Guo and Shmaya [2019], the receiver type's t's acceptance set (i.e.,

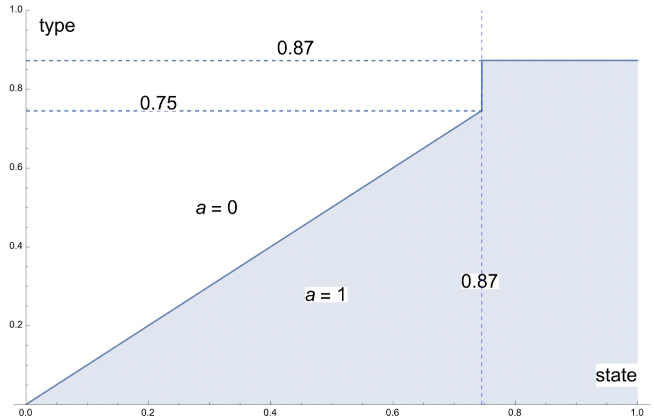


Figure 3: The optimal persuasion mechanism.

$\{\omega \mid \pi^*(\omega, t) = 1\}$) in the optimal persuasion mechanism shrinks in a nested manner as t increases. In other words, the optimal persuasion mechanism still takes the nested-interval form. The sender's maximum payoff under π^* is 0.2644. Appendix A.2 collects all the derivation details.

5 Counterexamples

Our results rely on three crucial conditions: (i) the measure μ^Ω on the underlying state space is atomless; (ii) the receivers' actions and the sender's messages are finite; and (iii) the players' utilities are pseudo-separable. If any are violated, we can find a counterexample. For condition (i), the famous judge-prosecutor example in Kamenica and Gentzkow [2011] is an immediate counterexample, where all conditions are satisfied except that the state space is atomic. In that example, an optimal information structure cannot be derandomized.[20]

The following example shows that our result may break down when condition (ii) is violated. In the example, the message space is uncountably infinite, and randomization can substantially enrich the mechanism structure in a way that deterministic mechanisms cannot replicate.

Example 1 (uncountably many messages)**.** Let $\Omega = [-\frac{1}{3}, \frac{4}{3}]$ with the distribution

$$\mu^\Omega(\mathrm{d}\omega) := \begin{cases} 0.5\,\mathrm{d}\omega & \text{on } [-\frac{1}{3}, \frac{1}{3}) \cup (\frac{2}{3}, \frac{4}{3}], \\ 1\,\mathrm{d}\omega & \text{on } [\frac{1}{3}, \frac{2}{3}]. \end{cases}$$

There is one sender and two receivers. Both receivers have null private types. As $\widehat{T} = T$, $\widehat{T}$ is also a null set. Receiver 1 takes an action from $\{0, 1\}$ while receiver 2's action set is a singleton. We assume that receiver 1's utility function is constant. If receiver 1 takes action 1, then both receiver 2 and the sender obtain a utility of 1; otherwise, they obtain 0. Consider the following information structure with

[20]Note that the probability measure in this example is purely atomic. It is known that any probability measure can be decomposed into a purely atomic part and a nonatomic part (see, for instance, Johnson [1970]). Thus, if a probability measure is not purely atomic, it is always possible to apply our result locally to its nonatomic part with some scaling.

the message space $[0,1]$:

$$\pi(\omega)=\begin{cases}\delta_{\{\omega+\frac{1}{3}\}} & \omega\in[-\frac{1}{3},\frac{1}{3}),\\ \delta_{\{\omega-\frac{1}{3}\}} & \omega\in(\frac{2}{3},\frac{4}{3}],\\ 0.5\delta_{\{\omega+\frac{1}{3}\}}+0.5\delta_{\{\omega-\frac{1}{3}\}} & \omega\in[\frac{1}{3},\frac{2}{3}].\end{cases} \tag{14}$$

Observe that the distribution of messages under π (denoted as ν) is the uniform distribution over $[0,1]$. Given that receiver 1 has a constant utility, any strategy is optimal. For any deterministic mechanism effectively equivalent to π, our definition of effective equivalence requires that, given receiver 1's strategy, receiver 2 must receive the same payoff in π and the deterministic mechanism subject to a message relabelling. For example, suppose receiver 1's strategy under π is to take action 1 if and only if the message is within the interval $[0,0.5]$. Consider a deterministic mechanism that communicates no information to receiver 1. If the message in the latter is in $[0,0.5]$, receiver 2's payoff increases from 0.5 to 1; if the message is anything else, their payoff decreases to 0.

As such, we claim that no deterministic information structure can be effectively equivalent to π. Suppose f is a deterministic information structure effectively equivalent to π. We will show that this implies that the distribution of messages under f subject to such relabelling must be equal to ν. However, we argue that such equivalence cannot hold, and thus such an f does not exits in the first place. The proof operates by assuming the existence of such a deterministic information structure, then discretizing the state space and using a convergence argument to derive a contradiction. See Appendix A.3 for more details.

Condition (iii) is about pseudo-separability. Let us first explain intuitively why our result may break down without such a condition in the setting of a single receiver with private types. Fix any arbitrary persuasion mechanism π. For any receiver type, to derandomize π effectively equivalent given their specific utility requires certain conditions to be satisfied. If the receiver's utility satisfies pseudo-separability, then the type will affect the receiver's utility in a multiplicative way and each receiver type essentially requires the same set of conditions. If, on the contrary, the pseudo-separability is not satisfied, then the conditions required by different receiver types may contradict each other. Such potential contradiction makes it not always possible to derandomize π in an effectively equivalent way. We formalize this idea in the following example.

Example 2. There is only one receiver who has a private type. The underlying state space and the receiver's private type space are unit intervals, i.e., $\Omega=T=[0,1]$. It is commonly known that the random variables ω and t are independent and both are uniformly distributed. We allow the sender to have any bounded measurable utility function. The receiver has two actions $\{0,1\}$ and utility function $u(\omega,t,a)$, where

$$u(\omega,t,a):=\begin{cases}\mathbb{1}_{\{(\omega',t')|\omega'\geq t'\}}(\omega,t) & \text{if he takes action } a=1,\\ \mathbb{1}_{\{(\omega',t')|\omega'<t'\}}(\omega,t) & \text{if he takes action } a=0.\end{cases}$$

For any type t, if he ignores the information he may receive and always takes action 0, his expected payoff is t, and if he always takes action 1, then his payoff is $1 - t$. Consider a information structure π with the sender's message set $\{0, 1\}$ such that $\pi(0|\omega) = 0.8$ and $\pi(1|\omega) = 0.2$ if $\omega < 0.5$, and $\pi(1|\omega) = 0.8$ and $\pi(0|\omega) = 0.2$ if $\omega \geq 0.5$. We claim there does not exist a deterministic persuasion mechanism that preserves all the equilibria under π as well as the corresponding equilibrium payoffs for each player. See Appendix A.3 for the proof details.

In our basic setting, the state and types are distributed independently. It is worth pointing out that, though the independence between the underlying state and receivers' type profile can be weakened (as in Theorem $2'$ in the Appendix), we cannot relax this assumption completely.

There are two reasons why such independence is required. The first reason is aligned with the idea in Aumann [1987] that randomization can provide coordination among players in the multi-player environment. As in Example 3, when all players' information is perfectly correlated, then deterministic persuasion mechanisms cannot be equivalent to random mechanisms.

Example 3. There is a sender (the social planner), and two receivers, each with an action set $A_i = \{L, R\}$, $i \in \{1, 2\}$. The underlying state space is $\Omega = [0, 1]$. All three players have perfect information, and thus the underlying state and receivers' information are perfectly correlated. The sender's payoff is the sum of all receivers' payoffs, i.e., $u_s(\omega, a_1, a_2) = u_1(\omega, a_1, a_2) + u_2(\omega, a_1, a_2)$. The receivers' payoff matrix is described below. Whenever receivers are indifferent, they take the sender-preferred actions.

$\omega, (u_1, u_2)$	L	R
L	(6ω,6ω)	(1ω,8ω)
R	(8ω,1ω)	(0,0)

In this example, the optimal information structure is the following: conditional on any realized ω, with probability 0.2, the sender privately recommends both receivers to play L; with probability 0.4, the sender privately recommends the first receiver to play L and the second receiver to play R; and with probability 0.4, the sender privately recommends the first receiver to play R and the second receiver to play L. Such an information structure would induce receivers to play a correlated equilibrium at each realized state. But any deterministic information structure can only induce Nash equilibria at each realized state in this game. Thus, the optimal information structure achieves a strictly higher payoff than any deterministic one.

The second reason relates to the previous condition (iii). To explain the intuition, consider the more general case when the joint distribution of state and types admits a density function $f(\omega, t)$ with respect to a fixed probability measure $\mu^\Omega \times \mu^T$. Following Section 5.1 of Guo and Shmaya [2019], we can view $u_i(t, \omega, a) \cdot f(\omega, t)$ as a generalized utility function for each player i where the joint distribution of state and type is replaced by $\mu^\Omega \times \mu^T$. When there is a player with constant utility 1, condition (iii) implies that if his generalized utility $f(\omega, t)$ violates pseudo-separability, then our result may not hold. Note that the independence between ω and t is a special case of $f(\omega, t)$ being pseudo-separable.

To provide an explicit example, we modify Example 2 and show that, if there does not exist some separability between states and types, then our main result may fail, even in a single receiver setting.

Example 4. There is one sender, and one receiver who has a private type. The underlying state space and the receiver's private type space are unit intervals, i.e., $\Omega = T = [0,1]$. It is commonly known that the joint distribution of random variables ω and t is $f(\omega,t)\,\mathrm{d}\omega\,\mathrm{d}t$ where $f(\omega,t) = \frac{2}{3}\cdot(2\mathbb{1}_{\{(\omega',t')|\omega'\geq t'\}}(\omega,t) + \mathbb{1}_{\{(\omega',t')|\omega'<t'\}}(\omega,t))$. We allow the sender to have any bounded measurable utility function. The receiver has two actions $\{0,1\}$ and utility function $u(\omega,t,a)$, where $u(\omega,t,a) := a\cdot(\omega - t)$.

For any type t, if he ignores the information he may receive and always takes action 0, then his expected payoff is 0, and if he always takes action 1, then his payoff is $\frac{2+t^2-4t}{4-2t}$. Consider an information structure π with the sender's message set $\{0,1\}$ that $\pi(0|\omega) = 0.8$ and $\pi(1|\omega) := 0.2$ if $\omega < 0.5$, and $\pi(1|\omega) = 0.8$ and $\pi(0|\omega) = 0.2$ if $\omega \geq 0.5$. We claim that there does not exist a deterministic information structure that is effectively equivalent to π. The proof is shown in Appendix A.3.

6 Conclusion

It is quite natural for a practical designer to restrict attention to deterministic persuasion mechanisms, given that they are relatively simple to design and implement in practice. However, such restriction may lose generality, thereby incurring suboptimality.

This paper proposes a criterion to compare different persuasion mechanisms based on their (informational) effectiveness. Based on this criterion, we provide tight conditions under which a potentially stochastic persuasion mechanism is effectively equivalent to some deterministic persuasion mechanism. Our results enhance the understanding of when it is without loss of generality to restrict attention to deterministic persuasion mechanisms, which may provide a simplified solution to the sender's optimization problem in practice. Potentially interesting future directions of research could be whether our derandomization technique can be useful in other economic situations with incomplete information, such as auction design, where the designer often has private information that may directly affect the allocation outcome.

Acknowledgements: I have benefited greatly from the thoughtful and constructive feedback of the editor, the associate editor, and three referees. I thank David Miller and Uday Rajan for their continued support. I would also like to thank Tilman Börgers, Marina Halac, Fei Hu, Ming Li, Heng Liu, and participants at the Michigan theory lunch seminar for helpful comments and suggestions. The first draft was written while the author was at the University of Michigan.

