Sequential Persuasion*

Fei Li[†] Peter Norman [‡]

March 4, 2018

Abstract

This paper studies sequential move persuasion games with multiple senders. We use convex analysis to transform a problem with infinite action spaces to a finite action model. This way we prove the existence of equilibria by the Zermelo-Kuhn backward induction algorithm, show that equilibrium outcomes are generically unique, and obtain a simple algorithm for finding equilibrium outcomes. We also obtain a simple condition for when full revelation is the unique equilibrium outcome and some comparative statics results. Adding a sender who moves first cannot reduce informativeness in equilibrium, and will result in a more informative equilibrium in the case with two states. Sequential persuasion cannot generate a more informative equilibrium than simultaneous persuasion and is always less informative when there are only two states.

Keywords: Communication, Bayesian Persuasion, Multiple Senders, Sequential Persuasion.

JEL Classification Codes: D82, D83

*We thank Yu Awaya, Gary Biglaiser, Navin Kartik, Keiichi Kawai, Kyungmin (Teddy) Kim, Elliot Lipnowski, Shuo Liu, Giuseppe Lopomo, Laurent Mathevet, Alexey Kushnir, Luca Rigotti, Joel Sobel, and participants at the 2017 CETC conference in Vancouver, the 2017 International Conference on Game Theory, 2017 Midwest Economic Theory Conference in Dallas for helpful comments. The usual disclaimer applies.

[†]University of North Carolina at Chapel Hill. lifei@email.unc.edu

[‡]*Corresponding Author.* University of North Carolina at Chapel Hill. normanp@email.unc.edu

1 Introduction

A lobbyist trying to influence legislators can use at least two different instruments. One can pay for votes, try to persuade the legislators that the policy advocated by the lobbyist is the right thing to do, or use a combination of money and persuasion to influence legislative outcomes. Accordingly, when there are different lobbyists representing different interests there is both price competition and competition in persuasion shaping eventual outcomes. There is an extensive literature studying how price competition influences outcomes of the political process, much of it based on contest theory. In contrast, competition in persuasion is not as well understood.

There are numerous real world examples of competition in terms of information structures beyond lobbying. Clinical trials by pharmaceutical companies, software companies competing by releasing free beta versions of their products, doctors and laboratories providing second medical opinions, home sellers using multiple realtors etc. Businesses specializing in selling second opinion services are common in medicine, finance, marketing, and law (see Sarvary (2002) and the references therein).

A few recent papers, such as Gentzkow and Kamenica (2017a,b) and Board and Lu (2017), have studied competition in persuasion using multi-sender extensions of *Bayesian persuasion* games. However, no paper in the existing literature provides a general treatment of how to analyze persuasion games in which the agents that seeks to influence the decision move sequentially. The purpose of this paper is to close this gap.

Many real world examples with competing experts are sequential in nature. The provider of the second opinion usually knows that the initial opinion triggered a second opinion, and often also the exact nature of the initial advice. Medical second opinions tend to be sequential when triggered by an internal rule or by an insurance provider. A realtor that is providing a secondary appraisal is at least aware of the existence of another appraisal. In competition between different news organizations, players do respond to information provided by other outlets.

Additionally, we like to understand how simultaneous move persuasion compares with sequential persuasion for many of the same reasons that we like to understand the difference between the Cournot and Stackelberg models: this touches on central issues such as the value of commitment and is also a necessary first step to understand how to design games with competing experts.

We consider a sequential model of information disclosure building on the recent advances in multi-sender *Bayesian persuasion* games in Gentzkow and Kamenica (2017a,b). An uninformed decision maker seeks to maximize her state dependent payoff. There is also a number of biased senders who move in sequence, each constructing an *experiment* or *signal* with a precision ranging from no information to full revelation of the state. Each sender observes the experiments designed by previous senders before constructing a new experiment.

The model is a finite horizon game of perfect information, but each player has an infinite action space, so off-the-shelf existence results don't apply. In a single-sender model, Kamenica and Gentzkow (2011) show that the sender's problem is reduced to splitting the decision maker's belief about the state subject to Bayes' rule. Existence of equilibria can be proven by having the decision maker breaking ties in favor of the sender, making the latter's payoff upper semi-continuous. In the multi-sender framework, there is no reason for different senders to agree on how to break ties, so extending the proof strategy from Kamenica and Gentzkow (2011) to the case with multiple senders does not seem easy.

However, the upper semi-continuity condition is sufficient but not necessary for equilibrium existence. We instead use the convexity of the model to guarantee existence. Using linearity in probabilities, the optimal choice rule by the decision maker can be characterized in terms of intersections of upper half spaces, or convex polytopes. Actions are constant in the interior of each polytope, so it is possible to replicate any outcome with an interior belief in some polytope with one where all probability is assigned to the vertices of the polytopes, provided that an appropriate tie breaking rule is used. Hence, the problem for the final sender is equivalent with a problem in which beliefs are restricted to belong to vertices of the polytopes, a finite problem.[1] After imposing an appropriate selection rule when there are multiple best responses, existence of a (Markov perfect) equilibrium is established by backward induction, where in each step senders without loss restrict attention to beliefs that are vertices of the convex polytopes that define the optimal actions for the decision maker.

The reader may notice that the existence question does not arise in the simultaneous move models considered in Gentzkow and Kamenica (2017a,b). If at least two senders play fully revealing signals, the information structure available for the decision maker is fully revealing regardless of any unilateral deviation by any player. Hence, no player is pivotal, so a fully revealing equilibrium exists even if this would be the worst possible outcome for every sender. In the sequential move case, all senders understand that they have influence unless a fully revealing signal has

[1] Lipnowski and Mathevet (2017) use this insight to simplify the *concavification* in a single-sender Bayesian persuasion game.

already been played. It follows that full revelation is not always an equilibrium, so equilibrium existence is not automatic.

Surprisingly, it is irrelevant whether or not senders observe the *outcomes* of the existing experiments, which is a consequence of signals being *arbitrarily correlated*, so that each sender can refine the pre-existing information structure *realization by realization*. Moreover, if realizations are observed, it doesn't matter whether signals are independent or coordinated. Hence, in a sequential framework we may reinterpret the convenient technical assumption of arbitrarily correlated signals as an assumption of publicly observable experimental results.

Being able to restrict attention on beliefs that are vertices defining the decision maker's "optimal action areas" drastically simplifies the analysis of the model. Additionally, we obtain a further simplification by showing that every subgame perfect equilibrium outcome can be supported using *one step equilibrium* strategies. In one step equilibrium, the first sender is the only sender that provides any non-trivial information. Every sender can always unilaterally change the information system implied by the previous senders into any Blackwell-more-informative signal. If we start with an arbitrary equilibrium and change it so that sender 1 plays the joint signal implied by the individual signals by senders $1, ..., n$, assume that no sender provides any additional information on the path, and assume that every sender plays in accordance with the initial equilibrium for any other history, this must also be an equilibrium. All senders earn exactly the same payoff as in the initial equilibrium, and any sender with an incentive to deviate in the one step profile would had an incentive to deviate under the initial profile as well. This applies on and off the equilibrium path, so any subgame perfect equilibrium can be replicated by strategies that are one step after any history of play.

Combining the one step equilibrium characterization with the insight that we without loss may restrict attention to signals with support on the finite set of vertex beliefs we next obtain a generic *essential uniqueness* result. The reason for the qualifier *essential* is that it is possible that all players can be indifferent across a continuum of distributions in the interior of a polytope that defines optimality for the decision maker, but such multiplicity is irrelevant as the joint distribution over states and actions is the same for all such equilibria. Hence, we focus on equilibria in which beliefs are on the (finite set of) vertices of the decision maker's choice areas. Essential uniqueness can fail for at most a set of sender payoff profiles of measure zero.

The generic uniqueness result is another consequence of the convexity of the model, combined with the fact that any *equilibrium outcome* can be replicated by a *one step equilibrium*. In a one step equilibrium, the *equilibrium path* is a distribution of beliefs such that no sender has an incentive to refine any belief in the support of this distribution. That is, any belief in the support of the equilibrium must be *stable*. We also know that the distribution has finite support without loss of generality, and, since the state space is finite, the convex hull of the beliefs in the equilibrium distribution is finite dimensional. A failure of uniqueness would imply that there are multiple convex combinations over the finite support of stationary beliefs that give some sender the same utility. But, by Carathéodory's theorem, every point in the convex hull can be spanned by $M+1$ vertices of the convex hull, where M is the dimensionality of the set. Hence, if there are more than $M+1$ vertices of the convex hull at least one can be spanned by the others. Indifference requires that a sender gets the same utility if moving for sure to the vector that is spanned by the others as if getting a lottery with probabilities defined by the spanning weights over the spanning vector. Such indifference holds for utility functions of measure zero. Repeating the argument for all possible combinations of vertices and all senders we have a finite set of possibilities for indifference, and it follows that equilibria are essentially unique for almost all sender payoff functions.

Converting the problem to a problem over a finite set of vertices is also helpful computationally. To illustrate this we consider a simple criminal trial example with three states and three actions where, because we can restrict attention to equilibria on a few vertices, it is very simple to find the equilibrium. The example is also of some interest because the unique equilibrium is fully revealing despite both senders being better off providing no information at all, so there is a flavor of a prisoner's dilemma. Despite its simplicity, the example suggests that one may be able to use the methodology of our paper to find interesting classes of "opposing" preferences where full revelation is the only equilibrium. Indeed, we provide a condition for when full revelation is the unique equilibrium that is easy to check and weaker than "zero sum conditions" suggested by the previous literature.

Moving on to comparative statics, we ask what happens if the number of senders is increased. We know from Li and Norman (2017) that this can result in a loss of information if an extra sender is added at the end or in the middle of the order. In contrast, if an additional sender is added at the root of the game we show that such loss is impossible. Either, the new (essentially unique) equilibrium is more informative in Blackwell's order or the equilibria cannot be compared. While it is unfortunate that one may not always be able to rank equilibria by informativeness, this is a consequence of the incompleteness of the Blackwell ordering that simply cannot be avoided in general. However, when we focus on models with two states, we have stronger results: after adding new senders at the root of the game, the new equilibrium is always more informative. We also demonstrate that sequential persuasion results in equilibria that are no more informative than simultaneous persuasion. That is, either the equilibria cannot be compared or simultaneous persuasion is more informative in Blackwell's order. With two states sequential persuasion is always less informative.

Our paper relates to a large body of work on information disclosure, but is most directly connected with the growing literature on Bayesian persuasion started by Kamenica and Gentzkow (2011) and Rayo and Segal (2010). This literature has recently been extended to incorporate multiple senders (see Gentzkow and Kamenica (2017a,b), Boleslavsky and Cotton (2016), Au and Kawai (2017a,b), and Hoffmann, Inderst, and Ottaviani (2014)), but none of these papers deal with sequential moves by the senders. There is also a growing body of work that embeds persuasion into dynamic models (see Ely, Frankel, and Kamenica (2015) and Ely (2017)), but the paper that is closest in spirit to ours is Board and Lu (2017), who incorporate Bayesian persuasion into a search model.[2] However, Board and Lu (2017) consider payoff functions that are more restrictive than in this paper, and the decision maker faces an optimal stopping problem. In contrast, the decision maker has no influence on the precision of her information in our model. Finally, our equilibrium characterization crucially depends on the finiteness of the game. See Wu (2017) for a discussion of infinite action spaces. Our formal analysis has some similarities with Lipnowski and Mathevet (2017), but their focus is to simplify the analysis of single-sender persuasion games.

The remainder of the paper is organized as follow. In section 2, we describe the model setup. Section 3 characterize the set of equilibria. We show that a Markov perfect equilibrium exists and that the equilibrium outcome is generically unique. Section 4 illustrates that our characterization is helpful for computing equilibria within a simple example. In section 5, we perform some comparative static. In appendix A, we present a non-generic example where a non-Markov equilibrium exists. All omitted proofs are in appendix B.

[2] Glazer and Rubinstein (2001) studies a finite horizon sequential persuasion model, but they consider a very different information structure. There are also papers in the cheap talk and disclosure literature that asks what the implications of multiple senders are. See Ambrus and Takahashi (2008), Battaglini (2002), Kawai (2015), Krishna and Morgan (2001), Kartik, Lee, and Suen (2017, 2016), Bhattacharya and Mukherjee (2013) and Milgrom and Roberts (1986).

2 Model

Consider an environment with $n \geq 1$ senders and one decision maker. Sender $i \in \{1,...,n\}$ has payoff function $u_i : A \times \Omega \to \mathbb{R}$ where A is a finite set of actions and Ω is the finite state space. The decision maker has a payoff function $u_D : A \times \Omega \to \mathbb{R}$ and there is a common prior belief $\mu_0 \in \Delta(\Omega)$. Payoff functions are common knowledge and players evaluate lotteries using expected utilities. Fix a belief μ and an action a, define player i's expected payoff as

$$v_i(a,\mu) \equiv \sum_{\omega\in\Omega} u_i(a,\omega)\mu(\omega), \tag{1}$$

for $i = 1,...,n,D$. We make a regularity assumption that guarantees that for every pair of actions there is some state that makes the decision maker strictly better off by taking the first action than the second and another state that reverses the rank of the payoffs. This assumption implies that the set of beliefs making the decision maker indifferent between any two actions, the set $\{\mu \in \Delta(\Omega) | v_D(a,\mu) = v_D(a',\mu)\}$, is the intersection between $\Delta(\Omega)$ and a hyperplane.

Assumption 1. *For each ordered pair of actions (a,a') there exists a state $\omega \in \Omega$ such that $u_D(a,\omega) > u_D(a',\omega)$.*

Players are uninformed about the state of the world, but the senders may provide information to the decision maker by creating *signals*. As in Gentzkow and Kamenica (2017a) we model signals using the *partition representation* introduced by McGuire (1959) and Marschak and Miyasawa (1968). A generic signal thus is a finite Lebesgue measurable partition of $\Omega \times [0,1]$.

Given *signal* π we assign probabilities as if a sunspot variable Z is drawn uniformly from $[0,1]$: the probability of signal $s \in \pi$ conditional on ω is $p(s|\omega) = \int_{z\in I_\omega(s)} dz$ where a generic signal realization is on form $s = \cup_\omega [\omega \times I_\omega(s)]$ and for each state ω the collection $\{I_\omega(s)\}_{s\in\pi}$ are disjoint sets such that $\cup_{s\in\pi} I_\omega(s) = [0,1]$. The posterior probability of state $\omega \in \Omega$ is thus

$$\mu(\omega|s) = \frac{p(s|\omega)\mu_0(\omega)}{\sum p(s|\omega')\mu_0(\omega')} = \frac{\left[\int_{z\in I_\omega(s)} dz\right]\mu_0(\omega)}{\sum_{\omega'\in\Omega}\left[\int_{z\in I_{\omega'}(s)} dz\right]\mu_0(\omega')}. \tag{2}$$

Denoting the unconditional probability of s by $p(s) = \sum_{\omega'\in\Omega} p(s|\omega')\mu_0(\omega')$ and noting that $\sum_{s\in\pi} p(s|\omega) = 1$ for each $\omega \in \Omega$ we note that every signal induces a distribution of posterior beliefs that satisfies the *Bayes plausibility* constraint

$$\sum_{s\in\pi} \mu(\omega|s)\, p(s) = \mu_0(\omega). \tag{3}$$

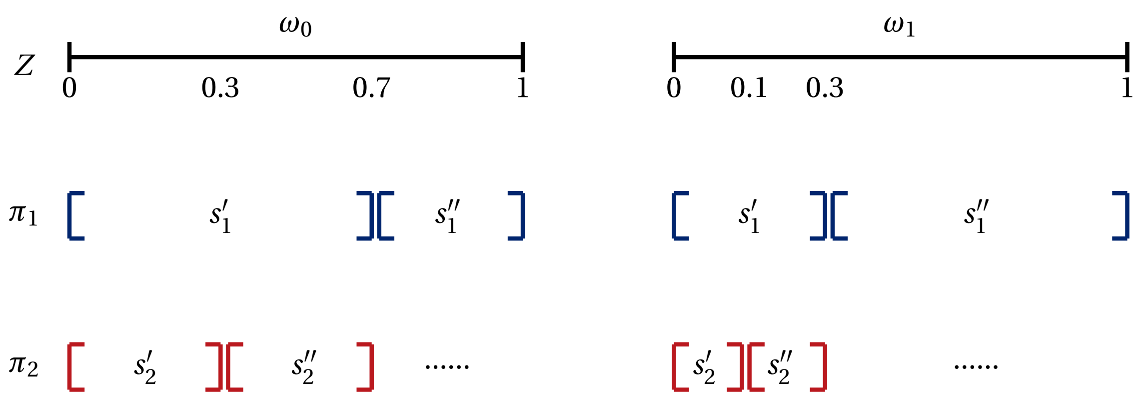

Figure 1: There are two states: ω_0 and ω_1. Consider two signals: $\pi_1 = \{s_1', s_1''\}$ and $\pi_2 = \{s_2'.s_2'', s_1''\}$. Signal π_1 has two possible realizations: $s_1' = \{(\omega_0, [0, 0.7]), (\omega_1, [0, 0.3])\}$ and $s_1'' = \{(\omega_0, (0.7, 1]), (\omega_1, (0.3, 1]),$ and signal π_2 has three possible realizations $\{s_2', s_2'', s_1''\}$. Suppose that the prior belief is $\mu_0 = 0.5$. If s_1' is realized, $(\omega, Z) \in (\omega_0, [0, 0.7])$ or $(\omega, Z) \in (\omega_1, [0, 0.3])$ with equal probability; and thus $\mu(s_1') = 0.3$ by (2). Since $s_1' = s_2' \cup s_2''$, π_2 is a finer partition than π_1.

We let Π denote the set of signals. Signals are partially ordered as one can partition each set in the partition:

Definition 1. *π is a finer partition than π', denoted $\pi \succeq \pi'$, if for every $s \in \pi$ there exists $s' \in \pi'$ such that $s \subset s'$.*

The space $(\Pi, \succeq)$ is a lattice and for any two signals π, π', the join $\pi \vee \pi'$ is the signal that consists of realization $s \cap s'$ for each $s \in \pi$ and $s' \in \pi'$. Notice that $\pi \vee \pi'$ is finer than both π and π' for any pair of signals (π, π'), so combining two signals by taking intersections generates a more informative new signal. See Figure 1 for a simple illustration of the information structure.

Senders move sequentially with sender $1, \ldots, n$ posting signals $\pi_1 \ldots., \pi_n$, and each sender observing previous senders' signals. Then nature draws ω and the sunspot variable. Finally, the decision maker observes $(\pi_1, \ldots., \pi_n)$ and a joint realization $\cap_{i=1}^n s_i$ with $s_i \in \pi_i$ for each $i \in \{1, \ldots, n\}$ and takes an action $a \in A$.[3] Allowing senders to move once only is without loss of generality because a sender moving multiple times can be replaced with multiple senders sharing the same payoff function

A pure strategy for sender i is a map $\sigma_i : \Pi^{i-1} \to \Pi$ where Π^0 is defined as the trivial null history. A generic history for the decision maker is a vector $(\pi_1, .., \pi_n, s)$ with $s = \cap_{i=1}^n s_i$ and $s_i \in \pi_i$ for each

[3]While drawing the state in the end or at the beginning doesn't affect equilibrium outcomes, we need to draw it at the end for subgame perfection to formally apply.

i. Let $\mathscr{H}_D$ be the set of all histories of the decision maker, and $\sigma_D : \mathscr{H}_D \to A$ denote his strategy. There is uncertainty about the state, but information is symmetric, and there is therefore never any point in the game where any player needs to update the beliefs about the type of other players. Hence, *subgame perfection* is applicable.

3 Results

We begin this section with an example that illustrates our approach of focusing on equilibria with support on a finite set of posterior beliefs for the decision maker. Unlike simultaneous move persuasion games, existence of a fully revealing equilibrium is not automatic, so our first general result is existence of a subgame perfect equilibrium. The proof uses convex analysis, the Zermelo-Kuhn algorithm, and a selection rule of best responses such that the best response problem for each sender can be transformed into a finite linear program.

As a by-product of our existence proof we also obtain a reinterpretation of fully coordinated signals for sequential persuasion games: such equilibria are also equilibria in a model in which senders post plain vanilla independent signals, but where senders observe the realizations of previous experiments.

We also show that the subgame perfect equilibrium distribution over actions and states is unique for almost all Bernoulli utility functions. This has several nice secondary implications. In particular, we may almost always without loss of generality focus on Markov equilibria, which allows us to construct a simple (finite) algorithm to characterize equilibrium outcomes. Additionally, a unique subgame perfect equilibrium rules out mixed-strategies. As allowing for mixed strategies makes a qualitative difference in simultaneous persuasion games this is an important implication of uniqueness.[4]

3.1 An Illustrative Example

To illustrate some of the challenges to be overcome in establishing our characterization results we will first discuss an example.[5] The example relies on a payoff function for one of the players which is non-generic in the sense that the set of payoff functions that can create the difficulties to be detailed below have Lebesgue measure zero, but it highlights how our approach to establishing

[4] See the discussion in Gentzkow and Kamenica (2017a,b) and the example in Li and Norman (2017).

[5] We are grateful to an anonymous referee for suggesting this example.

existence and (generic) essential uniqueness of equilibria differs from more standard continuity arguments.

We consider a persuasion game with two senders, two states, ω_0 and ω_1, and action space $A = \{a_0, a_1, a_2\}$. The decision maker has expected utility preferences with payoffs over states and actions given by.

$$\begin{array}{c|ccc} & a_0 & a_1 & a_2 \\ \hline \omega_0 & 1 & 0 & -1 \\ \omega_1 & -2 & 0 & 1/2, \end{array} \tag{4}$$

and the senders have state independent payoffs given by

$$u_1(a,\omega) = \begin{cases} 1 & \text{if } a = a_1 \\ 0 & \text{if } a \neq a_1 \end{cases}, \text{ and } u_2(a,\omega) = \begin{cases} 0 & \text{if } a = a_0 \\ 1 & \text{if } a = a_1 \\ 2 & \text{if } a = a_2 \end{cases}, \tag{5}$$

for $\omega \in \{\omega_0, \omega_1\}$.

First, sender 1 posts an information structure $\pi_1 \in \Pi$. Sender 2 observes π_1 and chooses $\pi_2 \in \Pi$, and then the decision maker observes the signals (π_1, π_2) and a realization s of the joint signal $\pi_1 \vee \pi_2$. We let μ denote a generic posterior probability that the state is ω_0.

In Section 3.2, we argue that the posterior belief μ is the only payoff relevant variable for the decision maker. It is easy to check that a_0 is the unique optimal action for the decision maker for $\mu < 1/3$, that a_1 is the unique optimal action on $(1/3, 2/3)$, and that a_2 is the unique optimal action given $\mu > 2/3$. At $\mu = 1/3$ the decision maker is indifferent between a_0 and a_1, and at $\mu = 2/3$, there is indifference between a_1 and a_2. To ensure that sender 2 has a well defined best response problem we break these ties in favor of sender 2, and assume that the decision maker follows

$$\sigma_D(\mu) = \begin{cases} a_0 & \text{if } 0 \le \mu < \frac{1}{3} \\ a_1 & \text{if } \frac{1}{3} \le \mu < \frac{2}{3} \\ a_2 & \text{if } \frac{2}{3} \le \mu \le 1 \end{cases}. \tag{6}$$

We next consider subgames when sender 2 moves. Sender 2 then observes some π_1, a signal played by sender 1, and each potential realization $s_1 \in \pi_1$ corresponds to an "interim belief" $\mu(s_1)$. Because payoffs are linear in probabilities we may separate the best response problem for sender 2 into $|\pi_1|$ distinct problems, one for each signal realization/interim belief generated by π_1.[6] Each such problem is equivalent to a single sender persuasion game in the framework of Kamenica and

[6] See the discussion on how to solve (16) by solving several problems on form (17) in Section 3.2

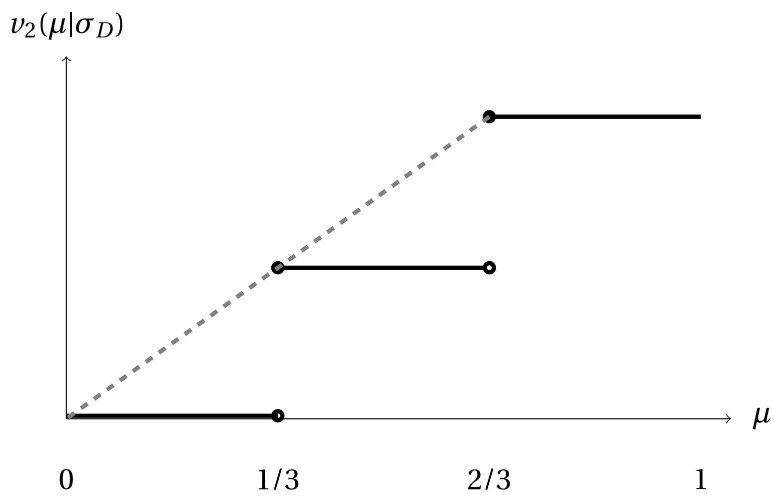

Figure 2: The solid line represents sender 2's payoff in the absence of providing further information, and the dashed line represents its concave closure.

Gentzkow (2011). One approach to the best response problem for sender 2 is therefore to work in belief space and find an optimal mean preserving spread of any possible interim belief. This is how we will proceed from here.

The reason that this example is pathological is that if sender 2 faces interim belief $\mu = 1/3$ then a feasible mean preserving spread is sending the posterior to 0 with probability 1/2 and to 2/3 with probability 1/2, and since

$$u_2(a_1;\omega) = \frac{1}{2}u_2(a_0;\omega) + \frac{1}{2}u_2(a_2;\omega). \tag{7}$$

this implies that any mean preserving spread over $\{0, 1/3, 2/3\}$ is optimal when $0 < \mu < 2/3$, so sender 2 has a continuum of best responses. To understand this geometrically, see Figure 2 which plots sender 2's payoff as a function of the interim belief of the decision maker.

One best response for sender 2 is

$$\tau_2(\mu) = \begin{cases} \text{the unique mean preserving spread onto } \left\{0, \frac{1}{3}\right\} & \text{if } \mu \leq \frac{1}{3} \\ \text{the unique mean preserving spread onto } \left\{\frac{1}{3}, \frac{2}{3}\right\} & \text{if } \frac{1}{3} < \mu \leq \frac{2}{3} \\ \text{the unique mean preserving spread onto } \left\{\frac{2}{3}, 1\right\} & \text{if } \mu > \frac{2}{3} \end{cases}. \tag{8}$$

Given that sender 2 uses the best response in (8) we may plot the payoff of sender 1 as a function of the interim belief for sender 2 as the solid lines in in the upper panel of Figure 3. When the prior belief $\mu_0 \in [0, 1/3]$ sender 1 may therefore create any mean preserving spread with all beliefs in the support being in the interval $[0, 1/3]$. Given the response in (8) any such mean preserving spread lead to the same ultimate outcome. In contrast, when $\mu_0 > 1/3$ sender 1 will respond by generating the unique mean preserving spread with support $\{1/3, 1\}$. Hence, we have constructed an equilibrium.

Another best response is

$$\tau_2'(\mu) = \begin{cases} \text{the unique mean preserving spread onto } \left\{0, \frac{2}{3}\right\} & \text{if } 0 \le \mu \le \frac{2}{3} \\ \text{the unique mean preserving spread onto } \left\{\frac{2}{3}, 1\right\} & \text{if } \mu > \frac{2}{3} \end{cases}. \tag{9}$$

Given the continuation play generated by σ_D and τ_2' there is nothing sender 1 can do to make the decision maker take action a_1 in equilibrium. Hence, anything is a best response for player 1 and we have constructed a family of subgame perfect equilibria with an outcome corresponding to belief distributions with support on $\{0, 2/3\}$ or $\{0, 2/3, 1\}$ depending on whether the prior is below or above $2/3$.

Both the equilibrium with sender 2 response (8) and the one with sender 2 response (9) led to multiplicity in what actions sender 1 takes. This kind of multiplicity is generic, but also irrelevant for the distribution over actions and states. However, the equilibria with sender 2 response (8) generate a different distribution over actions and states than and the ones with sender 2 response (9), and this type of multiplicity matters, so we call it *essential*.

We relied on a payoff function satisfying the equality in (7) to generate this multiplicity, and the set of payoff functions satisfying this equality defines a plane in $\mathbb{R}^3$, which is measure zero. In our general analysis we allow for richer action and state spaces and state dependent payoffs, but we nevertheless establish essential uniqueness for almost all payoff functions by ruling out more general versions of linear dependencies.

Finally, consider the best response

$$\tau_2''(\mu) = \begin{cases} \text{the unique mean preserving spread onto } \left\{0, \frac{1}{3}\right\} & \text{if } \mu < \frac{1}{3} \\ \text{the unique mean preserving spread onto } \left\{0, \frac{2}{3}\right\} & \text{if } \mu = \frac{1}{3} \\ \text{the unique mean preserving spread onto } \left\{\frac{1}{3}, \frac{2}{3}\right\} & \text{if } \frac{1}{3} < \mu \le \frac{2}{3} \\ \text{the unique mean preserving spread onto } \left\{\frac{2}{3}, 1\right\} & \text{if } \mu > \frac{2}{3} \end{cases}. \tag{10}$$

Sender 1 payoff as a function of sender 2 beliefs corresponding to the best response (10) is plotted in the lower panel of Figure 3. Clearly, the best reply to σ_D and τ_2'' is ill-defined because of the discontinuous best reply at $\mu = 1/3$, which illustrates that to establish existence of equilibrium we must be careful about which best replies to select when there is multiplicity.

To summarize, the main points of the example are as follows. Firstly, it is without loss of generality to focus on equilibria with support on a finite set of decision maker beliefs. In the example,

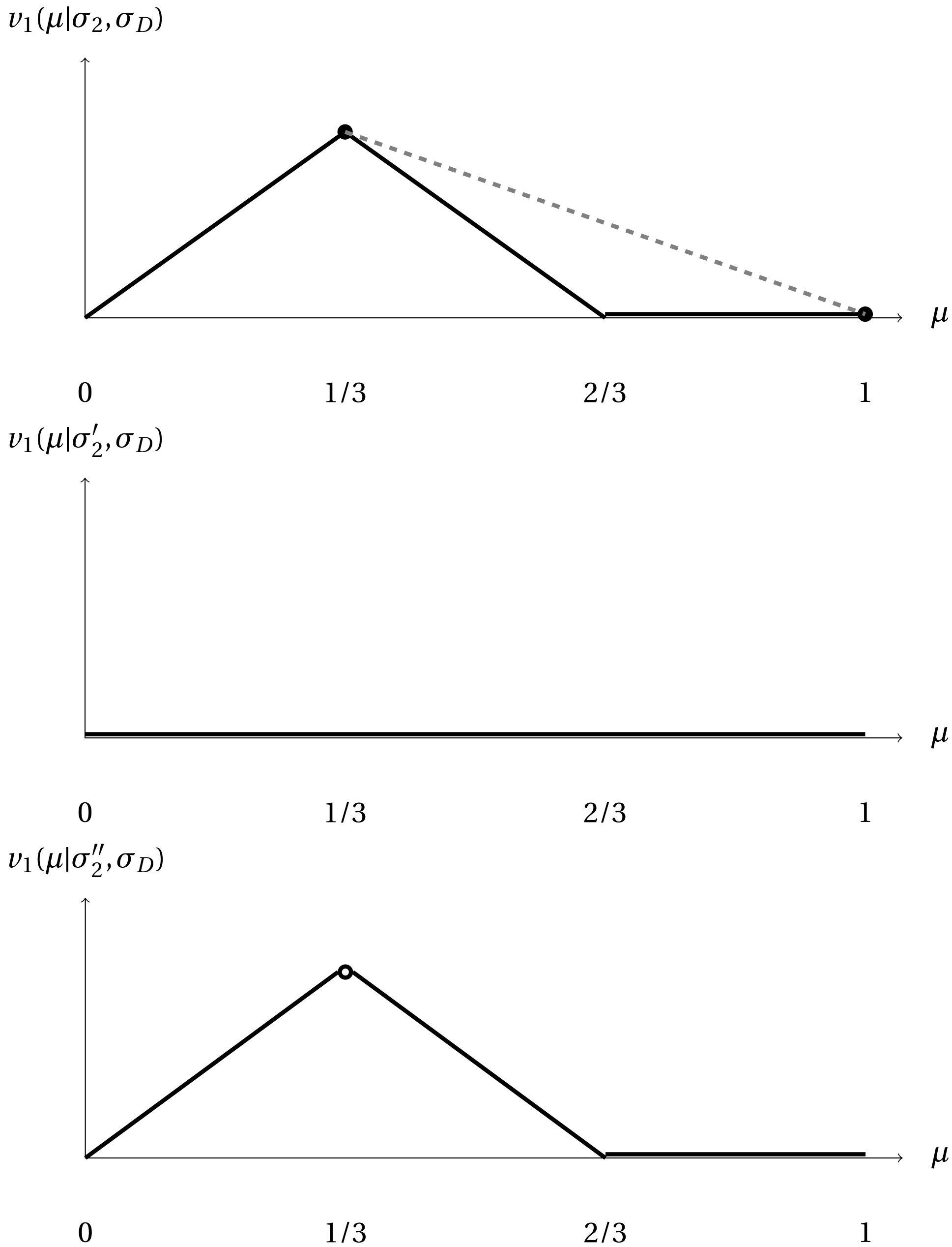

Figure 3: Sender 1's payoff as a function of the prior belief given three distinct best responses by sender 2.

the appropriate set are points of indifference for the decision maker. This generalizes to points spanning the convex hull of beliefs that make a certain action optimal in the general model.