A Appendix

A.1 Proof of Theorem 2

Before the main proof, we first introduce a technical lemma borrowing the insights from Mertens [2003].[21] For any Hausdorff topological space X, let $K^*(X)$ be the set of all nonempty compact subsets of X. Say a measurable correspondence $\mathbf{N}$ from the product space $E \times \Omega$ to $\mathbb{R}^l$ is $\mathbf{P}$-integrable if, for any measurable selection $\mathbf{f}$ of $\mathbf{N}$ (denoted as $f \in \mathbf{N}$), $\mathbf{f}(\omega, e)$ is $\mathbf{P}(\mathrm{d}\omega|e)$-integrable for any $e \in E$.

Lemma 3. *Suppose that $(\Omega, \mathcal{A})$ and $(E, \mathscr{E})$ are measurable spaces with the σ-algebra $\mathcal{A}$ countably generated; that $\mathbf{P}(\mathrm{d}\omega|e)$ is a $\mathbb{R}^k$-valued bounded kernel; and that $\mathbf{N}$ is a $\mathbf{P}$-integrable measurable correspondence from $(E \times \Omega, \mathscr{E} \otimes \mathcal{A})$ to $K^*(\mathbb{R}^l)$.[22] Let $\int \mathbf{N}\,\mathrm{d}\mathbf{P}$ be the map from E to the subsets of $\mathbb{R}^{l\cdot k}$ that $\int \mathbf{N}\,\mathrm{d}\mathbf{P}(e) := \left\{ \int \mathbf{f}(\omega, e)\mathbf{P}(\mathrm{d}\omega|e) \middle| \mathbf{f} \in \mathbf{N} \right\}$.*

(1) Fix an arbitrary element $e \in E$. Let $\mathbf{f}$ be a measurable selection of the convex hull of $\mathbf{N}(e, \cdot)$ and $\mathbf{f} \in \mathbf{N}(e, \cdot)$ on each atom of $\mathbf{P}(\cdot|e)$. Then, for any $\mathbb{R}^l$-valued bounded measurable function $\mathbf{u}$, we have $\int \mathbf{u} \cdot \mathbf{f}\, \mathbf{P}(\mathrm{d}\omega|e) \in \left(\int \mathbf{u} \cdot \mathbf{N}\,\mathrm{d}\mathbf{P}\right)(e)$.

(2) Fix an arbitrary $\mathbb{R}^{l\cdot M}$-valued function $\mathbf{g}^M := (\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_M)$ such that each $\mathbf{g}_m$, $1 \le m \le M$, is a $\mathbb{R}^l$-valued $\mathscr{E} \otimes \mathcal{A}$-measurable bounded function. Let $\int \mathbf{g}^M \odot \mathbf{N}\,\mathrm{d}\mathbf{P}$ be a mapping from E to the subsets of $\mathbb{R}^{M\cdot k}$ that

$$\int \mathbf{g}^M \odot \mathbf{N}\,\mathrm{d}\mathbf{P} : e \rightarrow \left\{ \int (\mathbf{g}_1 \cdot \mathbf{f}'(\omega, e), ..., \mathbf{g}_M \cdot \mathbf{f}'(\omega, e))\mathbf{P}(\mathrm{d}\omega|e) \middle| \mathbf{f}' \in \mathbf{N} \right\}.$$

Denote by $\mathbf{F}^v \subseteq E \times \mathbb{R}^{M\cdot k}$ its graph with $\mathcal{F}^v$ the corresponding σ-algebra. Then:

(i) $(\int \mathbf{g}^M \odot \mathbf{N}\,\mathrm{d}\mathbf{P})$ is an $\mathscr{E}$-measurable mapping to $K^*(\mathbb{R}^{M\cdot k})$, and $\mathbf{F}^v$ is a measurable set in $E \times \mathbb{R}^{M\cdot k}$;

(ii) There exists a measurable, $\mathbb{R}^l$-valued function $\mathbf{h}$ on $(\mathbf{F}^v \times \Omega, \mathcal{F}^v \otimes \mathcal{A})$ such that for any $(e, \mathbf{x}) \in \mathbf{F}^v \subseteq E \times \mathbb{R}^{M\cdot k}$ with $e \in E$ and $\mathbf{x} \in \mathbb{R}^{M\cdot k}$, the following assertions hold: (i) $\mathbf{h}(e, \mathbf{x}, \omega) \in \mathbf{N}(e, \omega)$; and (ii) $\mathbf{x} = \int \mathbf{g}^M(e, \omega) \odot \mathbf{h}(e, \mathbf{x}, \omega)\mathbf{P}(\mathrm{d}\omega|e)$, where $\mathbf{g}^M(e, \omega) \odot \mathbf{h}(e, \mathbf{x}, \omega) := (\mathbf{g}_1(e, \omega) \cdot \mathbf{h}(e, \mathbf{x}, \omega), ..., \mathbf{g}_M(e, \omega) \cdot \mathbf{h}(e, \mathbf{x}, \omega))$.

Proof of Lemma 3. (1) Let $\tilde{\mathbf{P}}(\cdot|\cdot)$ be a $\mathbb{R}^{l\cdot k}$-valued map where $\tilde{\mathbf{P}}(\mathrm{d}\omega|e) := (u_n \mathbf{P}(\mathrm{d}\omega|e))_{1 \le n \le l}$. Given the boundedness of $\mathbf{P}$ and $\mathbf{u}$, $\tilde{\mathbf{P}}$ is also a bounded kernel. Moreover, for the fixed element $e \in E$, an atom of $\tilde{\mathbf{P}}(\cdot|e)$ must be an atom of $\mathbf{P}(\cdot|e)$ as well. For any measurable selection $\mathbf{f} := (f_1, ..., f_l)$ from the convex hull of $\mathbf{N}(e, \cdot)$ such that $\mathbf{f} \in \mathbf{N}(e, \cdot)$ on the atoms of $\mathbf{P}(\mathrm{d}\omega|e)$, then $\mathbf{f}$ is also a measurable

[21] The result of Mertens [2003] is not directly applicable to our context since the domain of the mapping he constructed is on the graph the tensor product of the vector-valued integration. However, the version we need is on a different domain, that is, the graph of the inner product of the vector-valued integration in a particular way (see the definition of $\odot$ in P. 22). Such discrepancy requires us to reduce Mertens' domain by extra construction that involves other measurable selections.

[22] A $\mathbb{R}^k$-valued *kernel* $\mathbf{P}(\cdot|\cdot)$ is a map from $E \times \mathcal{A} \to \mathbb{R}^k$ such that (i) for any $e \in E$, $\mathbf{P}(\cdot|e)$ is a $\mathbb{R}^k$-valued measure on $(\Omega, \mathcal{A})$; (ii) $\mathbf{P}(A|\cdot)$ is $\mathscr{E}$-measurable, $\forall A \in \mathcal{A}$. A kernel $\mathbf{P}(\cdot|\cdot)$ is *bounded* if there exists a constant C such that $\|\mathbf{P}(A|e)\| \le C$ for all $A \in \mathcal{A}$ and $e \in E$.

selection of the convex hull of $\mathbf{N}(e,\cdot)$ such that $\mathbf{f} \in \mathbf{N}(e,\cdot)$ on the atoms of $\tilde{\mathbf{P}}(\mathrm{d}\omega|e)$. Let $\int \mathbf{N}\,\mathrm{d}\tilde{\mathbf{P}}(e) := \left\{\int \mathbf{f}'(\omega,e)\tilde{\mathbf{P}}(\mathrm{d}\omega|e)\middle|\mathbf{f}' \in \mathbf{N}\right\}$. Hence by Theorem (2), Mertens [2003], $\int \mathbf{f}\,\mathrm{d}\tilde{\mathbf{P}}(e) \in \int \mathbf{N}\,\mathrm{d}\tilde{\mathbf{P}}(e)$. Given that such inclusion holds for each term in Kronecker product, which implies $\int \mathbf{u}\cdot\mathbf{f}(\omega)\mathbf{P}(\mathrm{d}\omega|e) \in (\int \mathbf{u}\cdot\mathbf{N}\,\mathrm{d}\mathbf{P})(e)$.

(2) Fix an arbitrary M-component vector function $\mathbf{g}^M := (\mathbf{g}_1, ..., \mathbf{g}_M)$ which satisfies that each $\mathbf{g}_m$ is a $\mathbb{R}^l$-valued $\mathscr{E}\otimes\mathcal{A}$-measurable bounded function. Define a $\mathbb{R}^M$-valued correspondence $\hat{\mathbf{N}}$ from $(E\times\Omega, \mathscr{E}\otimes\mathcal{A})$ to $K^*(\mathbb{R}^M)$ such that for each $(e,\omega) \in E\times\Omega$, $\hat{\mathbf{N}}(e,\omega) := \left\{(\mathbf{g}_1(e,\omega)\cdot\mathbf{a}, ..., \mathbf{g}_M(e,\omega)\cdot\mathbf{a})\middle|\mathbf{a}\in\mathbf{N}(e,\omega)\right\}$. By the given conditions, the above correspondence $\hat{\mathbf{N}}$ is $\mathbf{P}$-integrable and $\hat{\mathbf{N}}$ is compact-valued.

(i) By Theorem (1) (a) and (b), Mertens [2003], $(\int \hat{\mathbf{N}}\,\mathrm{d}\mathbf{P})$ is an $\mathscr{E}$-measurable mapping to $K^*(\mathbb{R}^{M\cdot k})$, and its graph is a measurable set in $E\times\mathbb{R}^{M\cdot k}$.

(ii) Denote by $\hat{\mathbf{F}}$ the graph of $\int \hat{\mathbf{N}}\,\mathrm{d}\mathbf{P}$ with its sub σ-algebra $\hat{\mathcal{F}} \subseteq \mathscr{E}\otimes\mathcal{B}(\mathbb{R}^{M\cdot k})$. By Theorem (3), Mertens [2003], there exists a measurable, $\mathbb{R}^M$-valued function $\hat{\mathbf{f}}$ on $(\hat{\mathbf{F}}\times\Omega, \hat{\mathcal{F}}\otimes\mathcal{A})$ such that $\hat{\mathbf{f}}(e,\hat{\mathbf{x}},\omega) \in \hat{\mathbf{N}}(e,\omega)$ and

$$\hat{\mathbf{x}} = \int \hat{\mathbf{f}}(e,\hat{\mathbf{x}},\omega)\mathbf{P}(\mathrm{d}\omega|e). \tag{15}$$

For each element $(e,\hat{\mathbf{x}},\omega)$, by the definition of $\hat{\mathbf{N}}(e,\omega)$ and that $\hat{\mathbf{f}}(e,\hat{\mathbf{x}},\omega) \in \hat{\mathbf{N}}(e,\omega)$, there exists an element $\hat{\mathbf{a}}'_{e,\hat{\mathbf{x}},\omega} \in \hat{\mathbf{N}}(e,\omega)$ such that $\hat{\mathbf{f}}(e,\hat{\mathbf{x}},\omega) = (\mathbf{g}_1\cdot\hat{\mathbf{a}}'_{e,\hat{\mathbf{x}},\omega}, ..., \mathbf{g}_M\cdot\hat{\mathbf{a}}'_{e,\hat{\mathbf{x}},\omega})$. Based on the above $\hat{\mathbf{f}}$, define a correspondence $\mathbf{H}: E\times\hat{\mathbf{F}}\times\Omega$ to $\mathbb{R}^l$ such that for any $(e,\hat{\mathbf{x}},\omega)$,

$$\mathbf{H}(e,\hat{\mathbf{x}},\omega) := \{\hat{\mathbf{a}}''_{e,\hat{\mathbf{x}},\omega} \in \mathbf{N}(e,\omega) | (\mathbf{g}_n\cdot\hat{\mathbf{a}}''_{e,\hat{\mathbf{x}},\omega})_{1\le k\le M} = \hat{\mathbf{f}}(e,\hat{\mathbf{x}},\omega)\}.$$

By such a definition, $\mathbf{H}(e,\hat{\mathbf{x}},\omega)$ is measurable, nonempty-valued and compact-valued (for the measurability, see Corollary 18.8 in Aliprantis and Border 2006). Therefore, Kuratowski–Ryll-Nardzewski Selection Theorem implies that we can find a pointwise measurable selector $\mathbf{h}$ of $\mathbf{H}$. Therefore, for each $(e,\hat{\mathbf{x}},\omega)$, $\mathbf{h}(e,\hat{\mathbf{x}},\omega) \in \mathbf{N}(e,\omega)$, and by (15) and that $\mathbf{h} \in \mathbf{H}$ pointwisely, we have $\hat{\mathbf{x}} = \int \mathbf{g}^M(e,\omega)\odot\mathbf{h}(e,\hat{\mathbf{x}},\omega)\mathbf{P}(\mathrm{d}\omega|e)$. □

We introduce a more general environment under which our main results still hold, and this environment includes the setting in the main text (with independent information) as a special case.

Definition 6. The setting has *interdependent information amongst players with subjective priors* ("interdependent information" for short) if the following hold: let $(T\times\Omega, \mathcal{T}\otimes\mathcal{F}, \mu^T\times\mu^\Omega)$ be the product probability space of state and types. Assume that μ^Ω is atomless. Denote by $\tilde{\mu}_i \in \Delta(T\times\Omega)$ the subjective belief of each player i, $i \in \mathcal{I}\cup\{0\}$. Suppose that $\tilde{\mu}_i$ is absolutely continuous with respect to the product probability measure $\mu^T\times\mu^\Omega$ with the density function $\ell_i(t,\omega)$, i.e., $\mathrm{d}\tilde{\mu}_i = \ell_i(t,\omega)\,\mathrm{d}(\mu^T\times\mu^\Omega)$. For each $i \in \mathcal{I}\cup\{0\}$, the density function $\ell_i(t,\omega)$ is assumed to be bounded and measurable.

We will prove the conclusion of Theorem 2 in the setting of interdependent information, which we numbered as Theorem 2′. We say the density function of a player i's prior is *separable* if $\ell_i(t,\omega)$ can be written into the following form: $\ell_i(t,\omega) = \sum_{n=1}^{N} \ell_i^n(\omega) \cdot \tilde{\ell}_i^n(t)$, where N is a positive integer; ℓ_i^n and $\tilde{\ell}_i^n$, $1 \leq n \leq N$, are bounded measurable functions. In such a general setting, this separability condition of density functions is required.

Theorem 2′. *Suppose that the setting is of interdependent information and the density function of each player's prior is separable, then the statement of Theorem 2 holds.*

Proof of Theorem 2′. For convenience, let N be the largest index such that for each player $i \in \mathcal{I} \cup \{0\}$, $\ell_i(t,\omega) := \sum_{\tilde{n}=1}^{N} \ell_i^{\tilde{n}}(\omega) \cdot \tilde{\ell}_i^{\tilde{n}}(t)$, and $u_i(t,\omega,a) := \sum_{n=1}^{N} f_{i,n}(\omega,a) \cdot g_{i,n}(t,a)$ with bounded functions $f_{i,n}$, $g_{i,n}$, $\ell_i^{\tilde{n}}$ and $\tilde{\ell}_i^{\tilde{n}}$ for $1 \leq n, \tilde{n} \leq N$.

Let $P(\cdot|\cdot)$ be a stochastic kernel from $\widehat{T}$ to $\Delta(\Omega)$ such that $P(\cdot|\cdot) \equiv \mu^{\Omega}(\cdot)$. For convenience, we label the elements in $\hat{A}$ and A as $\hat{A} := \{\hat{a}^1, \hat{a}^2, ..., \hat{a}^{|\hat{A}|}\}$ and $A := \{a^1, ..., a^{|A|}\}$, respectively. For any m with $1 \leq m \leq |\hat{A}|$, let $\mathbf{e}_m \in \mathbb{R}^{|\hat{A}|}$ be the coordinate vector that all components in $\mathbf{e}_m$ are zero except that its m-th component is 1. Define a $\mathbb{R}^{|\hat{A}|}$-valued $\widehat{\mathcal{T}} \otimes \mathcal{F}$-measurable constant correspondence $\tilde{\mathbf{N}}$ on $\widehat{T} \times \Omega$ such that $\tilde{\mathbf{N}}(\cdot,\cdot) \equiv \left\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_{|\hat{A}|}\right\}$.

We will now construct a vector-valued mapping $\mathbf{F}^{\mathcal{N}} := (\mathbf{F}_1, \mathbf{F}_2, ..., \mathbf{F}_{\mathcal{N}})$ with each $\mathbf{F}_k$ a $\mathbb{R}^{|\hat{A}|}$-valued mapping and the largest index $\mathcal{N} := (|A|N+1) \cdot |\hat{A}|N(|\mathcal{I}|+1)$. Before the explicit construction, we define the following two sets for indexing purposes: let S_1 and S_2 be finite sets of integer-valued vectors, where

$$S_1 := \left\{(\hat{j}, j, i, n, \tilde{n}) \in \mathbb{N}^5 \middle| 1 \leq \hat{j} \leq |\hat{A}|, 1 \leq j \leq |A|, 0 \leq i \leq |\mathcal{I}|, 1 \leq n, \tilde{n} \leq N\right\};$$
$$S_2 := \left\{(\hat{j}, i, \tilde{n}) \in \mathbb{N}^3 \middle| 1 \leq \hat{j} \leq |\hat{A}|, 0 \leq i \leq |\mathcal{I}|, 1 \leq \tilde{n} \leq N\right\}.$$

For expositional convenience, we also label the elements in $S_{j'}$ as $S_{j'} := \{s^{j',1}, s^{j',2}, ..., s^{j',|S_{j'}|}\}$ for $j' = 1, 2$, where $|S_1| = |A||\hat{A}|N^2(|\mathcal{I}|+1)$ and $|S_2| = |\hat{A}|N(|\mathcal{I}|+1)$. We are now ready to define each $\mathbf{F}_k : \widehat{T} \times \Omega \to \mathbb{R}^{|\hat{A}|}$ for any $1 \leq k \leq \mathcal{N}$. For any integer k,

(1) if the integer k satisfies that $1 \leq k \leq |S_1|$, then we define $\mathbf{F}_k$ based on the vector $s^{1,k} = (\hat{j}_k, j_k, i_k, n_k, \tilde{n}_k)$ as follows:

$$\mathbf{F}_k(\hat{t},\omega) := f_{i_k,n_k}(\omega, a_{j_k}) \ell_{i_k}^{\tilde{n}_k}(\omega) \cdot \mathbf{e}_{\hat{j}_k} \text{ for any } (\hat{t},\omega). \tag{16}$$

(2) if the integer k satisfies that $|S_1|+1 \leq k \leq |S_1|+|S_2|$, then we define $\mathbf{F}_k$ based on the vector $s^{2,\hat{k}} = (\hat{j}_{\hat{k}}, i_{\hat{k}}, \tilde{n}_{\hat{k}})$ with $\hat{k} := k - |S_1|$ as follows:

$$\mathbf{F}_k(\hat{t},\omega) := \ell_{i_{\hat{k}}}^{\tilde{n}_{\hat{k}}}(\omega) \cdot \mathbf{e}_{\hat{j}_{\hat{k}}} \text{ for any } (\hat{t},\omega). \tag{17}$$

Given such $\mathbf{F}^{\mathcal{N}}$, we further define a vector-valued integral $\int \mathbf{F}^{\mathcal{N}} \odot \tilde{\mathbf{N}} \, \mathrm{d}P := (\int \mathbf{F}_k \cdot \tilde{\mathbf{N}} \, \mathrm{d}P)_{1 \leq k \leq \mathcal{N}}$ that each $\int \mathbf{F}_k \cdot \tilde{\mathbf{N}} \, \mathrm{d}P : \widehat{T} \twoheadrightarrow \mathbb{R}$ is a measurable correspondence and

$$\int \mathbf{F}_k \cdot \tilde{\mathbf{N}} \, \mathrm{d}P : \widehat{T} \twoheadrightarrow \Big\{ \int_{\Omega} \mathbf{F}_k(\hat{t},\omega) \cdot \mathbf{h}(\hat{t},\omega) P(\mathrm{d}\omega|\hat{t}) \Big|$$

$$\mathbf{h} \text{ is an } \widehat{\mathcal{T}} \otimes \mathcal{F}\text{-measurable selection from } \tilde{\mathbf{N}} \Big\}.$$

Denote by $(\mathbf{G}^u, \mathcal{G}^u)$ the graph of $\int \mathbf{F}^{\mathcal{N}} \odot \tilde{\mathbf{N}} \, \mathrm{d}P$ with the associated σ-algebra. Since any persuasion mechanism π is a $\widehat{\mathcal{T}} \otimes \mathcal{F}$-measurable function from $\widehat{T} \times \Omega \to [0,1]^{|\hat{A}|}$ such that $\sum_{\hat{j}=1}^{|\hat{A}|} \pi(\hat{a}^{\hat{j}}|\hat{t}, \omega) \equiv 1$ for any $(\hat{t}, \omega)$, by definition $\pi(\cdot|\hat{t}, \omega)$ is in the convex hull of $\tilde{\mathbf{N}}(\hat{t}, \omega)$ for any $(\hat{t}, \omega)$. By (1) of Lemma 3 and given that $P(\cdot|\cdot) \equiv \mu^{\Omega}(\cdot)$ is atomless, we can conclude that, fix any $\hat{t}$,

$$\int_{\Omega} \mathbf{F}_k(\hat{t}, \omega) \cdot \pi(\cdot|\hat{t}, \omega) P(\mathrm{d}\omega|\hat{t}) \in \int \mathbf{F}_k \cdot \tilde{\mathbf{N}} \, \mathrm{d}P(\hat{t}) \text{ for each } 1 \leq k \leq \mathcal{N}.$$

By (2ii) of Lemma 3, there exists a measurable, $\mathbb{R}^{|\hat{A}|}$-valued function $\mathbf{h}$ on $(\mathbf{G}^u \times \Omega, \mathcal{G}^u \otimes \mathcal{F})$ such that for any $(\hat{t}, \mathbf{x}) \in \mathbf{G}^u$, $\mathbf{h}(\hat{t}, \mathbf{x}, \omega) \in \tilde{\mathbf{N}}(\hat{t}, \omega)$ and

$$\mathbf{x} = \int_{\Omega} \mathbf{F}^{\mathcal{N}}(\hat{t}, \omega) \odot \mathbf{h}(\hat{t}, \mathbf{x}, \omega) P(\mathrm{d}\omega|\hat{t}). \tag{18}$$

By substituting $\mathbf{x}$ with the expression $\int_{\Omega} \mathbf{F}^{\mathcal{N}}(\hat{t}, \omega) \odot \pi(\cdot|\hat{t}, \omega) P(\mathrm{d}\omega|\hat{t})$ in the above equation, for any $\hat{t}$, we have

$$\begin{aligned} &\int_{\Omega} \mathbf{F}^{\mathcal{N}}(\hat{t}, \omega) \odot \pi(\cdot|\hat{t}, \omega) P(\mathrm{d}\omega|\hat{t}) \\ =& \int_{\Omega} \mathbf{F}^{\mathcal{N}}(\hat{t}, \omega) \odot \mathbf{h}\Big(\hat{t}, \int_{\Omega} \mathbf{F}^{\mathcal{N}}(\hat{t}, \omega) \odot \pi(\cdot|\hat{t}, \omega) P(\mathrm{d}\omega|\hat{t}), \omega\Big) P(\mathrm{d}\omega|\hat{t}). \end{aligned} \tag{19}$$