Secondly, the example illustrates that we need to be careful about how to select best responses for a sender in case there is *essential multiplicity* in the sense that the distribution over states and actions is affected by the selection. In the example this problem occurs when the second sender is indifferent between a point of indifference for the decision maker and a mean preserving spread over two other points of indifference. This implies that there are best responses for the second sender that makes the first senders problem ill defined. In the example, this problem is avoided by selecting a best response so that any belief in an interval is refined into the unique mean preserving spread to the same two points. We use convex analysis to demonstrate that this idea generalizes to the case with arbitrary finite state and action spaces. This allows us to avoid using any explicit references to continuity arguments because we construct best responses that transform the best response problems for previous senders into finite action optimization problems.

Finally, with a preference profile that is very special, we could generate multiplicity of equilibria in terms of the distribution over states and actions. It is not hard to see that if the preferences of sender 2 are perturbed in any direction, then best responses over points of indifference for the decision maker are unique. This also generalizes. For the general model we demonstrate that multiplicity of equilibria can only happen for a zero Lebesgue measure set of payoff functions.

3.2 Existence of Equilibria

In this section we will demonstrate:

Theorem 1. *There exists a subgame perfect equilibrium.*

We establish existence by showing that a *separable Markov* equilibrium exists. We say that the decision maker's strategy is *Markov* if the action only depends on the *posterior belief* induced by the signals and the realization, so, abusing notation, we write $\sigma_D : \Delta(\Omega) \to A$ for a Markov strategy for the decision maker. This is in general restrictive, and there is an example in Appendix A where the decision maker uses a non-Markov decision rule as carrots and sticks, which supports a qualitatively different outcome than any Markov equilibrium.[7]

[7]However, we show that subgame perfect equilibria are essentially unique and when it is unique, it is Markov, so such examples are pathological.

Any history of individual signals $h_i = \{\pi_j\}_{j=1}^{i-1}$ induces a joint signal $\pi^{i-1} = \vee_{j=1}^{i-1}\pi_j$, which is what would be available for the decision maker if senders $i, ..., n$ do not add any further information. For each $s \in \pi^{i-1}$, one can calculate the hypothetical posterior belief $\mu(s)$ by (3), which can be interpreted as the decision maker's posterior belief if no sender adds any signal in the continuation game and s is realized. We call such belief an *interim belief.* Each joint signal $\pi^{i-1} \in \Pi$ generates a distribution of interim posterior beliefs τ^{i-1} and we let $\Delta(\Delta(\Omega))$ denote the set of distributions of (interim or posterior) beliefs.

Given a joint signal π^{i-1} that induces a belief distribution τ^{i-1}, sender i can refine the information into any partition that is finer than π^{i-1}. Moreover, using a result from Green and Stokey (1978) together with the characterization in Gentzkow and Kamenica (2017a), we know that for any mean preserving spread of τ^{i-1} there exists a refinement of π^{i-1} that induces the mean preserving spread. Hence, we may replace the feasibility condition for the joint partition with a requirement that each sender is restricted to mean preserving spreads of the interim belief distribution. Hence, if sender i would be moving after a history that generates belief μ for sure (which then would have to be the prior), a belief distribution τ would be feasible for sender i if and only if it is a mean preserving spread of μ, that is

$$\sum_{\mu' \in supp(\tau')} \mu' \tau\left(\mu'\right) = \mu. \tag{11}$$

Given a non-degenerate distribution $\tau^{i-1} \in \Delta(\Delta(\Omega))$, this generalizes by summing over each μ in the support of τ^{i-1} so that τ is feasible for i if and only if

$$\sum_{\mu \in supp(\tau^{i-1})} \sum_{\mu' \in supp(\tau)} \mu' \tau\left(\mu'\right) \tau^{i-1}(\mu) = \sum_{\mu \in supp(\tau^{i-1})} \mu \tau^{i-1}(\mu). \tag{12}$$

One way to satisfy (12) is if for each belief $\mu \in \Delta(\Omega)$ and every τ with μ in its support there exists a distribution $\tau^i(\cdot|\mu)$ in $\Delta(\Delta(\Omega))$ with finite support such that the probability of any μ' is $\sum \tau^n\left(\mu'|\mu\right)\tau(\mu)$. We call such a strategy a *separable Markov strategy*:

Definition 2. *Sender i plays a (separable) Markov strategy if for every $\mu \in \Delta(\Omega)$ there exists a distribution $\tau^i(\cdot|\mu)$ in $\Delta(\Delta(\Omega))$ such for any $\mu' \in \Delta(\Delta(\Omega))$ and any history h_i that generate an interim belief distribution $\tau^{i-1} \in \Delta(\Delta(\Omega))$ the probability of belief μ' is $\sum \tau^i\left(\mu'|\mu\right)\tau^{i-1}(\mu)$.*

By Assumption 1, we pick any distinct pair $a, a' \in A$ and define the upper half-space of posterior beliefs such that the decision maker weakly prefers a to a' as

$$H\left(a \succeq a'\right) \equiv \{\mu \in \Delta(\Omega) | \sum_{\omega \in \Omega} \mu(\omega)[u_D(a,\omega) - u_D(a',\omega)] \geq 0\}. \tag{13}$$

It follows that the set of beliefs such that $a \in A$ is optimal is given by

$$M(a) = \cap_{a' \in A} H\left(a \succeq a'\right). \tag{14}$$

We note that $\cup_{a \in A} M(a) = \Delta(\Omega)$ and that for any μ in the interior of some $M(a)$ there is no $M\left(a'\right)$ such that $\mu \in M\left(a'\right)$. However, the boundary points of $M(a)$ are points where there exists some $a' \neq a$ such that $\mu \in M(a) \cap M\left(a'\right)$. By construction, $M(a)$ is a finite convex polytope for each $a \in A$. Such a convex polytope has a finite set of $J(a)$ vertices $\left\{\mu_j^a\right\}_{j=1}^{J(a)}$ and these vertices span $M(a)$ so that every $\mu \in M(a)$ can be represented as a convex combination of the vectors $\left\{\mu_j^a\right\}_{j=1}^{J(a)}$.[8] Denote

$$X = \cup_{a \in A} \left\{\mu_j\right\}_{j=1}^{J(a)} \tag{15}$$

as the set of all vertices that defines the optimal actions for the decision maker, which is finite because both Ω and A are finite. See Figure 4 for a simple example.

Next, we consider the last sender's best response. To ensure that this is well defined, we assume that the decision maker always break ties in favor of sender n so that $\sigma_D(\mu) = a$ if $\mu \in M(a)$ and $v_n(a,\mu) \geq v_n\left(a',\mu\right)$ for each $a' \neq a$. If there are multiple such rules, we arbitrarily pick one of them.

Following a history that generates belief distribution τ^{n-1} and decision rule σ_D, sender n plays a best response (using a separable Markov strategy) if generating belief distribution τ^* where $\tau^*\left(\mu'\right) = \sum \tau^n\left(\mu'|\mu\right)\tau^{n-1}(\mu)$ for some $\left\{\tau^n(\cdot|\mu)\right\}_{\mu \in supp(\tau^{n-1})}$ that solves the linear program

$$\begin{aligned} \max_{\tau} \sum_{\mu \in Supp(\tau^{n-1})} & \left[\sum_{\mu' \in supp(\tau(\cdot|\mu))} v_n(\sigma_D(\mu'),\mu')\tau(\mu'|\mu)\right]\tau^{n-1}(\mu) \\ \text{s.t.} \quad & \sum_{\mu'} \mu'\tau(\mu'|\mu) = \mu \text{ for each } \mu \in Supp(\tau^{n-1}), \end{aligned} \tag{16}$$

Program (16) is separable and can be solved by solving one linear program for each belief in the support of τ^{n-1}. Each such program is on form,

$$\begin{aligned} V_n(\mu) &= \max_{\tau \in \Delta(\Delta(\Omega))} \sum_{\mu'} v_n(\sigma_D\left(\mu'\right),\mu')\tau\left(\mu'\right) \\ \text{s.t.} \sum_{\mu'} \mu'\tau\left(\mu'\right) &= \mu, \end{aligned} \tag{17}$$

which shows that there are always best replies that are separable Markov strategies. For each pro-

[8] See Grünbaum, Klee, and Ziegler (1967)).

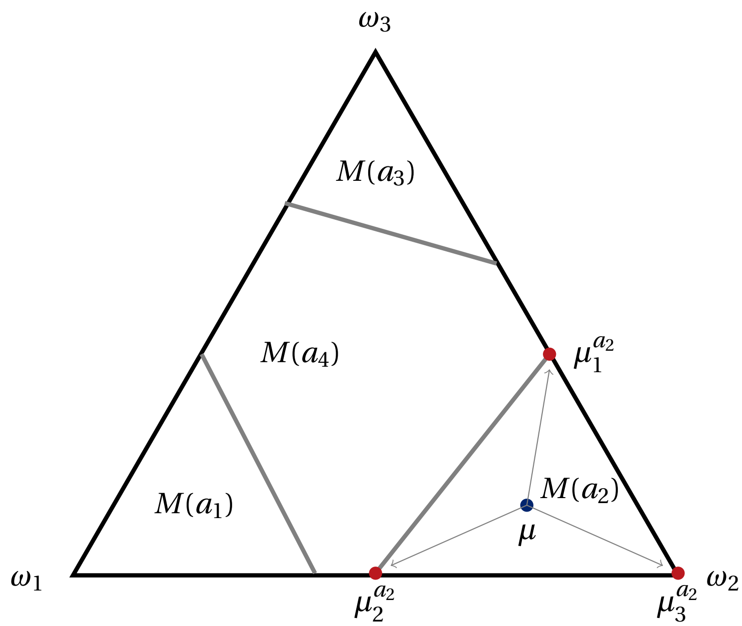

Figure 4: $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $A = \{a_1, a_2, a_3, a_4\}$.

gram on form (17), we consider a restricted *finite* linear program

$$\begin{aligned} \widetilde{V}_n(\mu) &= \max_{\tau \in \Delta(X)} \sum_{\mu' \in X} v_n(\sigma_D(\mu'), \mu')\tau(\mu') \\ \text{s.t.} \sum_{\mu'} \mu' \tau(\mu') &= \mu, \end{aligned} \tag{18}$$

Hence, (18) is well defined as it is a finite dimensional bounded linear program.

Lemma 1. *For each $\mu \in \Delta(\Omega)$ the linear program (17) is well defined. Moreover:*

1. $V_n(\mu) = \widetilde{V}_n(\mu)$ *for each* $\mu \in \Delta(\Omega)$;
2. *Every solution to (18) given belief μ also solves (17);*
3. *For every solution $\tau_n \in \Delta(\Delta(\Omega))$ to (17) there exists some $\widetilde{\tau}_n \in \Delta(X)$ that solves (18) that generates the same probability distribution over $A \times \Omega$.*

A proof is in Appendix B. The idea is that each $M(a)$ is spanned by its vertices. Hence it is possible for the sender to replace some belief μ that is not one of the vertices with a convex combination over the vertices. There are then two possibilities. The first is that the action $\sigma_D(\mu)$ is taken on all the vertices in the convex combination. In this case, the sender is indifferent between μ for sure and the convex combination over the vertices of $M(a)$. The second possibility is that a different action is taken on one or more of the vertices. Then, the sender may be strictly better off by using the convex combination. Hence, problem (18) generates a utility at least as great as (17). But, the feasible set in (18) is a subset of the feasible set in (17), so the two problems must have the same value.

See Figure 4 for a simple example. In the example, μ is in the interior of $M(a_2)$ and is induced by one of the solution to problem (17) with probability τ. There must be another solution in which $\{\mu_j^{a_2}\}_{j=1,2,.3}$ is induced with probability τ_1, τ_2, τ_3 s.t. $\sum_j \mu_j^{a_2} \tau_j = \tau$.

Lemma 1 suggests that we may characterize the best reply of every sender in terms of a finite optimization problem. The general idea is that the last sender always uses a best response with support on the vertex beliefs X, then previous senders may as well use strategies limited to the same set of vertices since the final sender will undo any attempt to generate any other beliefs by splitting them onto X. However, for this to always work we have to be careful in selecting the optimal response when sender n is indifferent between staying on a vertex belief and a mean preserving spread onto a set of different vertex beliefs.

The issue is illustrated in the example with the best response (10) in Section 3.1. In that case, the problem could be avoided by using the "consistent" best responses (eg. (8) or (9)). These best replies have the property that all beliefs in a given convex hull are split uniquely onto the spanning beliefs. Indeed, the same idea works in the general case: it is always possible to find at least one optimal solution which is "consistent" in the sense that every belief in a polytope is split onto a minimal set of vertices, and selecting best responses with this property is crucial for our existence proof.

Definition 3. *A set of beliefs $\{\mu_j\}$ is* ***minimal*** *if there is a unique $\tau \in \Delta\left(\{\mu_j\}\right)$ such that $\mu = \sum_j \tau_j \mu_j$ for any μ in the convex hull of $\{\mu_j\}$ which is denoted by $\mathrm{Co}(\{\mu_j\})$.*

When the set of beliefs are not minimal, one can always split them into multiple minimal sets. See Figure 5 for an example.

Lemma 2. *There exists an optimal solution to (18) such that:*

1. *For each belief $\mu \in \Delta(\Omega)$ the solution has a support $X(\mu)$ which is minimal.*[9]

[9] If the convex hull of $X(\mu)$ has interior points a minimal set consists of $|\Omega|$ spanning vectors. However, the convex hull could belong to a subspace of $\Delta(\Omega)$ in which the number of vectors would correspond to the dimensionality of the subspace.

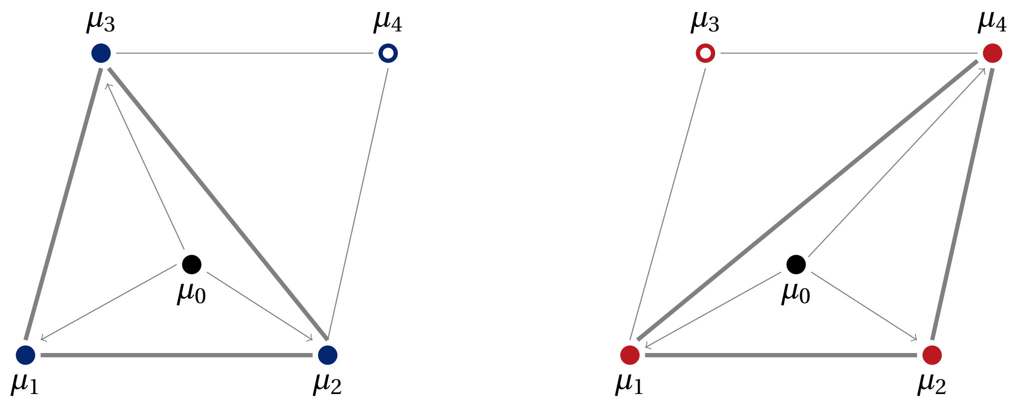

Figure 5: In the example, $|\Omega| = 3$. The set $Y = \{\mu_1, \mu_2, \mu_3, \mu_4\}$ is not minimal because there are multiple ways to split μ_0 on Y. Both $\{\mu_1, \mu_2, \mu_3\}$ (left panel) and $\{\mu_1, \mu_2, \mu_4\}$ (right panel) are minimal.