Let $\overline{\pi}$ be a function from $\widehat{T} \times \Omega \to \mathbb{R}^{|\hat{A}|}$ such that

$$\overline{\pi}(\cdot|\hat{t}, \omega) := \mathbf{h}\left(\hat{t}, \int_{\Omega} \mathbf{F}^{\mathcal{N}}(\hat{t}, \omega) \odot \pi(\cdot|\hat{t}, \omega) P(\mathrm{d}\omega|\hat{t}), \omega\right) \text{ for any } (\hat{t}, \omega). \tag{20}$$

That $\mathbf{h}(\hat{t}, \mathbf{x}, \omega) \in \tilde{\mathbf{N}}(\hat{t}, \omega)$ pointwisely implies that $\overline{\pi}(\cdot|\hat{t}, \omega) \in \{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_{|\hat{A}|}\}$ for all $(\hat{t}, \omega)$, and therefore $\overline{\pi}$ is a deterministic persuasion mechanism. Given that π is $\widehat{\mathcal{T}} \otimes \mathcal{F}$-measurable and that $P(\mathrm{d}\omega|\hat{t}) \equiv \mu^{\Omega}$, $\int_{\Omega} \mathbf{F}_k(\hat{t}, \omega) \cdot \pi(\cdot|\hat{t}, \omega) P(\mathrm{d}\omega|\hat{t})$ is $\widehat{\mathcal{T}}$-measurable. Thus, by the measurability of $\mathbf{h}$ and $\int_{\Omega} \mathbf{F}^{\mathcal{N}}(\hat{t}, \omega) \odot \pi(\cdot|\hat{t}, \omega) P(\mathrm{d}\omega|\hat{t})$, the measurability of $\overline{\pi}$ is immediate. Moreover, by Equation (19) and (20), we can conclude that for any $\hat{t}$,

$$\int_{\Omega} \mathbf{F}^{\mathcal{N}}(\hat{t}, \omega) \odot \pi(\cdot|\hat{t}, \omega) P(\mathrm{d}\omega|\hat{t}) = \int_{\Omega} \mathbf{F}^{\mathcal{N}}(\hat{t}, \omega) \odot \overline{\pi}(\cdot|\hat{t}, \omega) P(\mathrm{d}\omega|\hat{t}). \tag{21}$$

By replacing each $\mathbf{F}_k$ in Equation (21) with its explicit form for $1 \leq k \leq \mathcal{N}$, we have the following equations: for any $i \in \mathcal{I} \cup \{0\}$, any $\hat{t} \in \widehat{T}$, $1 \leq j \leq |A|$, $1 \leq \hat{j} \leq |\hat{A}|$, and $1 \leq n, \tilde{n} \leq N$,

$$\int_{\Omega} f_{i,n}(\omega, a^j) \, \ell_i^{\tilde{n}}(\omega) \, \pi(\hat{a}^{\hat{j}}|\hat{t}, \omega) \mu^{\Omega}(\mathrm{d}\omega) = \int_{\Omega} f_{i,n}(\omega, a^j) \, \ell_i^{\tilde{n}}(\omega) \, \overline{\pi}(\hat{a}^{\hat{j}}|\hat{t}, \omega) \mu^{\Omega}(\mathrm{d}\omega) \tag{22}$$

$$\int_{\Omega} \ell_i^{\tilde{n}}(\omega) \, \pi(\hat{a}^{\hat{j}}|\hat{t}, \omega) \mu^{\Omega}(\mathrm{d}\omega) = \int_{\Omega} \ell_i^{\tilde{n}}(\omega) \, \overline{\pi}(\hat{a}^{\hat{j}}|\hat{t}, \omega) \mu^{\Omega}(\mathrm{d}\omega). \tag{23}$$

By Equation (22), for $i \in \mathcal{I} \cup \{0\}$, any $\hat{t} \in \widehat{T}$, $1 \le j \le |A|$, $1 \le \hat{j} \le |\hat{A}|$ and for all $t \in T$, we have

$$\begin{aligned}
&\int_\Omega u_i(t,\omega,a^j)\pi(\hat{a}^{\hat{j}}|\hat{t},\omega)\,\ell_i(t,\omega)\mu^\Omega(\mathrm{d}\omega)\\
=&\sum_{\tilde{n}=1}^{N}\sum_{n=1}^{N}\tilde{\ell}_i^{\tilde{n}}(t)\,g_{i,n}(t,a^j)\int_\Omega f_{i,n}(\omega,a^j)\,\pi(\hat{a}^{\hat{j}}|\hat{t},\omega)\,\ell_i^{\tilde{n}}(\omega)\mu^\Omega(\mathrm{d}\omega)\\
=&\sum_{\tilde{n}=1}^{N}\sum_{n=1}^{N}\tilde{\ell}_i^{\tilde{n}}(t)\,g_{i,n}(t,a^j)\int_\Omega f_{i,n}(\omega,a^j)\,\overline{\pi}(\hat{a}^{\hat{j}}|\hat{t},\omega)\,\ell_i^{\tilde{n}}(\omega)\mu^\Omega(\mathrm{d}\omega)\\
=&\int_\Omega u_i(t,\omega,a^j)\overline{\pi}(\hat{a}^{\hat{j}}|\hat{t},\omega)\,\ell_i(t,\omega)\mu^\Omega(\mathrm{d}\omega).
\end{aligned} \tag{24}$$

Similarly, by Equation (23), for $i \in \mathcal{I} \cup \{0\}$, any $\hat{t} \in \widehat{T}$ and $1 \le \hat{j} \le |\hat{A}|$, we have

$$\int_\Omega \ell_i(t,\omega)\,\pi(\hat{a}^{\hat{j}}|\hat{t},\omega)\mu^\Omega(\mathrm{d}\omega) = \int_\Omega \ell_i(t,\omega)\,\overline{\pi}(\hat{a}^{\hat{j}}|\hat{t},\omega)\mu^\Omega(\mathrm{d}\omega).$$

Our conclusion follows directly from Lemma 4 presented below. □

Lemma 4. *In the setting of interdependent information, assume that players have separable densities. Suppose that for two persuasion mechanisms π and $\overline{\pi}$, the following hold for any $t \in T$, $\hat{t} \in \widehat{T}$, $\hat{a} \in \hat{A}$, $a \in A$ and any $i \in \mathcal{I} \cup \{0\}$:*

$$\int_\Omega u_i(t,\omega,a)\,\ell_i(t,\omega)\,\pi(\hat{a}|\hat{t},\omega)\mu^\Omega(\mathrm{d}\omega) = \int_\Omega u_i(t,\omega,a)\,\ell_i(t,\omega)\,\overline{\pi}(\hat{a}|\hat{t},\omega)\mu^\Omega(\mathrm{d}\omega) \tag{25}$$

$$\int_\Omega \ell_i(t,\omega)\,\pi(\hat{a}|\hat{t},\omega)\mu^\Omega(\mathrm{d}\omega) = \int_\Omega \ell_i(t,\omega)\,\overline{\pi}(\hat{a}|\hat{t},\omega)\mu^\Omega(\mathrm{d}\omega). \tag{26}$$

Then (i) π is effectively equivalent to $\overline{\pi}$; (ii) any equilibrium under π is still an equilibrium under $\overline{\pi}$, and vice versa; (iii) moreover, for any equilibrium (r^*,σ^*) under π, then the expected equilibrium payoff of each type of each receiver in (r^*,σ^*) is the same under π and under $\overline{\pi}$, and the expected equilibrium payoff of the sender in (r^*,σ^*) is the same under π and under $\overline{\pi}$.

Proof. Fix an arbitrary strategy profile $(\tilde{\sigma},\tilde{r})$. For any $i \in \mathcal{I}$ and the type profile $\tilde{t}$, any message profile $\hat{a}'$ and the action profile a', let $\tilde{\sigma}(a' \mid \tilde{t},\tilde{r},\hat{a}') := \prod_{j=1}^{|\mathcal{I}|}\tilde{\sigma}_j(a'_j \mid \tilde{t}_j,\hat{t}_j,\hat{a}'_j)$ be the probability of playing a' given the above $(\tilde{\sigma},\tilde{r},\tilde{t})$. By (25), the following holds

$$\begin{aligned}
&\int_{\widehat{T}}\int_\Omega \tilde{\sigma}(a' \mid \tilde{t},\tilde{r},\hat{a}')u_i(\tilde{t},\omega,a')\pi(\hat{a}'|\hat{t},\omega)\,\ell_i(\tilde{t},\omega)\mu^\Omega(\mathrm{d}\omega)\tilde{r}(\mathrm{d}\hat{t} \mid \tilde{t})\\
=&\int_{\widehat{T}}\int_\Omega \tilde{\sigma}(a' \mid \tilde{t},\tilde{r},\hat{a}')\,u_i(\tilde{t},\omega,a')\,\overline{\pi}(\hat{a}'|\hat{t},\omega)\,\ell_i(\tilde{t},\omega)\mu^\Omega(\mathrm{d}\omega)\tilde{r}(\mathrm{d}\hat{t} \mid \tilde{t}),
\end{aligned} \tag{27}$$

Given the above $(\tilde{r},\tilde{t})$, $\hat{a}' \in \hat{A}$ and $i \in \mathcal{I}$, by taking the expectations over $\widehat{T}$ in (26) according to $\tilde{r}(\mathrm{d}\hat{t} \mid \tilde{t})$, we have

$$\int_{\Omega}\int_{\widehat{T}}\pi(\hat{a}'|\hat{t},\omega)\,\ell_i(\tilde{t},\omega)\tilde{r}(\mathrm{d}\hat{t}\mid\tilde{t})\mu^{\Omega}(\mathrm{d}\omega)=\int_{\Omega}\int_{\widehat{T}}\overline{\pi}(\hat{a}'|\hat{t},\omega)\,\ell_i(\tilde{t},\omega)\tilde{r}(\mathrm{d}\hat{t}\mid\tilde{t})\mu^{\Omega}(\mathrm{d}\omega). \tag{28}$$

Let us use the same notation $U_i^{\pi}(\tilde{\sigma}_i,\tilde{\sigma}_{-i},\tilde{r}_i,\tilde{r}_{-i}|\tilde{t})$ (resp. $U_i^{\overline{\pi}}(\tilde{\sigma}_i,\tilde{\sigma}_{-i},\tilde{r}_i,\tilde{r}_{-i}|\tilde{t})$) to denote receiver i's expected utility conditional on type profile $\tilde{t}$ in persuasion mechanism π (resp. $\overline{\pi}$) in this more general setting. Label the elements in $\hat{A}$ and A as $\hat{A}:=\{\hat{a}^1,...,\hat{a}^{|\hat{A}|}\}$ and $A:=\{a^1,...,a^{|A|}\}$. By Equation (28) and (27), given the above $(\tilde{\sigma},\tilde{r})$, for any $i\in\mathcal{I}$ and any type profile $\tilde{t}$, we have

$$\begin{aligned}
&U_i^{\pi}(\tilde{\sigma}_i,\tilde{\sigma}_{-i},\tilde{r}_i,\tilde{r}_{-i}|\tilde{t})\\
=&\sum_{\hat{j}=1}^{|\hat{A}|}\sum_{j=1}^{|A|}\frac{\int_{\widehat{T}}\int_{\Omega}\tilde{\sigma}(a^j\mid\tilde{t},\tilde{r},\hat{a}^{\hat{j}})\,u_i(\tilde{t},\omega,a^j)\pi(\hat{a}^{\hat{j}}|\hat{t},\omega)\,\ell_i(\tilde{t},\omega)\mu^{\Omega}(\mathrm{d}\omega)\tilde{r}(\mathrm{d}\hat{t}\mid\tilde{t})}{\sum_{\hat{j}=1}^{|\hat{A}|}\int_{\Omega}\int_{\widehat{T}}\pi(\hat{a}^{\hat{j}}|\hat{t},\omega)\,\ell_i(\tilde{t},\omega)\tilde{r}(\mathrm{d}\hat{t}\mid\tilde{t})\mu^{\Omega}(\mathrm{d}\omega)}\\
=&\sum_{\hat{j}=1}^{|\hat{A}|}\sum_{j=1}^{|A|}\frac{\int_{\widehat{T}}\int_{\Omega}\tilde{\sigma}(a^j\mid\tilde{t},\tilde{r},\hat{a}^{\hat{j}})\,u_i(\tilde{t},\omega,a^j)\overline{\pi}(\hat{a}^{\hat{j}}|\hat{t},\omega)\,\ell_i(\tilde{t},\omega)\mu^{\Omega}(\mathrm{d}\omega)\tilde{r}(\mathrm{d}\hat{t}\mid\tilde{t})}{\sum_{\hat{j}=1}^{|\hat{A}|}\int_{\Omega}\int_{\widehat{T}}\overline{\pi}(\hat{a}^{\hat{j}}|\hat{t},\omega)\,\ell_i(\tilde{t},\omega)\tilde{r}(\mathrm{d}\hat{t}\mid\tilde{t})\mu^{\Omega}(\mathrm{d}\omega)}\\
=&U_i^{\overline{\pi}}(\tilde{\sigma}_i,\tilde{\sigma}_{-i},\tilde{r}_i,\tilde{r}_{-i}|\tilde{t}).
\end{aligned}\tag{29}$$

As such, by definition, $\pi\succcurlyeq_i\overline{\pi}$ and $\pi\preccurlyeq_i\overline{\pi}$ for each $i\in\mathcal{I}$. Thus we conclude that π is effectively equivalent to $\overline{\pi}$.

(ii) Again, let us use the same notation $U_i^{\pi}(\tilde{\sigma}_i,\tilde{\sigma}_{-i},\tilde{r}_i,\tilde{r}_{-i}|\tilde{t}_i)$ (resp. $U_i^{\overline{\pi}}(\tilde{\sigma}_i,\tilde{\sigma}_{-i},\tilde{r}_i,\tilde{r}_{-i}|\tilde{t}_i)$) to denote receiver i's expected utility conditional on his type $\tilde{t}_i$ in persuasion mechanism π (resp. $\overline{\pi}$) in the given setting.

Building on the above proof, for any strategy profile (r,σ), for any $t_i\in T_i$ and $i\in\mathcal{I}$, by taking the conditional expectation of t_{-i} over T_{-i} in equation (29), we have

$$U_i^{\pi}(\sigma,r|t_i)=U_i^{\overline{\pi}}(\sigma,r|t_i). \tag{30}$$

By equation (30), if receivers follow the strategy (r,σ), then the interim expected payoff of type t_i of receiver i is the same under π and under $\overline{\pi}$. Now consider any equilibrium $(r^*=(r_i^*)_{i\in\mathcal{I}},\sigma^*=(\sigma_i^*)_{i\in\mathcal{I}})$ under π. Then the following holds for any type t_i of any receiver i:

$$U_i^{\pi}(\sigma^*,r^*|t_i)\geq\sup_{\hat{\sigma}_i\in\Sigma_i,\hat{r}_i\in R_i}U_i^{\pi}(\hat{\sigma}_i,\sigma_{-i}^*,\hat{r}_i,r_{-i}^*|t_i).$$

By (30), the same strategy profile (r^*,σ^*) is also a best response under persuasion mechanism $\overline{\pi}$, and thus (r^*,σ^*) is also an equilibrium under $\overline{\pi}$. Similarly, if (r^*,σ^*) is an equilibrium under $\overline{\pi}$, then it is also an equilibrium under π.

(iii) For any equilibrium (r^*,σ^*), (30) implies that the equilibrium payoff of each type of each receiver is the same under π and $\overline{\pi}$. The preservation of the sender's expected utility in (r^*,σ^*) under π and $\overline{\pi}$ is follows by taking expectation of Equation (25) over A, $\hat{A}$ and $\widehat{T}$ for $i=0$. Thus we conclude our proof. $\square$

A.2 Proofs of Section 4

Proof of Lemma 1. **The "only if" direction**: Suppose that the pair $(q(\cdot), p(\cdot))$ can be implemented by a deterministic persuasion mechanism $1_{X(\cdot)}(\cdot)$ where $X(\hat{t})$ is a measurable subset in $[0, 1]$ for every $\hat{t}$. Then $0 \leq p(\hat{t}) = \int_{X(\hat{t})} \omega \, \mathrm{d}\omega \leq \int_{X(\hat{t})} \mathrm{d}\omega = q(\hat{t}) \leq 1$ is satisfied for every $\hat{t}$.

At each $\hat{t}$, note that the measure of $X(\hat{t})$ is $q(\hat{t})$. Under the uniform distribution, this is the same as that of the interval $[0, q(\hat{t})]$ and of $[1 - q(\hat{t}), 1]$. Thus $\int_{X(\hat{t})} \mathrm{d}\omega = q(\hat{t}) = \int_0^{q(\hat{t})} \mathrm{d}\omega = \int_{1-q(\hat{t})}^1 \mathrm{d}\omega$.

Since $f(\omega) = \omega$ is an increasing function, at each $\hat{t}$, we have

$$\frac{q(\hat{t})^2}{2} = \int_0^{q(\hat{t})} \omega \, \mathrm{d}\omega \leq \int_{X(\hat{t})} \omega \, \mathrm{d}\omega = p(\hat{t}) \leq \int_{1-q(\hat{t})}^1 \omega \, \mathrm{d}\omega = q(\hat{t})(1 - \frac{q(\hat{t})}{2}).$$

Thus (11) is satisfied and we conclude the "only if" direction.

The "if" direction: Suppose that a pair (q, p) satisfies (11) for every $\hat{t} \in [0, 1]$. We will construct a deterministic persuasion mechanism that implements this pair. Let $H(x, y) := \int_x^{x+y} \omega \, \mathrm{d}\omega$. Note that the inequalities in (11) implies that for any $\hat{t} \in [0, 1]$,

$$H(0, q(\hat{t})) \leq p(\hat{t}) \leq H(1 - q(\hat{t}), q(\hat{t})).$$

Given that $H(\cdot, q(\hat{t}))$ is continuous with respect to x, by the intermediate value theorem and that $H(\cdot, q(\hat{t}))$ strict increases as x increases, there exists a unique $x^*(\hat{t}) \in [0, 1 - q(\hat{t})]$ such that $p(\hat{t}) = H(x^*(\hat{t}), q(\hat{t}))$. It is straightforward to verify that the following persuasion mechanism

$$\pi(\omega, \hat{t}) := 1_{[x^*(\hat{t}), x^*(\hat{t}) + q(\hat{t})]}(\omega)$$

implements the given pair (q, p). Thus we conclude our proof. $\square$

In the following, we consider how to solve problem (12). We first develop a lemma that identifies an implicit requirement that the constraint (11) imposes on (q, U).

Lemma 5. *For any $t \in [0, 1]$, if $U(t) + q(t)t \leq q(t)(1 - \frac{q(t)}{2})$, then $U(t) \leq \frac{(1-t)^2}{2}$.*

Proof. For any $t \in [0, 1]$, note that $U(t) + q(t)t \leq q(t)(1 - \frac{q(t)}{2})$ is equivalent to $0 \leq -\frac{q(t)^2}{2} + q(t)(1 - t) - U(t)$. Since $-\frac{q(t)^2}{2} + q(t)(1 - t) - U(t)$ is a quadratic function of $q(t)$ that opens downward, it must have at least one root for this inequality to hold. Hence, the determinant must be nonnegative $(1 - t)^2 - 2U(t) \geq 0 \Rightarrow U(t) \leq \frac{(1-t)^2}{2}$. $\square$

We now identify the exact form of the solution to the control problem (12).

Lemma 6. *A solution to* (12) *takes the following form: there exists a cutoff* $t_0 \in [0,1]$ *such that*

$$q^*(t) = \begin{cases} 1-t & 0 \le t \le t_0; \\ 1-t_0 & t_0 \le t \le \frac{1+t_0}{2}; \\ 0 & t \ge \frac{1+t_0}{2}; \end{cases} \quad U^*(t) = \begin{cases} \frac{(1-t)^2}{2} & 0 \le t \le t_0; \\ \frac{1-t_0^2}{2} - (1-t_0)t & t_0 \le t \le \frac{1+t_0}{2}; \\ 0 & t \ge \frac{1+t_0}{2}. \end{cases} \tag{31}$$

Proof. Fix an arbitrary pair (q, U) that satisfies all the constraints in (12). In the following, we will argue that we can construct an alternative pair that satisfies (31) under which the sender's payoff is weakly higher than that under (q, U).

To construct an alternative pair, for notational convenience, let us first define a function $B(t) := \frac{(1-t)^2}{2}$ for any $t \in [0,1]$. Note that $U(0) = \frac{1}{2} = B(0)$ and $U'(0) = -q(0) = -1 = B'(0)$. Given that $U(1) = 0$, define

$$t^* := \min\{\frac{5}{6}, \quad \min\{t \in [0,1] \mid U(t) = 0\}\}.$$

By Lemma 5, $U(t) \le \frac{(1-t)^2}{2} = B(t)$ for every t. Hence we have

$$U(t^*) \le B(t^*) = B(t^*) + B'(t^*)(t^* - t^*). \tag{32}$$

As (q, U) satisfies the constraints in (12), we have $U'(t) = -q(t)$ with $q(t) \ge 0$ and weakly decreasing, which further implies that $U(t)$ is a convex function on $[0,1]$. Because $U(\cdot)$ is a convex and decreasing function, it follows that

$$U(t^*) \ge U(0) + U'(0)t^* = U(0) - q(0)t^* = B(0) + B'(0)(t^* - 0). \tag{33}$$

Given that $U(t^*)$ is a constant and that $B(t) + B'(t)(t^* - t)$ is continuous with respect to t, by intermediate value theorem, (32)-(33) imply that there must exist an $t_0 \in [0, t^*]$ such that

$$B(t_0) + B'(t_0)(t^* - t_0) = U(t^*). \tag{34}$$

Figure 4 further demonstrates the relation between $B(t)$, $U(t)$ and $B(t_0) + B'(t_0)(t - t_0)$.

Given such a t_0, consider the following alternative pair (q^*, U^*) in the form of (31):

$$q^*(t) := \begin{cases} 1-t & 0 \le t \le t_0; \\ 1-t_0 & t_0 < t \le \frac{1+t_0}{2}; \\ 0 & \text{otherwise}; \end{cases} \quad U^*(t) := \begin{cases} B(t) & 0 \le t \le t_0; \\ B(t_0) + B'(t_0)(t - t_0) & t_0 \le t \le \frac{1+t_0}{2}; \\ 0 & \text{otherwise}. \end{cases} \tag{35}$$

It is straightforward to check that the above pair satisfies all the constraints in (12). The rest of the proof is to argue that the sender achieves a weakly higher payoff under (q^*, U^*) as compared to that under (q, U).

Recall from equation (13) that the sender's payoff is $\int U(t)(2g(t) + g'(t)t)dt$. Note that the weight on

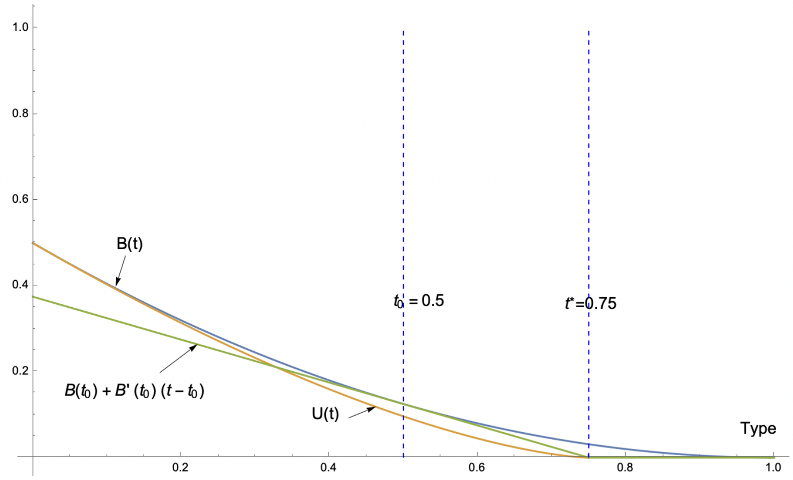


Figure 4: In this demonstration, we take $U(t) = \frac{(1-\frac{4t}{3})^{1.5}}{2}$, which plot is presented in the orange curve. Given this specification, we can determine the value of t^* by its definition that $t^* = 0.75$. The plot of $B(t)$ is the blue curve on the top of $U(t)$, and the green line is the resulting affine function $B(t_0) + B'(t_0)(t - t_0)$ that intersects $U(t)$ at t^*. We can further derive from (34) that $t_0 = 0.5$.