2. *If $X(\mu)$ is the solution selected for μ and $\tilde{\mu}$ is in the convex hull of $X(\mu)$ then the solution selected for $\tilde{\mu}$ is the unique distribution with support on $X(\mu)$.*

Proof. The first part of the Lemma is part of Theorem 1 in Lipnowski and Mathevet (2017), but we relegate a simple proof in Appendix B for convenience of the reader, and we present the proof of the second part here. Let $\tilde{\mu} \in \text{Co}(X(\mu))$ and let $X(\tilde{\mu}) \neq X(\mu)$ be the support for some solution $\tilde{\tau}$ to the problem given belief $\tilde{\mu}$. We thus have that

$$\begin{aligned} \mu &= \sum_{\mu' \in X(\mu)} \mu' \tau(\mu') \\ \tilde{\mu} &= \sum_{\mu' \in X(\tilde{\mu})} \mu' \tilde{\tau}(\mu') \end{aligned} \tag{19}$$

since τ solves (18) given belief μ and $\tilde{\tau}$ solves (18) given belief $\tilde{\mu}$. Also, $\tilde{\mu} \in \text{Co}(X(\mu))$ implies that there exists λ such that

$$\tilde{\mu} = \sum_{\mu' \in X(\mu)} \mu' \lambda(\mu'). \tag{20}$$

Combining the first equality in (19) with (20) and then using (19) we have that

$$\mu = \sum_{\mu' \in X(\mu)} \mu' \left(\tau(\mu') - \beta\lambda(\mu')\right) + \beta\tilde{\mu} = \sum_{\mu' \in X(\mu)} \mu' \left(\tau(\mu') - \beta\lambda(\mu')\right) + \sum_{\mu' \in X(\tilde{\mu})} \mu' \beta\tilde{\tau}(\mu') \tag{21}$$

where since $\tau(\mu') > 0$ for every $\mu' \in X(\mu)$ we have that $\tau(\mu') - \beta\tilde{t}(\mu') > 0$ for β small enough. Hence, it is feasible to split μ over $X(\mu) \cup X(\tilde{\mu})$ using coefficients

$$\left\{\tau(\mu') - \beta\lambda(\mu') + \beta\tilde{\tau}(\mu')\right\}_{\mu' \in X(\mu) \cup X(\tilde{\mu})}. \tag{22}$$

But, τ is optimal given belief μ so

$$\begin{aligned}\sum_{\mu'\in X(\mu)} v_n\left(\sigma_D\left(\mu'\right),\mu'\right)\tau\left(\mu'\right) &\geq \sum_{\mu'\in X(\mu)\cup X(\widetilde{\mu})} v_n\left(\sigma_D\left(\mu'\right),\mu'\right)\left(\tau\left(\mu'\right)-\beta\lambda\left(\mu'\right)+\beta\widetilde{\tau}\left(\mu'\right)\right) \\ &= \sum_{\mu'\in X(\mu)} v_n\left(\sigma_D\left(\mu'\right),\mu'\right)\tau\left(\mu'\right) \\ &\quad -\beta\left[\sum_{\mu'\in X(\mu)} v_n\left(\sigma_D\left(\mu'\right),\mu'\right)\lambda\left(\mu'\right)-\sum_{\mu'\in X(\widetilde{\mu})} v_n\left(\sigma_D\left(\mu'\right),\mu'\right)\widetilde{\tau}\left(\mu'\right)\right]\end{aligned} \tag{23}$$

This implies that λ is weakly better than $\widetilde{\tau}$, which is an optimal solution to (18) given belief $\widetilde{\mu}$. Hence, λ is also optimal. □

We can now complete the proof of Theorem 1:

Proof of Theorem 1. Pick some decision area $M(a)$ spanned by vertices $\left\{\mu_j^a\right\}_{j=1}^{J(a)}$ and let $\mu\in M(a)$ be arbitrarily chosen. By Lemma 1, there exists an optimal solution to program (17) with belief μ where sender n puts positive probability on μ' only if $\mu'\in\left\{\mu_j^a\right\}_{j=1}^{J(a)}$. Moreover, we may break ties so that $v_n(\sigma_D(h),\mu_j^a)\geq v_n(a',\mu_j^a)$ for every $a'\in A$. Let $\tau_n^*(\cdot|\mu)\in\Delta(X)$ be a separable Markov best reply for sender n which is defined as the solution to problem (18) for a given μ. By Lemma 2 we may pick solution τ_n^* with minimal support and where $t_n^*\left(\cdot|\widetilde{\mu}\right)$ has support on $X\left(\mu\right)$ for every $\widetilde{\mu}\in\mathrm{Co}\left(X\left(\mu\right)\right)$. Define

$$X_n=\cup_{\mu\in\Delta(\Omega)}Supp(\tau_n^*(\cdot|\mu)), \tag{24}$$

which is a finite subset of X. For each $\mu\notin X_n$, sender n finds it (weakly) optimal to generate a mean preserving spread with support X_n.

Given sender n's separable Markov strategy, sender $n-1$'s problem is also to split his interim beliefs. Using the same argument as when breaking up (16) into separate problems for each μ, we can find a best reply for sender $n-1$ by solving the linear program

$$\begin{aligned} V_{n-1}(\mu) &= \max_{\tau\in\Delta(\Delta(\Omega))}\sum_{\mu'}\left[\sum_{\mu''}v_{n-1}(\sigma_D(\mu''),\mu'')\tau_n^*(\mu''|\mu')\right]\tau(\mu') \\ \text{s.t.}\ \sum_{\mu'}\mu'\tau\left(\mu'\right) &= \mu, \end{aligned} \tag{25}$$

where $\left\{\tau_n^*(\mu''|\mu')\right\}_{\mu''\in X_n}$ is the conditional probability measure defined by the best response of sender n. Assume for contradiction that sender $n-1$ has a strict incentive to use some belief $\mu'\notin X_n$. By Lemma 2, we select a best reply by n that splits beliefs uniquely onto a minimal set of vectors $X\left(\mu'\right)$. Moreover, for any $\mu\in\mathrm{Co}\left(X\left(\mu'\right)\right)$ sender n will split beliefs uniquely onto $X\left(\mu'\right)$.

Hence, any belief in the support of the strategy of sender $n-1$ that is contained in $\text{Co}\left(X\left(\mu'\right)\right)$ will be split uniquely onto $X\left(\mu'\right)$ so sender $n-1$ may replace all beliefs in $\text{Co}\left(X\left(\mu'\right)\right)$ with a single distribution over $X\left(\mu'\right)$. Given a sender n best reply on the form of the one in Lemma 2 this is true for any μ' and the union of all the vertices of convex polytopes that define the best reply by n is given by X_n, so sender $n-1$ may without loss restrict attention to beliefs in X_n, a finite set. This problem has a solution since it is finite and bounded above and below. Hence, $n-1$ has a best reply in which each belief is split onto a minimal vector, implying that we can make the same argument for senders $n-2,...,1$. Existence of subgame perfect equilibria (in separable Markov strategies) follows by induction.

□

3.3 Reinterpreting Rich Signal Spaces in Sequential Persuasion

As a by-product of our existence argument we note if we focus on separable Markov strategies it is irrelevant whether sender i can observe a joint realization $s^{i-1} \in \pi^{i-1}$ or not. As Theorem 2 proves unique distributions over actions and states for generic preferences this is typically the relevant case.

Corollary 1. *Let $(\sigma_1,...,\sigma_n,\sigma_D)$ be a separable Markov perfect equilibrium. Then it is also a perfect Bayesian equilibrium of the game where nature draws the state ω and the sunspot variable Z at the root of the game and where each sender observes all predecessors' signal realizations in addition to the signals.*

This follows since every best response problem can be separated into one problem for each belief in the support of the distribution generated by the joint signals of previous senders. This is how we justified using (17) instead of (16), (25) instead of the analogue to (16) for sender $n-1$ and so on. This, came from using a signal structure where a joint signal realization is an intersection of individual realizations, which is why for each $s^{i-1} \in \pi^{i-1}$ we may let $\pi_i\left(s^{i-1}\right)$ define a contingent partition of s^{i-1} and note that the state dependent probability of signal $s^i \in \pi_i\left(s^{i-1}\right)$ conditional on s^{i-1} is $p(s^i|s^{i-1},\omega) \equiv \frac{p(s^i|\omega)}{p(s^{i-1}|\omega)}$. Hence, the only difference between observing the realized outcome of π^{i-1} and not is that player i conditions on s^{i-1} when it is observed and takes expectations over s^{i-1} when it is not observed. But, this is irrelevant as the expected payoff is to maximized when each component in the sum is maximal.

We further note that if senders observe the signal and the realization before adding more information it is irrelevant whether signals are represented as partitions of $\Omega \times [0,1]$ or as "plain vanilla" independent noisy signals. This implies that, if all senders move in sequence, we can interpret the assumption that signals can be arbitrarily correlated as saying that experiments are conducted sequentially with experimental results revealed in each stage.

In Board and Lu (2017) results depend critically on whether signal realizations are observable to competitors and they refer to this as a distinction between *public* and *private* persuasion. In this language, our sequential model with coordinated signals can thus be thought as a representation of a public persuasion model.

3.4 Simplifying the Problem

Players ultimately only care about the distribution over actions and states, which motivates the following definition:

Definition 4. *Two strategy profiles are **outcome equivalent** if they generate an identical joint distributions over* $\Omega \times A$.

This is in important concept because there may exist many equilibrium information structures, but where, because all players only care about distributions over $\Omega \times A$, everyone is indifferent across all equilibria. Some of these equilibria may be Blackwell comparable, but players don't care directly about informativeness, so we consider them equivalent.

Next, we define an equilibrium in which no sender except the first provides any additional information as follows:

Definition 5. *Consider a strategy profile* σ' *and let* h_i' *denote the implied outcome path before the move by sender* i*. We say that* σ' *is **one step** if* $\vee_{j=1}^{n} \sigma_i'(h_i') = \sigma_1'$.

We use one-step equilibria because they are convenient for characterizing equilibrium outcomes. It is possible that the first sender would be better off at some strictly less informative signal, in which case it may be more reasonable to assume that the first sender would add less information in the hope that the senders that follow make mistakes and don't further refine the signal. However, this is irrelevant for the set of equilibrium outcomes.

Proposition 1. *For any subgame perfect equilibrium there exists an outcome equivalent subgame perfect equilibrium in which senders play a one step continuation strategy profile after any history of play.*

The idea behind Proposition 1 is similar to the proof of the revelation principle. Consider an arbitrary subgame perfect equilibrium σ^* and let $h_i^* = \{\pi_1^*, ..., \pi_{i-1}^*\}$ be the equilibrium path history when it is sender i's turn to move. This equilibrium generates a joint signal $\pi^* = \vee_{i=1}^n \pi_i^*$. To construct a one step equilibrium we let sender 1 plays π^* and assume that *on the equilibrium path* players $i = 2, ..., n$ provide only redundant information. It then follows that the decision maker may as well generate the same distribution over $A \times \Omega$ as in the initial equilibrium after observing the one step equilibrium path history. Moreover, because π^* is finer than π_i^* for each $i < n$, any deviation that is feasible from the one step equilibrium path is feasible also in the original equilibrium, so it is possible to replicate continuation play following deviations from the one step equilibrium from the original equilibrium just like in the proof of the revelation principle. Finally, *off the equilibrium path*, we can just follow the original equilibrium strategies.

The paragraph above contains the key idea in proving that one step equilibria are sufficient, but ignores a subtle issue. We need to make sure that the consequences from a deviation by i on the one step equilibrium path replicates the consequences of an equivalent deviation from the original path, which requires us to keep track of "equivalent" histories for the two equilibria. See the proof of Proposition 1 in Appendix B for details.

Additionally, for the one step equilibrium characterization to be a significant simplification, we need it to apply not only on the equilibrium path, but also following arbitrary histories of play. We therefore need to generalize the definition to continuation strategy profiles. The logic for a one step equilibrium generalizes immediately to any *continuation equilibrium* following an arbitrary history of play, but notation gets heavy.

We next show that if we only care about the distribution over $\Omega \times A$, we may as well focus on equilibria with support on the vertices spanning the optimal decisions for the decision maker. The intuitive idea is the same as for Lemma 1, but the proof is notationally more cumbersome as we need to replicate the "incentives" corresponding to an arbitrary equilibrium with continuation play that is on the vertices only.

Proposition 2. *For every subgame perfect equilibrium there exists an outcome subgame perfect equivalent equilibrium in which senders play one step strategies with implied beliefs with support on X after every history of play.*

In the same spirit we note that when checking for subgame perfection it is without loss to consider one step deviations onto the set of vertices X. This is more or less a direct consequence of Proposition 2, but a proof is provided for completeness.

Proposition 3. *Let σ be a strategy profile that is one step after any history of play and suppose that there is no history h_i and no deviation π_i with beliefs having support on X only with π_i being the signal in a one step continuation equilibrium (possibly different from $\sigma|_{h_i}$). Then, σ is subgame perfect.*

Taken together, Propositions 1, 2, and 3 vastly simplifies the task of finding equilibria. It is useful to introduce the notion of a *stable belief* to summarize how:

Definition 6. *A belief $\mu \in \Delta(\Omega)$ is* ***stable*** *if it is in the support of a one step equilibrium.*

No player has an incentive to refine a stable belief, but the construction of the set of stable beliefs is still recursive and the order of moves matter for which beliefs are stable. In a one step equilibrium, beliefs must be immune to any Bayes plausible deviation by sender n. This restriction defines a subset of $\Delta(\Omega)$ of candidate stable beliefs, which we may call S_n, which would be the stable beliefs in the game in which n is the only sender. Next, if S_i are the stable beliefs in the continuation game where sender $i, i+1, ..., n$ move sequentially in the same order as in the full game, then stability requires that every $\mu \in S_i$ is in the support of a one step equilibrium of this continuation game. By induction, we obtain a sequence of belief sets that gets (weakly) smaller in every step and the set of stable beliefs is given by the final set in the sequence, S_1. By construction, this recursive process corresponds with beliefs consistent with a one step equilibrium.

Using the notion of stable beliefs we can rephrase Proposition 1 and 2 as

Corollary 2. *An outcome is consistent with subgame perfection if and only if there is a distribution τ^* over X such that each μ in the support of τ is stable and sender 1 weakly prefers τ^* over any other distribution over X with support on stable beliefs only.*

Moreover, we can express stable beliefs recursively:

Corollary 3. *$\mu \in X$ is a stable (vertex) belief if and only if for each sender i there exists no mean preserving spread τ onto X that i prefers to μ such that each μ' in the support of τ is stable in the truncated game with prior μ' starting with sender $i+1$.*

Hence, there is a simple algorithm for characterizing equilibrium outcomes.

Algorithm 1. *For each sequential persuasion game where both $|\Omega|$ and $|A|$ are finite, one can find the essential unique equilibrium outcome by following the procedure below:*

1. *Use the the payoff functions of the decision maker and sender n to find $M(a), \forall a \in A$.*

2. *Find the finite set of vertex beliefs X of $\{M(a)\}_{a\in A}$.*

3. *Repeat the following procedure for $i = n, n-1, \ldots, 1$ to find X_{S_i}: define X_{S_i} as a subset of $X_{S_{i+1}}$ and $\forall \mu \in X_{S_i}$, $\nexists \tau \in \Delta(X_{S_{i+1}})$ s.t. (i) $\sum_{\mu_j \in X_{S_{i+1}}} \mu_j \tau(\mu_j) = \mu$, and (ii)*

$$v_i(\sigma_D(\mu), \mu) < \sum_{\mu_j \in X_{S_{i+1}}} v_i(\sigma_D(\mu_j), \mu_j)\tau(\mu_j). \tag{26}$$

   where $X = X_{S_{n+1}}$.

4. *For each prior belief μ_0, find the minimal spanning vectors on $X(\mu_0) \subset X_{S_1}$ and a distribution $\tau \in \Delta(X(\mu_0))$ s.t. $\sum_{\mu_j \in X(\mu_0)} \mu_j \tau_j = \mu_0$.*

5. *Construct the corresponding one step equilibrium signal π^* from τ.*

An illustration of the application of Algorithm 1 can be found in section 4.

3.5 On Fully-Revealing Equilibria

Besides drastically simplifying the task of analyzing examples and special cases, the one step equilibrium characterization also provides a nice set of tools for asking when equilibria are fully-revealing.