$U(t)$ (i.e., the function $g(t) + g'(t)t$) is strictly positive for $t \in (0, \frac{5}{6})$ and strictly negative for $t \in (\frac{5}{6}, 1)$. We consider each of these two intervals in turn.

We first show that $U^*(t) \geq U(t)$ for $t \in [0, \frac{5}{6}]$. By Lemma 5, $U^*(t) = B(t) \geq U(t)$ holds for any $t \in [0, t_0]$. Note that U is convex, that $B(t_0) \geq U(t_0)$ and that $B(t_0) + B'(t_0)(t^* - t_0) = U(t^*)$. Hence the affine function $B(t_0) + B'(t_0)(t - t_0) \geq U(t)$ for any $t_0 \leq t \leq t^*$. Thus, if $t^* = \frac{5}{6}$, then $U^*(t) \geq U(t)$ for $t \in [0, \frac{5}{6}]$. Otherwise, given that t^* is the smallest type at which $U(t) = 0$ and that U is decreasing, thus $U(t) = 0$ for any $t \in (t^*, \frac{5}{6}]$. Therefore $U^*(t) \geq U(t)$ for any $t \in (t^*, \frac{5}{6})$. Combining all these cases, we conclude that $U^*(t) \geq U(t)$ for $t \in [0, \frac{5}{6}]$.

We now show that $U^*(t) \leq U(t)$ for $t \in (\frac{5}{6}, 1]$. If $U(\frac{5}{6}) = 0$, then $U(t^*) = 0$. Given (34) and that $\frac{1+t_0}{2}$ is the unique solution to $B(t_0) + B'(t_0)(t - t_0) = 0$, we have $t^* = \frac{1+t_0}{2}$. By the construction of U^* in (35), $U^*(t) = 0$ for any $t \in [t^*, 1]$ and thus our claim follows immediately. We now consider when $U(\frac{5}{6}) > 0$, which implies that $t^* = \frac{5}{6}$. Recall that U is decreasing and convex, so for any $t \in [\frac{5}{6}, 1]$, it follows that $U(t) \geq U(\frac{5}{6}) + U'(\frac{5}{6})(t - \frac{5}{6})$. Given that $U^*(t) = B(t_0) + B'(t_0)(t - t_0)$ for any $t \in [\frac{5}{6}, \frac{1+t_0}{2}]$, and $U'(t) = -q(t)$, to show $U^*(t) \leq U(t)$ for any $t \in [\frac{5}{6}, \frac{1+t_0}{2}]$, it is sufficient to show

$$B(t_0) + B'(t_0)(t - t_0) \leq U(\frac{5}{6}) - q(\frac{5}{6})(t - \frac{5}{6}) \text{ for any } t \in [\frac{5}{6}, \frac{1 + t_0}{2}]. \tag{36}$$

First note that, the point $(\frac{5}{6}, U(\frac{5}{6}))$ satisfies the affine functions $y = B(t_0) + B'(t_0)(x - t_0)$ and $y = U(\frac{5}{6}) - q(\frac{5}{6})(x - \frac{5}{6})$ since the former is implied by (34) and the latter is by its definition. Moreover, we claim that the slope of the affine function $B(t_0) + B'(t_0)(t - t_0)$ is steeper than that of $U(\frac{5}{6}) - q(\frac{5}{6})(t - \frac{5}{6})$, that is, $B'(t_0) \leq -q(\frac{5}{6})$. As such, these two affine function crossing at the same point $(\frac{5}{6}, U(\frac{5}{6}))$ with the former steeper than the latter will imply (36). To show our claim that $B'(t_0) \leq -q(\frac{5}{6})$, note that the definition of t_0 and t^* implies that $t_0 \le t^* = \frac{5}{6}$, by which,

$$\begin{aligned} B(t_0) - U(\frac{5}{6}) &\ge U(t_0) - U(\frac{5}{6}) \quad \text{(as } B \ge U) \\ &\ge U'(\frac{5}{6})(t_0 - \frac{5}{6}) \quad \text{(by the convexity of } U). \end{aligned} \tag{37}$$

By plugging in (34) into (37) (recall that $t^* = \frac{5}{6}$ in this case), we have $B'(t_0) \le U'(\frac{5}{6}) = -q(\frac{5}{6})$. Finally for any $t \in (\frac{1+t_0}{2}, 1]$, since $U^*(t) = 0$, it follows immediately that $U^* \le U$ on this interval. Hence we conclude that $U^*(t) = B(t_0) + B'(t_0)(t - t_0) \le U(t)$ for any $t \ge \frac{5}{6}$.

In conclusion, we have shown $U^*(t) \ge U(t)$ for $t \in [0, \frac{5}{6}]$ and $U^*(t) \le U(t)$ for $t \in [\frac{5}{6}, 1]$. Given that the weight function $g(t) + g'(t)t$ is positive on $[0, \frac{5}{6}]$ and negative on $[\frac{5}{6}, 1]$. Therefore the sender obtains a weakly higher payoff under (q^*, U^*) than that under (q, U). We conclude our proof. □

By Lemma 6, the optimal (q^*, U^*) in the form of (31) is completely determined by the cutoff t_0. With the transformation in (13), the sender's problem thus boils down to finding the optimal $t_0 \in [0, 1]$ that solves the following problem:

$$\begin{aligned} &\max_{q^*, U^*} \int_0^1 U^*(t)(2g(t) + g'(t)t)\,\mathrm{d}t \text{ subject to the constraints in (12)} \\ =& \max_{t_0 \in [0,1]} \left(\int_0^{t_0} \frac{(1-t)^2}{2}(2g(t) + g'(t)t)\,\mathrm{d}t + \int_{t_0}^{\frac{1+t_0}{2}} (\frac{1-t_0^2}{2} - (1-t_0)t)(2g(t) + g'(t)t)\,\mathrm{d}t \right) \\ =& \max_{t_0 \in [0,1]} \Big(\frac{25t_0^4}{2} - 32t_0^5 + \frac{85t_0^6}{3} - \frac{60t_0^7}{7} + \\ &+ \frac{1-t_0^2}{2}(-20(-\frac{5(\frac{1+t_0}{2})^4}{4} + \frac{6(\frac{1+t_0}{2})^5}{5}) + 20(-\frac{5t_0^4}{4} + \frac{6t_0^5}{5})) \\ &- (1-t_0)(20((\frac{1+t_0}{2})^5 - (\frac{1+t_0}{2})^6) - 20(t_0^5 - t_0^6)) \Big). \end{aligned}$$

The sender's payoff when she restricts attention to the mechanisms taking the form (31) is a function of t_0 We plot this function as follows:

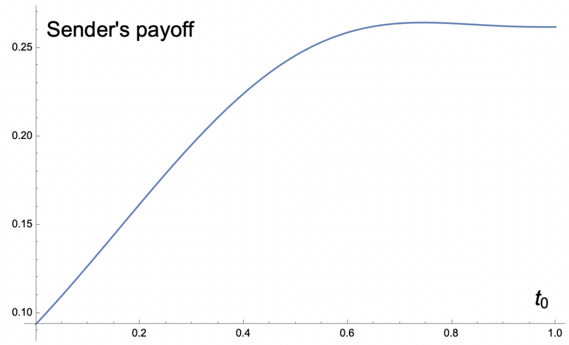


Figure 5: The plot of $\int_0^{t_0} \frac{(1-t)^2}{2}(2g(t) + g'(t)t)\,\mathrm{d}t + \int_{t_0}^{\frac{1+t_0}{2}} (\frac{1-t_0^2}{2} - (1-t_0)t)(2g(t) + g'(t)t)\,\mathrm{d}t$.

The optimal cutoff is 0.7452. Thus the optimal persuasion mechanism π^* is

$$\pi^*(\omega, t) = \begin{cases} 1 & \text{when } 0 \le t \le 0.7452, \omega \ge t; \\ 1 & \text{when } 0.7452 \le t \le 0.8726, \omega \ge 0.7452; \\ 0 & \text{otherwise.} \end{cases}$$

The sender achieves the maximum payoff of 0.2644 under π^*.

A.3 Proofs of Section 5

Proof of Example 1. We show that there is no deterministic information structure that is effectively equivalent to π. Suppose to the contrary that there exists such an information structure f. As receivers have no private types, f does not depend on the receiver's reports. We can write it as $f : [-\frac{1}{3}, \frac{4}{3}] \to [0, 1]$, where, with a slight abuse of notation $f(\omega)$ is the message sent with probability 1 by the information structure. Recall that ν is the distribution of messages induced by π and is a uniform distribution over $[0, 1]$. Moreover, we will focus on the following special class of receiver 1's strategies: For any fixed $t \in [0, 1]$, say a strategy is t-above if receiver 1 takes action 1 if and only if the message realization is greater or equal to t; and t-below if he takes action 1 if and only if the message is strictly less than t.

We show first that effective equivalence implies that the distribution of messages induced by f is the same as under π subject to a message relabelling. Given π and for any $t \in [0, 1]$, receiver 2's ex ante expected payoff is $\nu([t, 1])$ if receiver 1 plays t-above strategy and is $\nu([0, t))$ if he plays the t-below strategy. By $\pi \preccurlyeq_2 f$ in the effective equivalence, there exists a bijection mapping $\psi : [0, 1] \to [0, 1]$ such that for any fixed $t \in [0, 1]$ and the receiver's t-above and below strategy, receiver 2's ex ante payoff under π is less or equal to that under f subject to the given message relabelling, that is, for any fixed $t \in [0, 1]$,

$$\int_\Omega \pi([t, 1]|\omega) \, \mathrm{d}\mu^\Omega = \nu([t, 1]) \le \mu^\Omega \circ f^{-1}(\psi^{-1}([t, 1])), \tag{38}$$

$$\int_\Omega \pi([0, t)|\omega) \, \mathrm{d}\mu^\Omega = \nu([0, t)) \le \mu^\Omega \circ f^{-1}(\psi^{-1}([0, t))). \tag{39}$$

The above two inequalities together imply that $\nu([t, 1]) = \mu^\Omega \circ f^{-1}(\psi^{-1}([t, 1]))$. Hence $\mu^\Omega \circ (\psi \circ f)^{-1} = \nu$.

We then show that ν cannot be induced by any deterministic information structure, which leads to a contradiction. To achieve this goal, we first construct a sequence of $\mu_n \in \Delta([-\frac{1}{3}, \frac{4}{3}])$ with finite support such that μ_n converges to μ^Ω in weak topology. Specifically, for any fixed $n \in \mathbb{N}$, we construct the probability measure μ_n with support $Spt(\mu_n) := \{x_1, ..., x_{3n}\} \subseteq [-\frac{1}{3}, \frac{4}{3}]$ such that

(1) $\{x_1, ..., x_n\}$ are evenly distributed on $[-\frac{1}{3}, \frac{1}{3}]$; $\{x_{n+1}, ..., x_{2n}\}$ are evenly distributed on $(\frac{1}{3}, \frac{2}{3})$ and $\{x_{2n+1}, ..., x_{3n}\}$ evenly distributed on $[\frac{2}{3}, \frac{4}{3}]$.

(2) for any x_j, $\mu_n(\{x_j\}) = \frac{1}{2n}$ if $n + 1 \le j \le 2n$; otherwise, $\mu_n(\{x_j\}) = \frac{1}{4n}$.

By the above construction, we can verify that $\int c(\omega) \, \mathrm{d}\mu_n$ converges to $\int c(\omega) \, \mathrm{d}\mu^\Omega$ as n goes to ∞ for any bounded continuous function $c(\cdot)$ defined on Ω. Thus μ_n weakly converges to μ^Ω as n goes to ∞.

Define the function $\hat{f}$ as follows:

$$\hat{f}(\omega) := \begin{cases} \omega & \omega \in \cup_{n\in\mathbb{N}} Spt(\mu_n), \\ \psi(f(\omega)) & \text{otherwise}. \end{cases}$$

Since the union $\cup_{n\in\mathbb{N}} Spt(\mu_n)$ contains countably many points, $\mu^\Omega(\cup_{n\in\mathbb{N}} Spt(\mu_n)) = 0$ and thus $\hat{f} = \psi \circ f$, μ^Ω-almost surely. Hence $\hat{f}$ is also a deterministic information structure which under μ^Ω induces a distribution of messages equal to ν . Moreover, the definition of $\hat{f}$ implies that $\mu_n \circ \hat{f}^{-1}(E) = \mu_n(E)$ for any measurable set $E \subseteq [0,1]$, which, combines with μ_n weakly converges to μ, implies that

$$\mu_n \circ \hat{f}^{-1} \text{ weakly converge to } \mu^\Omega. \tag{40}$$

We now show that

$$\lim_{n\to\infty} \mu_n \circ \hat{f}^{-1}(E) = \mu^\Omega \circ \hat{f}^{-1}(E) = \nu(E), \forall \text{ Borel measurable set } E \subseteq [0,1], \tag{41}$$

which leads to a contradiction to (40) and concludes the proof. To show (41), note that, for any measurable set E, if the boundary of $\hat{f}^{-1}(E)$ (denoted as $\partial(\hat{f}^{-1}(E))$) is of μ^Ω-measure zero, then the above claim (41) follows directly from Portemanteau theorem (cf. Theorem 13.16 in Klenke [2014]) and the weak convergence of $\{\mu_n\}_{n\in\mathbb{N}}$. So the only thing left is to show that $\mu^\Omega(\partial \hat{f}^{-1}(E)) = 0$ for any measurable set $E \subseteq [0,1]$. For any such $E \subseteq [0,1]$, since $\hat{f}^{-1}(E) \subseteq \mathbb{R}$, its interior is an open set in $\mathbb{R}$. It is known that every open set of $\mathbb{R}$ can be written as a countable union of mutually disjoint open intervals, which therefore implies $\partial(\hat{f}^{-1}(E))$ consists of at most countably many points, which are the endpoints of the intervals in the interior of $\hat{f}^{-1}(E)$. Hence for any measurable set $E \subseteq [0,1]$, $\mu^\Omega(\partial(\hat{f}^{-1}(E))) = 0$. So we conclude the proof. □

Proof of Example 2. The proof is by contradiction. Suppose to the contrary that there exists a deterministic information structure $\overline{\pi}$ that is effectively equivalent to π, where $\overline{\pi}(\cdot) := \mathbb{1}_S(\cdot)$ for some fixed measurable set S. Note that for any receiver type, there exists a pure strategy that is optimal under the given mechanism π, which is one of the following: (i) $\sigma(\cdot) \equiv \delta_0$; (ii) $\sigma(\cdot) \equiv \delta_1$; (iii) $\sigma(1) \equiv \delta_1$ and $\sigma(0) \equiv \delta_0$; (iv) $\sigma(1) \equiv \delta_0$ and $\sigma(0) \equiv \delta_1$. We will consider only the types in $(0.5, \frac{11}{16})$.

For any $t \in (0.5, \frac{11}{16})$, one can check that the optimal pure strategy under π is to always obey the recommendation, under which the expected payoff is the following:

$$\int_\Omega \left(\mathbb{1}_{[t,1]}(\omega)\cdot 0.8 + \mathbb{1}_{[0.5,t]}(\omega)\cdot 0.2 + 0.8\cdot \mathbb{1}_{[0,0.5)}(\omega)\right) \mathrm{d}\omega = 1.1 - 0.6t.$$

Similarly, for any such receiver type, there exists a pure strategy that is optimal under $\overline{\pi}$. Given that π and $\overline{\pi}$ are effectively equivalent, an optimal pure strategy under $\overline{\pi}$ must give this type t the same payoff as well. For any $t \in (0.5, \frac{11}{16})$, the strategies $\sigma \equiv \delta_0$ and $\sigma \equiv \delta_1$ cannot be optimal under $\overline{\pi}$, since the respective payoffs of these two strategies are t and $1-t$. So we consider the two only possible cases for the optimal strategy under $\overline{\pi}$: one is always following the recommendation (obedient strategy), i.e., $\sigma(1) \equiv \delta_1$ and $\sigma(0) \equiv \delta_0$; the other is always defying the recommendation (defiant strategy), i.e., $\sigma(1) \equiv \delta_0$ and $\sigma(0) \equiv \delta_1$. By playing the obedient strategy, each type $t \in (0.5, \frac{11}{16})$ gets the following expected payoff:

$$\begin{aligned} u^o(t) &:= \int_\Omega \Big(\overline{\pi}(\omega) \cdot 1_{[t,1]}(\omega) + \big(1 - \overline{\pi}(\omega)\big) \cdot 1_{[0,t)}(\omega)\Big)\, \mathrm{d}\omega \\ &= \int_\Omega (-2 \cdot \overline{\pi}(\omega) + 1) \cdot 1_{[0,t)}(\omega)\, \mathrm{d}\omega + \int_\Omega \overline{\pi}(\omega)\, \mathrm{d}\omega. \end{aligned} \tag{42}$$

Similarly, by playing the defiant strategy, type $t \in (0.5, \frac{11}{16})$ gets the following payoff:

$$\begin{aligned} u^d(t) &:= \int_\Omega \big((1 - \overline{\pi}(\omega)) \cdot 1_{[t,1]}(\omega) + \overline{\pi}(\omega) \cdot 1_{[0,t)}(\omega)\big)\, \mathrm{d}\omega \\ &= \int_0^t \big(-2 \cdot (1 - \overline{\pi}(\omega)) + 1\big)\, \mathrm{d}\omega + \int_\Omega \big(1 - \overline{\pi}(\omega)\big)\, \mathrm{d}\omega. \end{aligned} \tag{43}$$

Note that the function $u^o(t)$ and $u^d(t)$ are almost everywhere differentiable with respect to t. Let T_O be the subset of types in $(0.5, \frac{11}{16})$ such that the obedient strategy is uniquely optimal amongst all pure strategies, i.e., $T_O := \{t \in (0.5, \frac{11}{16}) \mid u^o(t) > u^d(t)\}$. Let T_D be the subset of types in $(0.5, \frac{11}{16})$ such that the defiant strategy is uniquely optimal amongst all pure strategies, i.e., $T_D := \{t \in (0.5, \frac{11}{16}) \mid u^o(t) < u^d(t)\}$. Let T_B be the subset of types in $(0.5, \frac{11}{16})$ such that both strategies are equally optimal amongst all pure strategies, i.e., $T_B := \{t \in (0.5, \frac{11}{16}) \mid u^o(t) = u^d(t)\}$. As $\sigma \equiv 1$ and $\sigma \equiv 0$ are both suboptimal for any type in $(0.5, \frac{11}{16})$, thus we have $(0.5, \frac{11}{16}) = T_O \cup T_D \cup T_B$.

Given that $\overline{\pi}$ is effectively equivalent to π, the optimal pure strategy under $\overline{\pi}$ must give each receiver type t the same payoff as that under π, therefore $u_o(t) = 1.1 - 0.6t$ for any $t \in T_O$, and similarly, $u_d(t) = 1.1 - 0.6t$ for any $t \in T_D$ and $u_o(t) = u_d(t) = 1.1 - 0.6t$ for any $t \in T_B$. However, the following Claim 1 shows that none of the resulting sets is of strictly positive measure, which leads to a contradiction. The rest is to prove the claim.

Claim 1. *The set* $\{t \in T_O | u_o(t) = 1.1 - 0.6t\}$, $\{t \in T_D | u_d(t) = 1.1 - 0.6t\}$ *and* $\{t \in T_B | u_o(t) = u_d(t) = 1.1 - 0.6t\}$ *are of zero measure.*

The proof of Claim 1 is by contradiction. Consider the case that $\{t \in T_O | u_o(t) = 1.1 - 0.6t\}$ is of strictly positive measure (note that by definition T_O is an open set). Thus $u^{o\prime}(t) = -0.6$ for any $t \in T_O$, i.e., $-0.6 = -2 \cdot \overline{\pi}(t) + 1$ for every $t \in T_O$, and thus $\overline{\pi}(t) = 0.8$ for every $t \in T_O$. This contradicts with the definition that $\overline{\pi}$ is deterministic, i.e., $\overline{\pi} \in \{0, 1\}$ for almost everywhere. For the second case that $\{t \in T_D | u_d(t) = 1.1 - 0.6t\}$ is of strictly positive measure, we can use similar argument to derive $\overline{\pi}(t) = 0.2$ for every $t \in T_D$, which again leads to a contradiction. Since the above two sets are of zero measure, then we have $\{t \in T_B | u_o(t) = u_d(t) = 1.1 - 0.6t\}$ is only different from $(0.5, \frac{11}{16})$ by a zero measure set. However, the same argument as above implies $\overline{\pi} \notin \{0, 1\}$ for almost everywhere $\omega \in T_B$, which is again a contradiction. So we have shown that all of these sets cannot be of positive measure. □

Proof of Example 4. The logic behind the proof is quite similar to Example 2 which is also by contradiction. Suppose to the contrary that there exists a deterministic information structure $\overline{\pi}$ that is effectively equivalent to π, where $\overline{\pi}(\cdot) := \mathbb{1}_S(\cdot)$ for some fixed measurable set S.

For each type t, its belief about the underlying state condition on type t will be the following distribution: $\frac{f(t,\omega)}{\frac{2}{3}(2-t)}\,\mathrm{d}\omega$. Moreover, for any such type, there exists a pure strategy that is optimal under π. We will consider only the types in $(0.5, 0.6)$. For any $t \in (0.5, 0.6)$, one can check that the optimal pure strategy under π is to always obey the recommendation, under which the sender's expected payoff is:

$$\frac{\int_0^{0.5}(\omega-t)0.2\,\mathrm{d}\omega + \int_{0.5}^t 0.8(\omega-t)\,\mathrm{d}\omega + \int_t^1 1.6(\omega-t)\,\mathrm{d}\omega}{2-t} = \frac{0.725+0.4t^2-1.3t}{2-t}.$$

Similarly, for each receiver type, there exists a pure strategy that is optimal under $\overline{\pi}$. Given that π and $\overline{\pi}$ are effectively equivalent, an optimal pure strategy under $\overline{\pi}$ must give this type t the same payoff as well. For any $t \in (0.5, 0.6)$, the strategies $\sigma \equiv \delta_0$ and $\sigma \equiv \delta_1$ cannot be optimal under $\overline{\pi}$, since both $\frac{0.725+0.4t^2-1.3t}{2-t} > 0$ and $\frac{0.725+0.4t^2-1.3t}{2-t} > \frac{2+t^2-4t}{4-2t}$ hold for any $t \in (0.5, 0.6)$. So we consider the two only possible cases for the optimal strategy under $\overline{\pi}$: the obedient strategy and the defiant strategy. Under the obedient strategy, each type $t \in (0.5, 0.6)$ gets an expected payoff:

$$u^o(t) := \frac{\int_t^1 2\overline{\pi}(\omega)(\omega-t)\,\mathrm{d}\omega + \int_0^t \overline{\pi}(\omega)(\omega-t)\,\mathrm{d}\omega}{2-t} \tag{44}$$

Similarly, by playing the defiant strategy, type $t \in (0.5, 0.6)$ gets the following payoff:

$$u^d(t) := \frac{\int_t^1 2(1-\overline{\pi}(\omega))(\omega-t)\,\mathrm{d}\omega + \int_0^t (1-\overline{\pi}(\omega))(\omega-t)\,\mathrm{d}\omega}{2-t}. \tag{45}$$

Note that the function $u^o(t)$ and $u^d(t)$ are almost everywhere differentiable with respect to t. Let T_O be the set of types in $(0.5, 0.6)$ such that the obedient strategy is uniquely optimal amongst all pure strategies, i.e., $T_O := \{t \in (0.5, 0.6) \mid u^o(t) > u^d(t)\}$. Let T_D be the set of types in $(0.5, 0.6)$ such that the defiant strategy is uniquely optimal amongst all pure strategies, i.e., $T_D := \{t \in (0.5, 0.6) \mid u^o(t) < u^d(t)\}$. Let T_B be the set of types in $(0.5, 0.6)$ such that both strategies are equally optimal amongst all pure strategies, i.e., $T_B := \{t \in (0.5, 0.6) \mid u^o(t) = u^d(t)\}$. Based on the above argument that $\sigma \equiv 1$ and $\sigma \equiv 0$ are both suboptimal for any type in $(0.5, 0.6)$, thus $(0.5, 0.6) = T_O \cup T_D \cup T_B$.