Discussed by Sobel (2010), in most multi-sender strategic communication models, a fully revealing equilibrium exists under very weak conditions. One of the key reasons is that: when others fully reveal the state, a sender has no way to further affect the outcome in a simultaneous move game, so fully revealing is a best response. Full revelation can therefore be supported as an equilibrium outcome even if it is Pareto dominated.

Since the sequential model cannot support full revelation as a coordination failure, it is natural to ask under what conditions fully revealing is the only equilibrium outcome. We are able to identify a simple sufficient condition: one can rule out non-fully revealing equilibria as long as at each non-degenerate vertex belief of the convex polytope, there exists at least one sender who prefers full revelation to the current belief being observed by the decision maker.

Proposition 4. *Denote $\hat{X}$ as the set of vertices of $\Delta^{|\Omega|-1}$. All equilibria are fully revealing if for each $\mu \in X \setminus \hat{X}$, there exists a sender i and a $\tau \in \Delta(\hat{X})$ such that*

$$v_i(\sigma_D(\mu), \mu) < \sum_{\mu_j \in \hat{X}} v_i(\sigma_D(\mu_j), \mu_j)\tau(\mu_j). \tag{27}$$

Proof. Suppose there is a non-fully revealing equilibrium. By Proposition 2, there is an outcome equivalent one step equilibrium with support on X. If $\tau(\mu) = 0, \forall \mu \not\in \hat{X}$, then the proof is complete. Otherwise, take such a vertex belief $\mu^* \not\in \hat{X}$. In the one step equilibrium, sender 1 will split the belief and μ^* will be realized with positive probability. There exists a sender i who has the incentive to unilateral split μ^* on $\hat{X}$. Since the sequential model cannot support full revelation as a coordination failure among the senders, If $i = 1$, then it is not optimal to endow μ^* positive probability in a one step separating Markov equilibrium. If $i > 1$, sender i has the incentive to deviate at his turn. Therefore, any equilibrium must be fully revealing. □

It is easy to check condition (27) as it only depends on the decision maker's strategy and the current sender's payoff at a small number of vertices. Although persuasion is sequential, the one step characterization makes it unnecessary to take subsequent senders' action into account. In addition, the condition (27) is weaker than the zero-sum condition suggested by Gentzkow and Kamenica (2017a).

Proposition 4 contributes to the discussion on advocacy. As noted by Dewatripont and Tirole (1999), many organizations deliberately turn their (prior unbiased) members to advocates with goals differing from the one of the organization. The organization then benefits from the competition among advocates who defend their specific "causes". A leading example can be found in courts. Although the legal system is meant to "find the truth", the defense attorney is expected to stand for the defendant rather than revealing information that may potentially hurt the defendant's case to the judge. In our model, it is the decision maker's best interest to fully learn the state. An obvious way to do so is to recruit one sender who share the preference with the decision maker. However, such a goal may be too difficult or too costly to achieve for several reasons. First, a sender may have some intrinsic bias. It is time-consuming to find a sender agreeing with the decision maker in "every dimension". Second, acquiring information is typically costly. It may be too expensive to incentivize the sender to acquire perfect information. However, Proposition 4 suggests that competition among multiple carefully selected senders may be ideal. Although it is difficult to make one sender always agree with the decision maker, it may be easier to make a sender dislike a particular non-degenerate vertex belief. To learn the true state, the decision maker only needs to find or incentivize a set of senders so that each of the non-degenerate vertex beliefs will be opposed by one sender.

3.6 Essential Uniqueness

In this section, we prove that all subgame perfect equilibria are outcome equivalent for generic preferences.

Definition 7. *An equilibrium is **essentially unique** if all equilibria are outcome equivalent.*

Theorem 2. *Fix any μ_0. Then there is an essentially unique subgame perfect equilibrium that induces distribution τ over a minimal set of vectors $X(\mu_0)$ for a set of (bounded) Bernoulli payoff function profiles with full Lebesgue measure.*

The proof of Theorem 2 combines two ideas. The first is that we may consider “one step vertex equilibria”. The second is that it is rare that players have preferences that make them indifferent between a belief and a “relevant" mean preserving spread.

By Proposition 2 it is without loss to consider subgame perfect equilibria in which, after any history of play, continuation strategies are one step strategies inducing beliefs with support on vertices in the finite set X. Moreover, Proposition 3 establishes that we only need to check for deviations generating beliefs on X. By separability, we may check this belief by belief. Hence, if there are multiple continuation equilibria for some history there is a belief such that some sender is indifferent between staying at this belief and some mean preserving spread over X.

There are two cases to consider. The first is that each belief in the mean preserving spread induces the same action as providing no additional information which does not violate essential uniqueness because the distribution over $A \times \Omega$ is unchanged. The second is that there are distinct actions induced by the mean preserving spread, but then indifference is only possible for a lower dimensional subset of preferences. This is because indifference means indifference between no additional information and a mean preserving spread over a minimal set of vectors. Each minimal set of vectors defines a unique mean preserving spread and for each of these there is a lower dimensional set of preferences consistent with indifference. There is a finite set of possible minimal vectors, so by adding a finite set of measure zero possibilities we get a measure zero possibility.

3.6.1 Proof of Theorem 2

We now provide a detailed proof of Theorem 2. First, we need a few intermediate results. non-Markov

Lemma 3. *The decision maker must follow a Markov strategy for a set of bounded Bernoulli payoff functions for the decision maker and sender n with full Lebesgue measure.*

There are two pathological cases that needs to be addressed. Firstly. it may be that there is some state $\omega \in \Omega$ in which the decision maker is indifferent between two actions. In that case payoff irrelevant aspects of the history, such as which sender revealed state ω, can be used to construct non-Markov strategies for the decision maker. See Appendix A for detail. The second case is that there is some interior vertex associated with some decision area $M(a)$ where both sender n and the decision maker are indifferent. Both these cases are rare in the sense that the associated payoff functions is a measure zero subset of all conceivable payoff functions.

By Lemma 3 we may for generic preference profiles assume that the decision maker uses a Markov strategy. We abuse previous notation and write $\sigma_D : \Delta(\Omega) \to A$. By substituting the equilibrium decision rule σ_D into (1) we define

$$\hat{v}_i(\mu) \equiv v_i(\sigma_D(\mu), \mu). \tag{28}$$

That is, given that the decision maker plays a Markov strategy we may work with the a reduced form payoff function directly over beliefs for each sender i.

Next, we show that for most stable beliefs, no sender has a weak incentive to add information. To state this result, let X_{S_i} be the intersection between the stable beliefs and X in the truncated game starting with sender i:

Lemma 4. *Suppose that the decision maker plays a Markov strategy. Then, for any $\mu \in X_{S_i}$, any $i = 1, 2, ..., n$, and any non-empty set $Y \subset X_{S_i} \setminus \{\mu\}$ either:*

1. $\sigma_D(\mu') = \sigma_D(\mu)$ *for every* $\mu' \in Y$, *or,*
2. *there exists* $\mu' \in Y$ *such that* $\sigma_D(\mu') \neq \sigma_D(\mu)$. *In this case*

$$\hat{v}_i(\mu) > \sum_{\mu' \in Y} \hat{v}_i(\mu') \tau(\mu') \tag{29}$$

for every τ such that $\sum_{\mu' \in Y \setminus \{\mu\}} \mu' \tau(\mu') = \mu$ and a set of (bounded) sender i Bernoulli utility functions over $A \times \Omega$ with full Lebesgue measure.

The idea is that if some sender i has a weak incentive to split a stable belief μ into a distribution over $X_{S_i} \setminus \{\mu\}$ then it must be that

$$\hat{v}_i(\mu) = \sum_{\mu' \in X_{S_i} \setminus \{\mu\}} \hat{v}_i(\mu') \tau(\mu') \tag{30}$$

for some sender i and some τ that satisfies Bayes plausibility. If the action is the same for each $\mu' \in X_{S_i}$ equality (30) always holds, but then moving probability from μ to $X_{S_i}\backslash\{\mu\}$ is irrelevant for the joint distribution over actions and states, so the multiplicity is not essential.

The interesting case is when X_{S_i} contains beliefs that result in at least two distinct actions. Since X_{S_i} contains the vertices of $\Delta(\Omega)$ the dimensionality of $\mathrm{Co}(X_{S_i}\backslash\{\mu\})$ is the same as the dimensionality of Ω. It follows from Carathéodory's theorem that μ can be spanned by $|\Omega|$ vectors in $X_{S_i}\backslash\{\mu\}$. Depending on whether μ is in the interior of $\mathrm{Co}\left(X_{S_i}\backslash\{\mu\}\right)$ or on a face, a set of $|\Omega|$ spanning vectors may not be minimal. However, if there are redundant vectors, we can eliminate vectors until μ is spanned by a minimal set of vectors in $X_{S_i}\backslash\{\mu\}$.

If sender i is indifferent between μ and some distribution over $X_{S_i}\backslash\{\mu\}$ there is some minimal set of vectors that creates indifference. But, provided that there are at least two distinct actions, the indifference condition for any particular minimal set of vectors is satisfied with probability zero as the indifference condition defines a lower dimensional subset of (bounded) utility functions. Since the number of possible minimal set of vectors and senders are both finite, the result follows by induction. See Appendix B for details.

In a similar spirit we establish that indifferences over distinct distributions over stable continuation beliefs are rare.

Lemma 5. *Fix any $i \in \{1,...,n\}$, $\mu \in X\backslash X_{S_i}$ and any distinct pair $(X(\mu), \widetilde{X}(\mu)) \in X_{S_i}$ of minimal spanning vectors of μ such that at least two distinct actions a, a' are taken for beliefs in $X(\mu) \cup \widetilde{X}(\mu)$. Then*

$$\sum_{\mu' \in X(\mu)} \widehat{v}_i(\mu')\,\tau(\mu') \neq \sum_{\mu' \in \widetilde{X}(\mu)} \widehat{v}_i(\mu')\,\widetilde{\tau}(\mu') \tag{31}$$

for a set of (bounded) sender i Bernoulli utility functions over $A \times \Omega$ with full Lebesgue measure.

Using Lemmas 4 and 5, we are now in position to prove our second main result.

Proof of Theorem 2. Using Proposition 2, every subgame perfect equilibrium is outcome equivalent to an equilibrium in which senders play one step strategies with beliefs on X after every history of play. Hence, it is without loss to assume that every sender i generates a signal with corresponding beliefs with support on X_{S_i} after any history of play. By Lemma 3 it is without loss to assume that the decision maker follows a Markov decision rule given preferences for the decision maker and sender n chosen from a set with full Lebesgue measure. We also know from the separability of the problem facing each sender i that if essential uniqueness fails there exists some sender i and some $\mu \in X$ such that sender i has multiple best responses to this belief that generate distinct distributions over $A \times \Omega$. Suppose such a belief exists and suppose first that $\mu \in X_{S_i}$. Since μ is stable, equation (30) holds for some mean preserving spread of μ onto $X_{S_i} \setminus \{\mu\}$ with at least two actions taken by the decision maker. Given that preferences for the decision maker and sender n are generic, Lemma 4 rules this possibility out for full Lebesgue measure set of payoff functions. The other possibility is that $\mu \in X \setminus X_{S_i}$ and that the sender is indifferent between two distinct mean preserving spreads that do not always generate the same action, which for generic preferences is ruled out by Lemma 5. Since the possibility is non-generic for each player i it is also non-generic in the space of payoff profiles. □

Our final characterization result says that if prior μ_0 generates an essentially unique equilibrium with minimal support $X(\mu_0)$ (which is generically true), then the essentially unique equilibrium given a prior $\widetilde{\mu}_0$ in the convex hull of $X(\mu_0)$ is the unique Bayes plausible distribution over $X(\mu_0)$ given prior $\widetilde{\mu}_0$. Equilibrium outcomes are robust with respect to perturbations of prior beliefs that are in the interior of $X(\mu_0)$.

Proposition 5. *Suppose that the essentially unique equilibrium given prior μ_0 is distribution τ over (minimal) support $X(\mu_0)$. Then, for every $\widetilde{\mu}_0 \in \mathrm{Co}(X(\mu_0))$ every equilibrium outcome is outcome equivalent with the unique Bayesian plausible belief distribution λ over $X(\mu_0)$ for prior $\widetilde{\mu}_0$.*

4 An Application

In this section, we apply results derived in section 3 to study a criminal trial model. This application illustrates (i) that to find an equilibrium one only needs to check the stability of a few vertices; (ii) the logic behind the condition for full revelation in (27), and ; (iii) how, in practice, essential uniqueness falls out from our algorithm.

A prosecutor (sender 1) and a defense attorney (sender 2) move sequentially to persuade the judge (decision maker). There are three states, the defendant being innocent (ω_0), guilty of a misdemeanor (ω_1), and guilty of a felony (ω_2). The judge can choose one of the three actions, a_0, a_1, a_2, which represent ideal sentences in accordance to the true state respectively.

We assume that the judge's payoff is

	a_0	a_1	a_2
ω_0	1	0	-1
ω_1	0	1	0
ω_2	-1	0	1

(32)

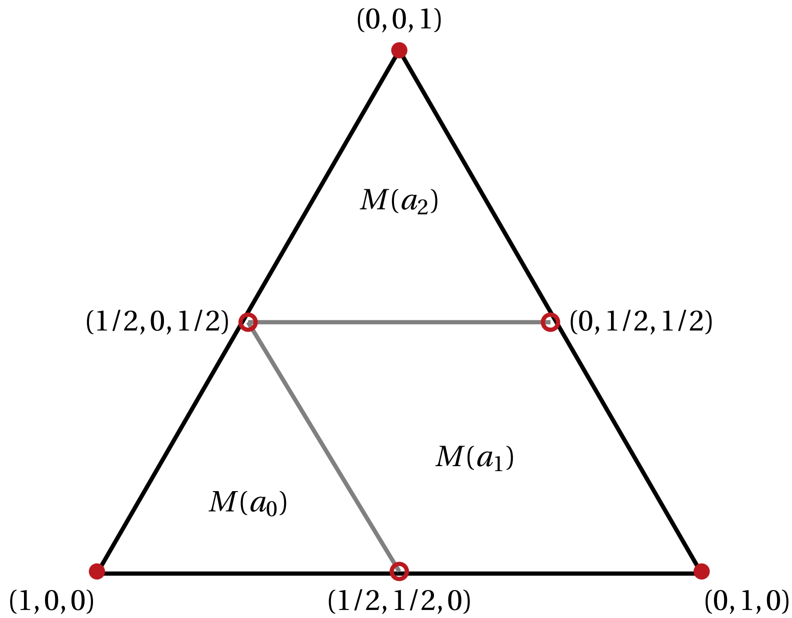

Figure 6: In the application, $X = \{(0,0,1),(1,0,0),(0,1,0),(1/2,0,1/2),(0,1/2,1/2),(1/2,1/2,0)\}$. The vertices of the simplex, $(1,0,0),(0,1,0),(0,0,1)$ represent the degenerated beliefs that the state being $\omega_0,\omega_1,\omega_2$ respectively.

Assume that senders only care about "winning", so that their preferences are state independent. Specifically, for $\alpha > 0$ and $\beta > 0$, the payoffs of the prosecutor and the defense attorney are

$$u_1(a) = \begin{cases} -1 & \text{if } a = a_0 \\ 0 & \text{if } a = a_1 \\ \beta & \text{if } a = a_2 \end{cases}, u_2(a) = \begin{cases} \alpha & \text{if } a = a_0 \\ 0 & \text{if } a = a_1 \\ -1 & \text{if } a = a_2 \end{cases}, \tag{33}$$

implying that the prosecutor prefers a guilty conviction to a misdemeanor to acquittal and that the defense attorney has the opposite preferences.

The beliefs that make each action optimal for the decision maker are depicted in Figure 6. That is, a_0 is optimal for the decision in the convex hull of $(1,0,0),(1/2,1/2,0),(1/2,0,1/2)$, a_1 is optimal in the convex hull of $(0,1,0),(1/2,1/2,0),(0,1/2,1/2)$, and a_2 is optimal in the convex hull of $(0,0,1),(1/2,0,1/2),(0,1/2,1/2)$. Hence, the set of vertices spanning $M(a_0), M(a_1)$, and $M(a_2)$ consists of the vertices of the simplex together with $(1/2,1/2,0)$, $(1/2,0,1/2)$, and $(0,1/2,1/2)$. Now apply Algorithm 1 to find the essentially unique equilibrium:

1. If $(1/2,1/2,0)$ is a stable belief then the decision maker must brake the tie in favor of the defense. Hence, the decision is a_0. But, then the prosecutor has an incentive to split the beliefs to $(1,0,0)$ and $(0,1,0)$.

2. Symmetrically, if $(1/2, 0, 1/2)$ is a stable belief the decision maker brakes the tie in favor of the defense, meaning that the decision again is a_0. But, then splitting beliefs to $(1, 0, 0)$ and $(0, 0, 1)$ is a profitable deviation for the prosecutor.

3. If $(0, 1/2, 1/2)$ is a stable belief the decision maker again brakes the tie in favor of the defense, which in this case results in decision a_1. The prosecutor is better off splitting beliefs to $(0, 1, 0)$ and $(0, 0, 1)$.

The conclusion is that full revelation is the unique equilibrium outcome for any prior belief μ_0, which confirms the sufficient condition for full revelation in Proposition 4. Also, if $\mu_0 = (1/3, 1/3, 1/3)$, the expected payoffs are $\left(\beta - 1\right)/3$ and $(\alpha - 1)/3$ for the prosecution and the defense. It is noteworthy that, when $\alpha < 1$ and $\beta < 1$, both senders strictly prefer providing no information (results in misdemeanor conviction for sure) to the equilibrium outcome. Thus there is a flavor of a prisoners dilemma in the model even if both senders are purely "career motivated".

5 Comparative Statics

There may be multiple equilibrium belief systems that can be ranked according to the Blackwell order, but where the difference in informativeness is irrelevant because all equilibria induce the same joint distribution over $A \times \Omega$. We therefore treat π and π' as equivalent in terms of the information content provided that they are outcome equivalent:

Definition 8 (Essential Blackwell Order)**.** *π is **essentially less informative** than π' if the finest signal that is outcome equivalent to π is less informative than the finest signal that is outcome equivalent to π' in the Blackwell order*

In the rest of this section, we adopt the essential Blackwell order to study two comparative statics exercises.

5.1 Adding Experts

Intuition suggests that competition from an increase in the number of experts should increase the amount of information revealed in the market. This view may even be seen as an intellectual foundation for freedom of speech, a free press, the English common law system, and many other institutions. While the literature provides a somewhat mixed support for this view, Gentzkow and

Kamenica (2017a,b) provide some sufficient conditions under which additional senders does not reduce the amount of information revealed in equilibrium in a simultaneous move Bayesian persuasion game.

The support for competition being beneficial becomes even weaker in our sequential framework. To start with, the implications of adding additional experts depend on the ordering of the players as shown in the following simple example. Let the state space be $\{\omega_0, \omega_1\}$ and the available actions be $\{a_0, a_1, a_2\}$ and assume that

$$M(a) = \begin{cases} \left[0, \frac{1}{3}\right] & \text{for } a = a_0 \\ \left[\frac{1}{3}, \frac{2}{3}\right] & \text{for } a = a_1 \\ \left[\frac{2}{3}, 1\right] & \text{for } a = a_2 \end{cases} . \tag{34}$$

Assume the senders have state independent preferences with sender 1 having strict preference order $a_1 \succ_1 a_2 \succ_1 a_0$ and sender to ranking the actions in accordance with $a_2 \succ_2 a_0 \succ_2 a_1$. Let the prior probability that the state is ω_0 be some $\mu_0 > 1/3$.

In the single sender persuasion game with the players being sender 1 and the decision maker, sender 1 may without loss optimally choose the unique mean preserving spread with support on $\{1/3, 2/3\}$ if $\mu_0 \in [1/3, 2/3]$, and the unique mean preserving spread with support $\{2/3, 1\}$ when $\mu_0 > 2/3$.

Now, suppose sender 2 moves *before* sender 1. Then any interim belief μ such that $0 < \mu < 1/3$ will be split by 1 onto $\{0, 1/3\}$ which implies that action a_0 is chosen with some probability and a_1 is chosen with some probability. Any belief on $[1/3, 2/3]$ ultimately leads to action a_1 for sure and a belief on $(2/3, 1)$ implies that action a_1 is taken with some probability and a_2 with some probability. It follows that the unique best response by 2 is to play the unique mean preserving spread onto $\{0, 1\}$, so the state is fully revealed. In contrast, if 1 moves before 2 we note that 2 will split any $0 < \mu < 1/3$ onto $\{0, 2/3\}$, implying that a best response for 1 is to split the prior to $\{0, 2/3\}$ if $\mu_0 < 2/3$ and to $\{2/3, 1\}$ if $\mu_0 \geq 2/3$. Hence, we see that the informativeness and equilibrium payoffs depend on the order of moves, and also that the equilibrium in the model with 1 moving before 2 is strictly less informative in the Blackwell order than when 1 is the only sender, provided that $\mu_0 \in [2/3, 1)$.

The example is special, but it implies that it is hopeless to get any general results unless we make restrictions on when a new sender moves. Also, in the example the equilibrium is more informative with the new sender when added as a first mover. This is not quite general due to the incompleteness of the Blackwell ordering, but we are able to establish an analogue of the comparative statics result in simultaneous persuasion games:

Proposition 6. *Suppose that there is an essentially unique equilibrium with n senders. Then, if a sender is added who moves before all other senders, there is no equilibrium with $n+1$ senders that is essentially less informative than the equilibrium in the original game.*

Proof. Suppose not. Let $X_n(\mu_0)$ be the set of spanning vectors of an essentially unique equilibrium with n senders that cannot be refined further into a Blackwell dominant outcome equivalent equilibrium. For contradiction suppose that a coarser equilibrium exists in the game with $n+1$ senders. Then there exist at least one belief μ' in the support of the equilibrium with $n+1$ senders that can be split into $X_n(\mu_0)$. For μ' to be in the convex hull of $X_n(\mu_0)$ but not part of the (unique) equilibrium at least one sender $i \in \{1,...,n\}$ strictly prefers a mean preserving spread onto $X_n(\mu_0)$ to μ' and for μ' to be part of the unique equilibrium with $n+1$ senders every $i \in \{1,...,n+1\}$ is weakly better off at μ' than at any mean preserving spread to $X_n(\mu_0)$, contradiction. □

In the special case when there are only two states we have a stronger result:

Proposition 7. *Suppose that $\Omega = \{\omega_0, \omega_1\}$ and that the equilibrium with n senders is essentially unique. Then, if a sender is added who moves before all other senders, every equilibrium with $n+1$ senders is weakly essentially more informative in the Blackwell ordering.*

Proof. Without loss of generality, consider a one step equilibrium with support on X, which now are beliefs where the decision maker is indifferent between two actions together with 0 and 1. Suppose that $\{\mu_L, \mu_H\}$ are stable vertex beliefs in the equilibrium with n senders, that $\mu_L < \mu_0 < \mu_H$, and that μ_M with $\mu_L < \mu_M < \mu_H$ is a stable vertex belief in some equilibrium with $n+1$ senders. Without loss we can assume that there at least two distinct actions that are taken at beliefs $\{\mu_L, \mu_M, \mu_H\}$ as otherwise μ_M would not be on the set of vertices X. But, for μ_M to be stable with $n+1$ players every $i \in \{1,...,n+1\}$ must be weakly better off at μ_M than at the unique mean preserving spread onto $\{\mu_L, \mu_H\}$. This implies that transferring probability from $\{\mu_L, \mu_H\}$ to μ_M is consistent with equilibrium in the model with n senders, contradicting uniqueness with n senders. □

5.2 Simultaneous vs Sequential Persuasion

Finally, we compare the information revealed in the sequential persuasion model with the one in Gentzkow and Kamenica (2017a) where senders choose their signals simultaneously. They show that:

Lemma 6 (Gentzkow and Kamenica, 2017a)**.** *Consider a persuasion game with the same structure as our model except that all senders move simultaneously. Assume that τ is the distribution of beliefs in a Markov perfect equilibrium. Then, τ is Bayes plausible given the prior belief μ_0 and for each μ in the support of τ there exists no τ' that is Bayes plausible given μ that is preferred to any sender.*

Hence, the difference between the sequential model and the simultaneous model boils down to a comparison that can be done belief by belief. A belief in the support of an equilibrium of the sequential model must be unimprovable with respect to Bayes plausible deviations over the set of stable beliefs, that is, beliefs that no player would like to further refine. In contrast, a belief in the support of an equilibrium in the simultaneous move game must be unimprovable with respect to any Bayes plausible deviation.

We can then finally conclude that in terms of the distribution over $A \times \Omega$ it cannot be that the simultaneous move game creates an outcome corresponding to a less informative signal.[10]

Proposition 8. *Consider the generic case in which the sequential game has an essentially unique equilibrium. Then, there exists no equilibrium in the simultaneous game that is essentially less informative than the equilibrium in the sequential game.*

In the special case with two states, we have a stronger result.

Proposition 9. *Suppose that $\Omega = \{\omega_0, \omega_1\}$ and that there is an essentially unique equilibrium in the sequential game. Then, any equilibrium in the simultaneous move game is weakly essentially more informative.*

The intuition is similar to the one for Proposition 7.

6 Concluding Remarks

We consider a sequential Bayesian persuasion model with multiple senders. Because it is without loss of generality to focus on equilibria corresponding to a finite set of beliefs we can show that Markov equilibria exist and generate a unique joint distribution over states and outcomes for generic preferences. The fact that a finite set of stable beliefs characterizes the equilibrium is also very convenient computationally. We suggest an algorithm to compute the essentially unique

[10] A similar comparison is made in the multi-sender cheap talk literature. The conditions under which a fully revealing equilibrium exists is weaker in a simultaneous-move cheap talk model than a sequential-move one. See Ambrus and Takahashi (2008), Battaglini (2002), Kawai (2015), and Krishna and Morgan (2001).

equilibrium. We also provide a simple condition under which full revelation is the only equilibrium outcome. In Section 4, we apply these insights on a simple criminal trial example, where we can easily show that there is a unique equilibrium that is fully revealing despite both senders being better off providing no information at all. Some comparative statics results are derived. Adding a sender who moves first cannot reduce informativeness in equilibrium, and will result in a more informative equilibrium in the case with two states. Sequential persuasion cannot generate a more informative equilibrium than simultaneous persuasion, and is less informative in the case with two states.

A Appendix: Non-Markov Equilibrium

Suppose that $\Omega = \{\omega_0, \omega_1\}$ and the optimal choice correspondence for the decision maker is

$$\sigma\left(\mu\right) = \begin{cases} \{a_1, a_2\} & \text{if } \mu \leq 1/10 \\ a_3 & \text{if } 0.1 \leq \mu \leq 9/10 \\ \{a_4, a_5\} & \text{if } \mu \geq 9/10 \end{cases} \tag{A.1}$$

Also suppose that two senders have state independent preferences

$$u_1\left(a,\omega\right) = \begin{cases} 3 & \text{if } a \in \{a_1, a_4\} \\ 1 & \text{if } a = a_3 \\ 0 & \text{if } a \in \{a_2, a_5\} \end{cases}, \text{ and } u_2\left(a,\omega\right) = \begin{cases} 3 & \text{if } a \in \{a_2, a_5\} \\ 1 & \text{if } a = a_3 \\ 0 & \text{if } a \in \{a_1, a_4\} \end{cases} \tag{A.2}$$

Consider a Markov equilibrium. Allowing for mixed strategies let $\sigma_1(0)$ be the probability for a_1 given belief $\mu = 0$ and $\sigma_4(1)$ be the probability of a_4 given belief $\mu = 1$. Suppose that the decision maker has full information. Then, the payoffs of sender 1 and 2 are $3[\sigma_1(0) + \sigma_4(1)]/2$ and $3[2 - \sigma_1(0) - \sigma_4(1)]/2$ respectively, so the payoff is greater than or equal to $3/2$ for at least one sender. Hence, beliefs in $[1/10, 9/10]$ can be ruled out in any Markov equilibrium. In contrast, if the decision maker always breaks the tie against the sender who first split the belief into $[0, 1/10]$ or $[9/10, 1]$ each sender may as well not provide any information and qualitatively different equilibria with action a_3 can be supported by such non-Markov strategies.

B Appendix: Omitted Proofs

Proof of Lemma 1. Pick any feasible solution τ to program (17). For each $a \in A$ write $\tau^a\left(\mu'\right)$ for μ' such that $\sigma_D\left(\mu'\right) = a$ and $\tau = \left\{\left\{\tau^a\left(\mu'\right)\right\}_{\mu' \in \widehat{M}(a)}\right\}_{a \in A}$ where $\widehat{M}(a) = \{\mu \in \Omega | \sigma_D(\mu) = a\}$ is the "decision area" of action a defined by $\sigma_D(\cdot)$. Obviously, $\widehat{M}(a) \subset M(a), \forall a$.

For each $\mu' \in M(a)$ there exists $\lambda' \in \Delta\left(\left\{\mu_j^a\right\}_{j=1}^{J(a)}\right)$ such that $\mu' = \sum_{j=1}^{J(a)} \lambda'_j \mu_j^a$. For every $a \in A$ and μ_j^a spanning $M(a)$ let $\widehat{\tau}\left(\mu_j^a\right) = \sum_{\mu' \in \widehat{M}(a)} \tau\left(\mu'\right) \lambda'_j$ so that

$$\sum_{j=1}^{J(a)} \widehat{\tau}\left(\mu_j^a\right) = \sum_{\mu' \in \widehat{M}(a)} \tau\left(\mu'\right) \sum_{j=1}^{J(a)} \lambda'_j = \sum_{\mu' \in \widehat{M}(a)} \tau\left(\mu'\right). \tag{B.1}$$

Since it is possible that $v_n\left(a, \mu_j^a\right) < v_n\left(a', \mu_j^a\right)$ for some $\mu_j^a \in M(a)$ (and $\mu_j^a \notin \widehat{M}(a)$, because breaking the tie in favor of a' may be better than a) it follows that the solution to (18) satisfies

$$\begin{aligned} \widetilde{V}_n\left(\mu\right) &\geq \sum_{a \in A} \sum_{j=1}^{J(a)} v_n(a, \mu_j^a) \widehat{\tau}\left(\mu_j^a\right) = \sum_{a \in A} \sum_{j=1}^{J(a)} \sum_{\omega \in \Omega} u_n(a, \omega) \mu_j^a(\omega) \widehat{\tau}\left(\mu_j^a\right) \\ &= \sum_{a \in A} \sum_{\omega \in \Omega} u_n(a, \omega) \sum_{j=1}^{J(a)} \left[\mu_j^a(\omega) \lambda'_j\right] \left[\sum_{\mu' \in \widehat{M}(a)} \tau\left(\mu'\right)\right] = \sum_{a \in A} \sum_{\omega \in \Omega} u_n(a, \omega) \mu' \left[\sum_{\mu' \in \widehat{M}(a)} \tau\left(\mu'\right)\right] \\ &= \sum_{\mu'} v_n(\sigma_D\left(\mu'\right), \mu') \tau\left(\mu'\right). \end{aligned} \tag{B.2}$$

This holds for any feasible solution to (17). Hence, $\widetilde{V}_n\left(\mu\right) \geq V_n\left(\mu\right)$. Moreover, any optimal solution to (18) is a feasible solution to (17), so $\widetilde{V}_n\left(\mu\right) \leq V_n\left(\mu\right)$. This establishes that solutions to (17) exist and that $\widetilde{V}_n\left(\mu\right) = V_n\left(\mu\right)$ and that every $\widetilde{\tau}_n \in \Delta(X)$ that solves (18) also solves (17). Finally, if τ_n solves (17) and μ' is such that $\tau_n(\mu') > 0$ there can be no $\mu_k^a \in M(a)$ such that $v_n\left(a, \mu_k^a\right) < v_n\left(a', \mu_k^a\right)$ and $\lambda'_k > 0$ for the weight on vector μ_k^a in the convex combination such that $\mu' = \sum_{j=1}^{J(a)} \lambda'_j \mu_j^a$. This is seen from noting that this would generate a strict inequality in the first inequality of (B.2). □