Given that $\overline{\pi}$ is effectively equivalent to π, the optimal pure strategy under $\overline{\pi}$ must give each receiver type t the same payoff as that under π, therefore $u_o(t) = \frac{0.725+0.4t^2-1.3t}{2-t}$ for any $t \in T_O$, and similarly, $u_d(t) = \frac{0.725+0.4t^2-1.3t}{2-t}$ for any $t \in T_D$ and $u_o(t) = u_d(t) = \frac{0.725+0.4t^2-1.3t}{2-t}$ for any $t \in T_B$. However, the following Claim 2 shows that none of the resulting sets is of strictly positive measure, which leads to a contradiction. The rest is to prove the claim.

Claim 2. *The set* $\{t \in T_O | u_o(t) = \frac{0.725+0.4t^2-1.3t}{2-t}\}$, $\{t \in T_D | u_d(t) = \frac{0.725+0.4t^2-1.3t}{2-t}\}$ *and* $\{t \in T_B | u_o(t) = u_d(t) = \frac{0.725+0.4t^2-1.3t}{2-t}\}$ *are of zero measure.*

The proof of Claim 2 is again by contradiction. Consider the case that $\{t \in T_O | u_o(t) = \frac{0.725+0.4t^2-1.3t}{2-t}\}$ is of strictly positive measure (note that by definition T_O is an open set). Thus $\int_t^1 2\overline{\pi}(\omega)(\omega-t)\,\mathrm{d}\omega +$

$\int_0^t \overline{\pi}(\omega)(\omega - t)\,\mathrm{d}\omega = 0.725 + 0.4t^2 - 1.3t$, for any $t \in T_O$. By taking derivative w.r.t t twice, the following holds: $\overline{\pi}(t) = 0.8$, for any $t \in T_O$. This contradicts with the definition that $\overline{\pi}$ is deterministic, i.e., $\overline{\pi} \in \{0, 1\}$ for almost everywhere. For the second case that $\{t \in T_D | u_d(t) = \frac{0.725+0.4t^2-1.3t}{2-t}\}$ is of strictly positive measure, we can use a similar argument to derive $\overline{\pi}(t) = 0.2$ for every $t \in T_D$, which again leads to a contradiction. Since the above two sets are of zero measure, then we have $\{t \in T_B | u_o(t) = u_d(t) = \frac{0.725+0.4t^2-1.3t}{2-t}\}$ is only different from $(0.5, 0.6)$ by a zero measure set. The same argument again implies $\overline{\pi} \notin \{0, 1\}$ for almost everywhere $\omega \in T_B$, thus we again derive a contradiction. So we have shown that all of these sets cannot be of positive measure. □

References

Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag, Berlin, Heidelberg, 3rd edition, 2006. https://doi.org/10.1007/3-540-29587-9.

Itai Arieli, Yakov Babichenko, Rann Smorodinsky, and Takuro Yamashita. Optimal persuasion via bi-pooling. *Theoretical Economics*, forthcoming. https://doi.org/10.1145/3391403.3399468.

Robert J. Aumann. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, 55(1): 1–18, 1987. https://doi.org/10.2307/1911154.

Dirk Bergemann and Stephen Morris. Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11:487–522, 2016. https://doi.org/10.3982/TE1808.

Dirk Bergemann and Stephen Morris. Information design: A unified perspective. *Journal of Economic Literature*, 1:1–57, 2019. https://doi.org/10.1257/jel.20181489.

Alexander W. Bloedel and Ilya Segal. Persuading a rationally inattentive agent. *working paper*, 2019.

Yi-Chun Chen, Wei He, Jiangtao Li, and Yeneng Sun. Equivalence of Stochastic and Deterministic Mechanisms. *Econometrica*, 87(4):1367–1390, July 2019. https://doi.org/10.3982/ECTA14698.

Joseph L. Diestel and John J. Uhl. *Vector measures*. American Mathematical Society, Providence, R.I., 1977. http://dx.doi.org/10.1090/surv/015.

A. Dvoretzky, A. Wald, and J. Wolfowitz. Elimination of randomization in certain problems of statistics and of the theory of games. *Proceedings of the National Academy of Sciences of the United States of America*, 36:256–260, 1950. https://doi.org/10.1073/pnas.36.4.256.

A. Dvoretzky, A. Wald, and J. Wolfowitz. Elimination of Randomization in Certain Statistical Decision Procedures and Zero-Sum Two-Person Games. *The Annals of Mathematical Statistics*, 22(1):1 – 21, 1951. https://doi.org/10.1214/aoms/1177729689.

Piotr Dworczak and Anton Kolotilin. The persuasion duality. *working paper*, 2022. https://arxiv.org/abs/1910.11392.

Piotr Dworczak and Giorgio Martini. The simple economics of optimal persuasion. *Journal of political economy*, 127(5), Oct 2019. https://doi.org/10.1086/701813.

Matthew Gentzkow and Emir Kamenica. A Rothschild-Stiglitz Approach to Bayesian Persuasion. *American Economic Review, Papers & Proceedings*, 106(5):597–601, 2016. https://doi.org/10.1257/aer.p20161049.

Matthew Gentzkow and Emir Kamenica. Bayesian persuasion with multiple senders and rich signal spaces. *Games and Economic Behavior*, 104(5):411–429, 2017. https://doi.org/10.1016/j.geb.2017.05.004.

Yingni Guo and Eran Shmaya. The interval structure of optimal disclosure. *Econometrica*, 87(2):653–675, 2019. https://doi.org/10.3982/ECTA15668.

J.C. Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *Int J Game Theory*, 2:1 – 23, 1973. https://doi.org/10.1007/BF01737554.

Sergiu Hart and Philip J. Reny. Maximal revenue with multiple goods: nonmonotonicity and other observations. *Theoretical Economics*, 10(3), September 2015. https://doi.org/ 10.3982/TE1517.

Wei He and Yeneng Sun. Pure-strategy equilibria in bayesian games. *Journal of Economic Theory*, 180: 11–49, 2019. https://doi.org/10.1016/j.jet.2018.11.007.

Roy A. Johnson. Atomic and nonatomic measures. *Proc. Amer. Math. Soc.*, 25:650–655, 1970. https://doi.org/10.2307/2036664.

Emir Kamenica and Matthew Gentzkow. Bayesian Persuasion. *American Economic Review*, 101(6): 2590–2615, 2011. https://doi.org/10.1257/aer.101.6.2590.

M. Ali Khan and Kali P. Rath. On games with incomplete information and the dvoretsky-wald-wolfowitz theorem with countable partitions. *Journal of Mathematical Economics*, 45(12):830–837, 2009. https://doi.org/10.1016/j.jmateco.2009.06.003.

M. Ali Khan, Kali P. Rath, and Yeneng Sun. The dvoretzky–wald–wolfowitz theorem and purification in atomless finite-action games. *Int J Game Theory*, 34(91), 2006. https://doi.org/10.1007/s00182-005-0004-3.

M.Ali Khan and Yeneng Sun. Pure strategies in games with private information. *Journal of Mathematical Economics*, 24(7):633–653, 1995. https://doi.org/10.1016/0304-4068(94)00708-I.

M.Ali Khan and Yongchao Zhang. On the existence of pure-strategy equilibria in games with private information: A complete characterization. *Journal of Mathematical Economics*, 50:197–202, 2014. https://doi.org/10.1016/j.jmateco.2013.12.005.

Andreas Kleiner, Benny Moldovanu, and Philipp Strack. Extreme points and majorization: Economic applications. *Econometrica*, 89(4):1557–1593, 2021. https://doi.org/10.3982/ECTA18312.

Achim Klenke. *Probability theory*. Springer, London, 2014. https://doi.org/10.1007/978-1-4471-5361-0.

Anton Kolotilin. Optimal information disclosure: A linear programming approach. *Theoretical Economics*, 13(2):607–635, 5 2018. https://doi.org/10.3982/TE1805.

Anton Kolotilin, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li. Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964, 11 2017. https://doi.org/10.3982/ECTA13251.

Heng Liu. Efficient dynamic mechanisms in environments with interdependent valuations: The role of contingent transfers. *Theoretical Economics*, 13(2):795–829, 2018. https://doi.org/10.3982/TE2234.

Alejandro M. Manelli and Daniel R. Vincent. Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 127(1):1–35, March 2006. https://doi.org/10.1016/j.jet.2005.08.007.

Alejandro M. Manelli and Daniel R. Vincent. Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic Theory*, 137(1):153–185, November 2007. https://doi.org/10.1016/j.jet.2006.12.007.

R. Preston McAfee and John McMillan. Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory*, 46(2):335–354, December 1988. https://doi.org/10.1016/0022-0531(88)90135-4.

Jean-François Mertens. A measurable "measurable choice" theorem. *In: Neyman A., Sorin S. (eds) Stochastic Games and Applications. NATO Science Series, (Series C: Mathematical and Physical Sciences)*, 507:107–130, 2003. https://doi.org/10.1007/978-94-010-0189-2_9.

Paul R. Milgrom and Robert J. Weber. Distributional strategies for games with incomplete information. *Mathematics of Operations Research*, 10(4):619–632, 1985. https://doi.org/10.1287/moor.10.4.619.

Roger Myerson. Optimal coordination mechanisms in generalized principal-agent problems. *Journal of Mathematical Economics*, 10:67–81, 1982. https://doi.org/10.1016/0304-4068(82)90006-4.

Roger B. Myerson. Multistage games with communication. *Econometrica*, 54(2), 1986. https://doi.org/10.2307/1913154.

Konrad Podczeck. On purification of measure-valued maps. *Economic theory*, 38:399–418, 2009. https://doi.org/10.1007/s00199-007-0319-3.

Roy Radner and Robert W. Rosenthal. Private information and pure-strategy equilibria. *Mathematics of Operations Research*, 7(3):401–409, 1982. https://doi.org/10.1287/moor.7.3.401.

Luis Rayo and Ilya Segal. Optimal information disclosure. Journal of Political Economy, 118(5), 2010. https://doi.org/10.1086/657922.

Roland Strausz. Deterministic versus stochastic mechanisms in principal-agent models. *Journal of Economic Theory*, 128(1):306–314, May 2006. https://doi.org/10.1016/j.jet.2004.11.008.

Dong Wei and Brett Green. (Reverse) price discrimination with information design. Working Paper (December 2019), 2019.