Proof of Part 1 of Lemma 2. Let $X\left(\mu\right)$ be the support of a solution τ to (18) given interim belief μ. By linearity of payoffs in probabilities, any mean preserving spread over $X\left(\mu\right)$ solves (18). Write $\text{Co}(X)$ for the convex hull of a set X, and, without loss, assume that $\text{Co}\left(X\left(\mu\right)\right)$ has dimension $|\Omega|-1$ (otherwise the same argument applies in the relevant subspace). Also suppose that $X\left(\mu\right)$ contains at least $|\Omega|+1$ points. Then, by Radon's theorem (see Radon (1921)), $X\left(\mu\right)$ can be partitioned into sets I and J such that $\text{Co}(I)$ and $\text{Co}(J)$ has at least one point in common. Label the sets so that I contains at least as many points as J. Suppose first that $J = \left\{\mu_j\right\}$ is a singleton set. Then the only way for $\text{Co}(I)$ and $\text{Co}(J)$ to intersect is for $\mu_j \in \text{Co}(I)$. Hence, there is an optimal solution to (18) in $X\left(\mu\right) \backslash \mu_j$. Suppose instead that J contains two or more points. Then, for $\text{Co}(I)$ and $\text{Co}(J)$ to intersect there must exist some pair $\left(\mu_j^1, \mu_j^2\right) \in J$ such that either one of these points belongs to $\text{Co}(I)$ or μ_j^1 and μ_j^2 are in opposite half spaces relative the hyperplane defined by $\text{Co}(I)$. If μ_j^1

or μ_j^2 belongs to $\text{Co}(I)$ one can pick an optimal solution with support in $X(\mu)\setminus\mu_j^1$ or $X(\mu)\setminus\mu_j^2$. If μ_j^1 and μ_j^2 are in opposite half spaces we note that $\text{Co}\left(X(\mu)\setminus\mu_j^1\right)\cup\text{Co}\left(X(\mu)\setminus\mu_j^2\right)=\text{Co}\left(X(\mu)\right)$ so $\mu\in\text{Co}\left(X(\mu)\setminus\mu_j^1\right)$ or $\mu\in\text{Co}\left(X(\mu)\setminus\mu_j^2\right)$. Hence, again one can pick an optimal solution with support in $X(\mu)\setminus\mu_j^1$ or $X(\mu)\setminus\mu_j^2$. By induction there is thus a solution with support on exactly $|\Omega|$ vectors. Therefore, we can without loss to assume $|X(\mu)|=|\Omega|$. Because $\text{Co}(X(\mu))$ and $\Delta^{|\Omega|-1}$ are isomorphic, $X(\mu)$ is minimal □

Proof of Proposition 1. To proceed, we extend the definition of one step equilibrium to off-the-path of play:

Definition 9. *Consider strategy σ' and let h_i be an arbitrary history when sender $i\in\{1,...,n-1\}$ moves. Also for $j\geq i$ let $h'_j|h_i$ be the implied continuation outcome path induced if each player $j\geq i$ follows σ'_j after history h_i and let $\sigma'|_{h_i}$ denote the continuation strategy profile.*[11] *We say that $\sigma'|_{h_i}$ is **one step** if $\vee_{j=i}^{n}\sigma'_j(h'_j|h_i)=\sigma'_i(h_i)$.*

Now, we are ready to proceed the proof. Fix a subgame perfect equilibrium σ^* and $h_i=(\pi_1,...,\pi_{i-1})$ be an arbitrary history when i moves. Let $\left(\pi_i^*|_{h_i},...,\pi_n^*|_{h_i}\right)$ be the continuation equilibrium path following h_i. Let

$$\pi^*|_{h_i}=\left(\vee_{i=i}^{i-1}\pi_i\right)\vee\left(\vee_{i=i}^{n}\pi_i^*|_{h_i}\right) \tag{B.3}$$

be the joint signal generated by the continuation equilibrium path. Replace the continuation equilibrium strategies following h_i by $\left(\sigma'_i,....,\sigma'_n,\sigma_D\right)$ where on the continuation outcome path

$$\begin{aligned}\sigma'_i(h_i) &= \pi^*|_{h_i} \\ \sigma'_j\left(h_i,\pi^*|_{h_i},...,\pi^*|_{h_i}\right) &= \pi^*|_{h_i}\text{ for }j\in\{i+1,..,n\} \\ \sigma'_D\left(h_i,\pi^*|_{h_i},...,\pi^*|_{h_i},s\right) &= \sigma_D\left(h_i,\left(\pi_i^*|_{h_i},...,\pi_n^*|_{h_i}\right),s\right),\end{aligned} \tag{B.4}$$

For a history in which i plays $\pi^*|_{h_i}$ but some $j\in\{i+1,...,n\}$ deviates let

$$\begin{aligned}\sigma'_k\left(h_i,\pi^*|_{h_i},...,\pi^*|_{h_i},\pi_j,....,\pi_k\right) &= \sigma_k^*\left(h_i,\pi_i^*|_{h_i},...,\pi_j^*|_{h_i},\pi_j,....,\pi_k\right) \\ \sigma'_D\left(h_i,\pi^*|_{h_i},...,\pi^*|_{h_i},\pi_j,....,\pi_n\right) &= \sigma_D^*\left(h_i,\pi_i^*|_{h_i},...,\pi_j^*|_{h_i},\pi_j,....,\pi_n\right),\end{aligned} \tag{B.5}$$

and for any other history let

$$\begin{aligned}\sigma'_j\left(h_i,\pi_i,..,\pi_{j-1}\right) &= \sigma_j^*\left(h_i,\pi_i,..,\pi_{j-1}\right)\text{ for }j\in\{i+2,..,n\} \\ \sigma'_D\left(h_i,\pi_i,..,\pi_{j-1},s\right) &= \sigma_D^*\left(h_i,\pi_i,..,\pi_{j-1},s\right)\end{aligned} \tag{B.6}$$

[11] That is, $h'_i|h_i=h_i$, $h'_{i=1}|h_i=\left(h_i,\sigma'_i(h_i)\right)$, $h'_{i+2}|h_i=\left(h_i,\sigma'_i(h_i),\sigma'_{i+1}\left(h_i,\sigma'_i(h_i)\right)\right)$ and so on.

The decision maker plays an optimal response following any path of play after h_i as after each continuation path the response is selected as some response for an identical joint signal. Moreover, if each $j \geq i$ plays in accordance with σ'_j it follows from (B.4) that the implied distribution over $\Omega \times A$ is identical to if each $j \geq i$ plays in accordance with the original equilibrium σ^*. Also, the strategies in (B.6) imply that continuation play after a deviation by i is the same under σ' as under σ^*, so i has no incentive to deviate. As σ^* is subgame perfect, the continuation play in (B.6) is trivially subgame perfect. Finally, (B.5) implies that play after if $j > i$ is the first player after i to deviate from $\pi^*|_{h_i}$ to π_j replicates continuation play in the σ^* equilibrium if j deviates from $\pi^*_j|_{h_i}$ to π_j following history $\left(h_i, \pi^*_i|_{h_i}, ..., \pi^*_{j-1}|_{h_i}\right)$ in the original equilibrium, so $j \in \{i+1, ..., n\}$ have no incentives to deviate. Clearly, σ' is not one step after any history, but i and h_i were arbitrary, so adjusting σ^* in accordance with (B.4), (B.5) and (B.6) following any history i and h_i we obtain a subgame perfect strategy profile which is one-step after every history h with the same equilibrium outcome. □

Proof of Proposition 2. Proposition 1 implies that for every subgame perfect equilibrium there is an outcome equivalent equilibrium in which strategies are one step for every history, so we assume that σ^* is such a strategy profile. Suppose that there is a sender i and history h_i with associated continuation signal $\pi^*|_{h_i}$ such that there exists some realization $s' \in \pi^*|_{h_i}$ that induces a decision maker posterior belief $\mu' \notin X$ with positive probability. Let $a' = \sigma_D\left(h_i, \pi^*|_{h_i}, ..., \pi^*|_{h_i}\right)$ be the equilibrium action induced by s'. Furthermore, let $M\left(a'\right)$ be the belief polytope where a' is optimal and $X\left(a'\right) = \left\{\mu_j\left(a'\right)\right\}_{j=1}^m$ the set of vertices of $M\left(a'\right)$. Since $M\left(a'\right)$ is the convex hull spanned by $X\left(a'\right)$ there exists $\lambda \in \Delta\left(X\left(a'\right)\right)$ such that $\mu' = \sum_{j=1}^m \lambda_j \mu_j\left(a'\right)$. Consider an alternative one step strategy with $\pi^*|_{h_i}$ replaced by some π' in which the realization s' is replaced by the set $\{s_1, ..., s_m\}$, where each s_j generates posterior $\mu_j\left(a'\right)$ and has unconditional probability $p\left(s'\right)\lambda_j$, and everything else in π' is like the original equilibrium.[12] We also assume that the decision maker follows a strategy

[12] It is possible that $\lambda_j = 0$ for some j. Instead of eliminating these beliefs we may simple generate a probability zero signal in order not to treat this case separately.

in which

$$\sigma'_D(h,s) = \begin{cases} a' & \text{if } h = (h_i, \pi', ...\pi') \text{ and } s \in \{s_1, ..., s_m\} \\ \sigma^*_D(\pi^*|_{h_i}, .., \pi^*|_{h_i}, s) & \begin{array}{c}\text{if } h = (h_i, \pi', ...\pi') \\ \text{and } s \text{ corresponds to some } s \in \pi^* \setminus s\end{array} \\ \sigma^*_D(\pi^*|_{h_i}, .., \pi^*|_{h_i}, \pi_j, ..., \pi_n, s) & \begin{array}{c}\text{if } h = (h_i, \pi', ...\pi', \pi_j, ..., \pi_n) \\ \text{where } j \geq i \text{ is the first player playing } \pi_j \neq \pi'\end{array} \\ \sigma^*_D(h,s) & \text{for any other } h. \end{cases}$$

where σ^*_D is the strategy of the decision maker in the original equilibrium. Since each $\mu_j(a') \in M(a')$ this must be a best response if σ^*_D is a best response. Also, assume the all senders with $j < i$ follow the original equilibrium strategy σ^*_i and that sender $j = \{i, .., n\}$ play

$$\sigma'_j(h_j) = \begin{cases} \pi' & \text{if } h_j = (h_i, \pi', ..., \pi') \\ \sigma^*_i(h_i, \pi^*, .., \pi^*, \pi_k, ..., \pi_{j-1}) & \text{if } h_j = (h_i, \pi', .., \pi', \pi_k, ..., \pi_{j-1}) \\ \sigma^*_i(h_j) & \text{if } h_j = (h_i, \pi_i, ..., \pi_{j-1}) \text{ is such that } \pi_i \neq \pi' \end{cases}, \tag{B.7}$$

and leave everything as in the original equilibrium is h_i is not played by $\{1, ..., i-1\}$ The continuation outcome path is following h_i is then $(\pi', .., \pi')$ and

$$v_i(a,\mu) = \sum_{j=1}^{m} \lambda_j v_n(a, \mu^a_j) = \sum_{j=1}^{m} \lambda_j v_n(\sigma_D(\pi', .., \pi', s_j), \mu^a_j) \tag{B.8}$$

while nothing is changed for signal realizations that are kept like in π^*, so the distribution over states and outcomes is the same as in the original equilibrium if no player deviates after h_i. Moreover, if $j \geq i$ is the first sender deviating from playing π' to π_j the path of play replicates what happens if j is the first sender to deviate from π^* to π_j in the original continuation equilibrium. Hence, there is no profitable deviation on the path. Finally, off-path play replicates off path continuation play in the original equilibrium, so there is no profitable deviation off the path. Repeating the same argument for each history h_i, every continuation signal $\pi^*|_{h_i}$ and every realization $s' \in \pi^*|h_i$ with corresponding belief $\mu' \notin X$ completes the proof. See Figure 7 for an intuitive illustration. □

Proof of Proposition 3. It is immediate from Lemma 1 that n has no profitable deviation following any history unless there is a profitable deviation onto X, so consider sender $i < n$ and history h_i with one step continuation signal $\pi^*|_{h_i}$. Consider a deviation π_i with some induced belief $\mu' \notin X$ given all subsequent senders play sequentially rational which is profitable for i. By Proposition 2

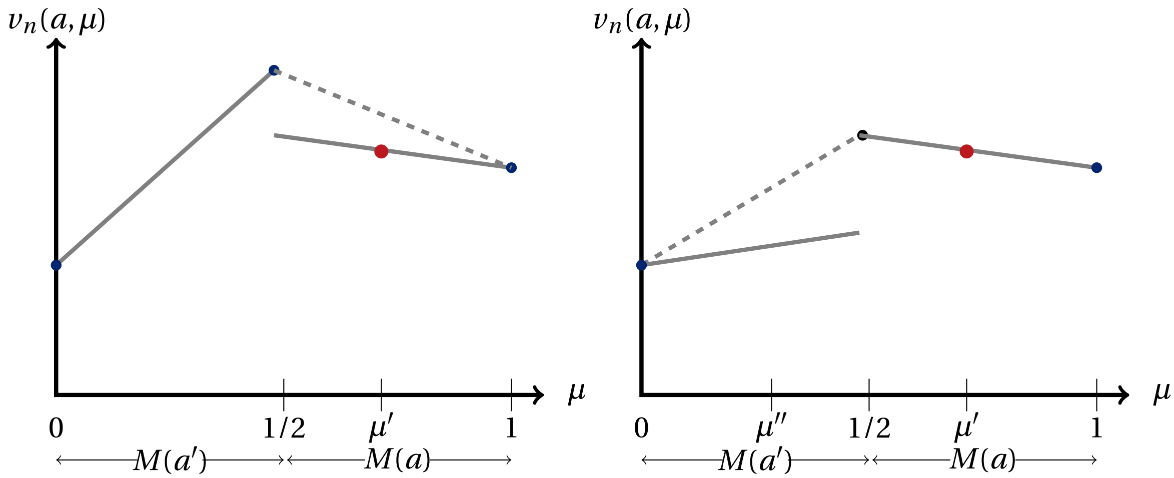

Figure 7: Consider an example with two states ω_0, ω_1 and two actions a, a'. Suppose that $M(a') = [0, 1/2]$ and $M(a) = [1/2, 1]$. Suppose there is an equilibrium where $\mu' \in (1/2, 1)$ is induced and therefore action a is played with positive probability. Then (1) $\sigma_D(1/2) = a$ or $v_n(a, 1/2) > v_n(a', 1/2)$, and (2) $\mu'' \in (0, 1/2)$ is not induced in the equilibrium. Otherwise, sender n has a profitable deviation. On the left panel, we show that, if $v_n(a, 1/2) < v_n(a', 1/2)$, then at μ', sender n has the incentive to deviate by inducing posterior beliefs on 1/2 and a. Notice that we ignore the non-generic case where $v_n(a, 1/2) = v_n(a', 1/2)$. On the right panel, we show that if μ'' is induced in equilibrium, sender n has the incentive to deviate by inducing 0 and 1/2.

(applied to truncated game with i moving first and using one of the beliefs in the interim belief distribution as the prior) there is a continuation one step equilibrium after (h_i, π_i) that generate the same distribution over $\Omega \times A$. Hence, there is a deviation over vertex beliefs and a one step continuation equilibrium that is profitable. □

Proof of Lemma 3. Consider some action a that is taken in equilibrium and some vertex $\mu_j(a) \in X(a)$ and assume that there exists an equilibria σ^* and σ^{**} and histories h^*, h^{**} that generate joint signals π^*, π^{**} with realizations $s^* \in \pi^*$ and $s^{**} \in \pi^{**}$ such that $\mu(s^*) = \mu(s^{**}) = \mu_j(a)$ but that

$$\sigma_D^*\left(h^*, s^*\right) = a \neq a' = \sigma_D^*\left(h^*, s^*\right) \tag{B.9}$$

Suppose first that $\mu_j(a)$ is a vertex of the simplex $\Delta(\Omega)$ then there must be some ω such that

$$v_D(a, \omega) = v_D\left(a', \omega\right). \tag{B.10}$$

A decision maker's Bernoulli payoff function is an element in $|\Omega \times A|$ dimensional Euclidean space and pure each triple $\left(a, a', \omega\right)$ the payoff functions that satisfy (B.10) defines a $|\Omega \times A| - 1$ dimensional subspace. As there is a finite number of triples $\left(a, a', \omega\right) \in A^2 \times \Omega$ the set of bounded Bernoulli payoff functions in which (B.10) holds for some triple (a, a', ω) is of Lebesgue measure zero. Next, consider the case with (B.9) holding at some $\mu_j(a)$ that is not a vertex of the simplex $\Delta(\Omega)$. Then sender n can deviate in a way so that either a or a' is chosen with probability arbitrarily close to one, implying that

$$\sum_{\omega \in \Omega} v_n(a, \omega) \mu_j(a) = \sum_{\omega \in \Omega} v_n(a', \omega) \mu_j(a), \tag{B.11}$$

which again defines a $|\Omega \times A| - 1$ dimensional subspace. Again, this implies that (B.11) can hold only for a set of sender n payoff functions with Lebesgue measure 0. For beliefs in the interior of some $M(a)$ the action a is the unique optimal action for the decision maker and the tie breaking rule on the faces of some $M(a)$ is irrelevant in an equilibrium in which continuation play is always on the vertices. Hence, we may without loss assume that the decision maker follows a Markov strategy. □

Proof of Lemma 4. If $\sigma_D(\mu') = \sigma_D(\mu)$ for each $\mu \in X_{S_n}$ there is nothing to prove. Suppose instead that there exists $\mu \in X_{S_n}$ and $Y \subset X_{S_n}$ and $\tau \in \Delta(Y)$ such that $\mu = \sum_{\mu' \in Y} \mu' \tau(\mu')$ and that (29) is violated for sender n. Then

$$\hat{v}_n(\mu) = \sum_{\mu' \in Y} \hat{v}_n(\mu') \tau(\mu') \tag{B.12}$$

because staying at μ must be weakly preferred to any feasible split of beliefs as otherwise sender n has a strict incentive to split the beliefs, which contradicts that $\mu \in X_{S_n}$.

Let $\mathrm{Co}(Y)$ be the convex hull of Y and $M \leq |\Omega| - 1$ be the dimensionality of $\mathrm{Co}(Y)$. By Carathéodory's theorem it follows that there is a set of $M+1$ points $(\mu_1, ..., \mu_{M+1})$ from Y such that $\mu = \sum_{i=1}^{M+1} \mu_i \tau_i$ for some $\tau \in \Delta^{M+1}$. Moreover, suppose that the M vectors $(\mu_2 - \mu_1, ..., \mu_{M+1} - \mu_1)$ are linearly dependent. Then there are scalars $(\alpha_2, ..., \alpha_{M+1}) \neq (0, ..., 0)$ such that $\sum_{i=2}^{M+1} \alpha_i (\mu_i - \mu_1) = 0$. So

$$\left(-\sum_{i=2}^{M+1} \alpha_i\right) \mu_1 + \sum_{i=2}^{M+1} \alpha_i \mu_i = \sum_{i=1}^{M+1} \alpha_i \mu_i = 0 \tag{B.13}$$

given that $\alpha_1 = -\sum_{i=2}^{M+1} \alpha_i$, which also implies that $\sum_{i=1}^{M+1} \alpha_i = 0$, where we also know that there exists j such that $a_j \neq 0$. Hence, $\alpha_j > 0$ for some j and we thus have that for every β

$$\mu = \sum_{i=1}^{M+1} \mu_i \tau_i = \sum_{i=1}^{M+1} \mu_i \tau_i - \beta \sum_{i=1}^{M+1} \alpha_i \mu_i = \sum_{i=1}^{M+1} (\tau_i - \beta \alpha_i) \mu_i \tag{B.14}$$

Let $I^+ = \{i \in \{1, ..., M+1\} \,|\, \mu_i > 0\}$ and let $j^* \in \arg\min_{j \in I^+} \frac{\tau_j}{\alpha_j}$ and consider $\beta^* = \frac{\tau_{j^*}}{\alpha_j^*}$. Then, let

$$\tau_i^* = \tau_i - \frac{\tau_{j^*}}{\alpha_j^*} \alpha_i. \tag{B.15}$$

It follows that $\tau_i^* \geq 0$ for all i,that $\sum_{i=1}^{M+1} \tau_i^* = 1$ and $\tau_{j^*} = 0$. Hence, we can remove one belief vector from $\left(\mu_1, ..., \mu_{M+1}\right)$ and still find a convex combination that generates μ. By induction, it is without loss of generality to assume that we have $M+1 \leq |\Omega|+1$ vectors $\left(\mu_1, ..., \mu_{M+1}\right)$ such that $\left(\mu_2 - \mu_1, ..., \mu_{M+1} - \mu_1\right)$ are linearly independent. Then we seek to solve

$$\mu = \sum_{i=1}^{M+1} \mu_i \tau_i \Leftrightarrow \mu - \mu_1 = \sum_{i=2}^{M+1} \left(\mu_i - \mu_1\right) \tau_i, \tag{B.16}$$

which is M equations in M unknowns with vectors being linearly independent, so it has a unique solution $\tau_2^*, ..., \tau_{M+1}^*$ where each τ_1^* is positive and $\sum_{i=2}^{M+1} \tau_i^* \leq 1$ because mu is in the convex hull of the other vectors. By letting $\tau_1^* = 1 - \sum_{i=2}^{M+1} \tau_i^*$ we then have that if

$$\hat{v}_n\left(\mu\right) = \sum_{i=1}^{M+1} \hat{v}_n\left(\mu_i\right) \tau_i^* \tag{B.17}$$

and if there are at least two distinct actions, then a utility function satisfying (B.17) belongs to a measure zero set of utility functions.[13] If there is another set of belief vectors $\left(\widetilde{\mu}_1, ..., \widetilde{\mu}_{\widetilde{M}+1}\right)$ that span μ as a convex combination and $\left(\widetilde{\mu}_2 - \widetilde{\mu}_1, ..., \widetilde{\mu}_{\widetilde{M}+1} - \widetilde{\mu}_1\right)$ are linearly dependent we again find that there is a measure zero set of presences that can make sender n indifferent. Since there is a finite set of possibilities to span μ using $\left(\widetilde{\mu}_1, ..., \widetilde{\mu}_{\widetilde{M}+1}\right)$ such that $\widetilde{\mu}_i \in X_{S_n}$ and $\left(\widetilde{\mu}_2 - \widetilde{\mu}_1, ..., \widetilde{\mu}_{\widetilde{M}+1} - \widetilde{\mu}_1\right)$ are linearly dependent and since indifference for a mixture of $\left(\widetilde{\mu}_1, ..., \widetilde{\mu}_{\widetilde{M}+1}\right)$ and $\left(\mu_1, ..., \mu_{M+1}\right)$ is only possible if the sender is indifferent between μ and the convex combination of $\left(\widetilde{\mu}_1, ..., \widetilde{\mu}_{\widetilde{M}+1}\right)$ and μ and the convex combination of $\left(\mu_1, ..., \mu_{M+1}\right)$ we conclude that

$$\hat{v}_n\left(\mu\right) > \sum_{\mu' \in Y} \hat{v}_n\left(\mu'\right) \tau\left(\mu'\right) \tag{B.18}$$

for every $Y \subset X_{S_n}$ and almost all utility functions of the sender given that at least two distinct actions are taken. Since X_{S_n} is finite, the set of utility functions for n that can result in a failure of essential uniqueness is measure zero. Next, consider sender $i < n$ and assume that the conclusion holds for each $j > i$. It is immediate that $X_{S_i} \subset X_{S_{i+1}} \subset .... \subset X_{S_n}$. We can thus repeat exactly the same argument as for sender n to conclude that any failure can occur only for a non-generic set of sender i preferences. The result follows. $\square$

Proof of Lemma 5. Suppose there are two distinct minimal sets of vectors $X\left(\mu\right) \subset X_{S_i}$ and $\widetilde{X}\left(\mu\right) \subset X_{S_i}$ such that

$$\sum_{\mu' \in X(\mu)} \widehat{v}_i\left(\mu'\right) \tau\left(\mu'\right) = \sum_{\mu' \in \widetilde{X}(\mu)} \widehat{v}_i\left(\mu'\right) \widetilde{\tau}\left(\mu'\right). \tag{B.19}$$

[13] Repeating the steps in (B.19), (B.20), we can show that the reduced form payoff functions in "vertex belief space" are non-generic if and only if the Bernoulli payoff functions in over $A \times \Omega$.

where τ is the unique mean preserving spread of μ onto $X(\mu)$ and $\widetilde{\tau}$ is the unique mean preserving of μ onto $\widetilde{X}(\mu)$. Also assume there are at least two distinct actions chosen by the decision maker. In terms of the primitive preferences over $A \times \Omega$ (B.19) can be rewritten as

$$\sum_{\mu' \in X(\mu)} \sum_{\omega \in \Omega} \left[u_i \left(\sigma_D \left(\mu' \right), \omega \right) \mu' (\omega) \right] \tau \left(\mu' \right) = \sum_{\mu' \in \widetilde{X}(\mu)} \sum_{\omega \in \Omega} \left[u_i \left(\sigma_D \left(\mu' \right), \omega \right) \mu' (\omega) \right] \widetilde{\tau} \left(\mu' \right) \tag{B.20}$$

Notice that if for each $a \in A$ we let $X(\mu, a) = \{\mu' \in X(\mu) \text{ s.t } \sigma_D(\mu') = a\}$ then we may rewrite the equality further as

$$\sum_{a \in A} \left\{ \sum_{\omega \in \Omega} u_i (a, \omega) \left[\sum_{\mu' \in X(\mu, a)} \mu' (\omega) \tau \left(\mu' \right) - \sum_{\mu' \in \widetilde{X}(\mu, a)} \mu' (\omega) \widetilde{\tau} \left(\mu' \right) \right] \right\} = 0 \tag{B.21}$$

Since τ and $\widetilde{\tau}$ are unique defined this defines a lower dimensional subspace of $|A \times \Omega|$ −dimensional Euclidean space, so the set of sender n payoff functions such that (B.19) holds is measure zero. Since $X \backslash X_{S_i}$ is finite and the set of possible minimal spanning vectors in X_{S_i} is finite for every $\mu \in X \backslash X_{S_i}$. □

Proof of Proposition 5. Let $X(\mu_0)$ be the support for the unique equilibrium given prior μ_0 and let τ be the associated equilibrium distribution. For contradiction, assume that there exists $\widetilde{\mu}_0 \in \text{Co}(X(\mu_0))$ such that an equilibrium distribution $\widetilde{\tau}$ exists with support $X(\widetilde{\mu}_0) \neq X(\mu_0)$. The argument is identical if $\widetilde{\mu}_0$ is a boundary point (it would just be restricted to a subspace) so we assume without loss of generality that $\widetilde{\mu}_0$ is an interior point in $\text{Co}(X(\mu_0))$. We note that τ and λ are unique vectors so that

$$\mu_0 = \sum_{\mu \in X(\mu_0)} \mu \tau (\mu) \tag{B.22}$$

$$\widetilde{\mu}_0 = \sum_{\mu \in X(\mu_0)} \mu \lambda (\mu). \tag{B.23}$$

Hence, for any β

$$\mu_0 = \sum_{\mu \in X(\mu_0)} \mu \left(\tau (\mu) - \beta \lambda (\mu) \right) + \beta \widetilde{\mu}_0, \tag{B.24}$$

and all coefficients are positive if β is small enough. Also, we assume that $\widetilde{\tau}$ has support on $X(\widetilde{\mu}_0) \neq X(\mu_0)$ so that

$$\widetilde{\mu}_0 = \sum_{\mu \in X(\widetilde{\mu}_0)} \mu \widetilde{\tau} (\mu). \tag{B.25}$$

This implies that when the prior is μ_0 it is feasible to split beliefs over $X(\mu_0) \cup X(\widetilde{\mu}_0)$ in accordance to

$$\{\tau(\mu) - \beta\lambda(\mu) + \beta\widetilde{\tau}(\mu)\}_{\mu \in X(\mu_0) \cup X(\widetilde{\mu}_0)} \tag{B.26}$$

provided that β small enough. But, since τ is the generically unique equilibrium given μ_0, this is suboptimal so

$$\begin{aligned} \sum_{\mu \in X(\mu_0)} \hat{v}_1(\mu)\tau(\mu) &> \sum_{\mu \in X(\mu_0) \cup X(\widetilde{\mu}_0)} \hat{v}_1(\mu)\left[\tau(\mu) - \beta\lambda(\mu) + \beta\widetilde{\tau}(\mu)\right] \\ &= \sum_{\mu \in X(\mu_0)} \hat{v}_1(\mu)\tau(\mu) + \beta\left[\sum_{\mu \in X(\widetilde{\mu}_0)} \hat{v}_1(\mu)\widetilde{\tau}(\mu) - \sum_{\mu \in X(\mu_0)} \hat{v}_1(\mu)\lambda(\mu)\right] \end{aligned}$$

Hence,

$$\sum_{\mu \in X(\widetilde{\mu}_0)} \hat{v}_1(\mu)\widetilde{\tau}(\mu) < \sum_{\mu \in X(\mu_0)} \hat{v}_1(\mu)\lambda(\mu), \tag{B.27}$$

which contradicts that $\widetilde{\tau}$ is better than λ for prior belief $\widetilde{\mu}_0$. □

Proof of Proposition 8. Consider the generic case and suppose that an equilibrium exists in the simultaneous game that is less informative than an equilibrium in the sequential game that cannot be refined further into a Blackwell dominant outcome equivalent equilibrium. Pick some μ in the support of the simultaneous move equilibrium that is in the interior of $\text{Co}(\mu_1, .., \mu_M)$, where $(\mu_1, .., \mu_M)$ are beliefs in X_S in support of the essentially unique equilibrium in the sequential move model. Refining μ to the unique Bayes plausible distribution over $(\mu_1, .., \mu_M)$ creates a Blackwell dominating information structure and either the decisions on each vector in $(\mu_1, .., \mu_M)$ is the same as in μ. Then, splitting beliefs onto $(\mu_1, .., \mu_M)$ gives the same distribution over $A \times \Omega$ as μ, so refining μ to the unique Bayes plausible distribution over $(\mu_1, .., \mu_M)$ is also an equilibrium in the simultaneous case. If there are distinct actions then (Lemma 5) generically some sender has a strict incentive to split the beliefs onto $(\mu_1, .., \mu_M)$ which by appeal to Proposition 6 implies that that sender has a profitable deviation in the simultaneous move game.

□

Proof of Proposition 9. Fix the prior μ_0 and begin by noting that for the result to fail some information must be provided in the sequential model. Hence, without loss there must be a pair $\mu_L, \mu_H \in X$ such that $\mu_L < \mu_0 < \mu_H$ where μ_L and μ_H are in the support of the equilibrium in the sequential model and some μ with $\mu_L < \mu < \mu_H$ that is in the support of an equilibrium in the simultaneous move model. Suppose first that the action is the same at μ_L and μ_H. Then putting positive probability on μ or the unique mean preserving spread onto $\{\mu_L, \mu_H\}$ has no effect on the distribution over actions and states, so putting positive probability on μ doesn't affect the essential informativeness. Suppose that μ_L and μ_H generates distinct actions. Then, for μ to be part of an equilibrium in the simultaneous game all senders must weakly prefer μ to the unique mean preserving spread to $\{\mu_L, \mu_H\}$. But then μ must be an equilibrium (not necessarily on a vertex) in the sequential game, which since μ and the mean preserving spread to $\{\mu_L, \mu_H\}$ generate different distribution over states and action contradicts essential uniqueness. □

References

AMBRUS, A., AND S. TAKAHASHI (2008): "Multi-sender cheap talk with restricted state spaces," *Theoretical Economics*, 3(1), 1–27.

AU, P. H., AND K. KAWAI (2017a): "Competitive Disclosure of Correlated Information," *Working Paper, Nanyang Technological University and the University of New South Wales.*

——— (2017b): "Competitive Information Disclosure by Multiple Senders," *Working Paper, Nanyang Technological University and the University of New South Wales.*

BATTAGLINI, M. (2002): "Multiple referrals and multidimensional cheap talk," *Econometrica*, 70(4), 1379–1401.

BHATTACHARYA, S., AND A. MUKHERJEE (2013): "Strategic information revelation when experts compete to influence," *The RAND Journal of Economics*, 44(3), 522–544.

BOARD, S., AND J. LU (2017): "Competitive Information Disclosure in Search Markets," *Journal of Political Economy, forthcoming.*

BOLESLAVSKY, R., AND C. COTTON (2016): "Limited capacity in project selection: competition through evidence production," *Economic Theory*, pp. 1–37.

DEWATRIPONT, M., AND J. TIROLE (1999): "Advocates," *Journal of political economy*, 107(1), 1–39.

ELY, J., A. FRANKEL, AND E. KAMENICA (2015): "Suspense and surprise," *Journal of Political Economy*, 123(1), 215–260.

ELY, J. C. (2017): "Beeps," *The American Economic Review*, 107(1), 31–53.

GENTZKOW, M., AND E. KAMENICA (2017a): "Bayesian persuasion with multiple senders and rich signal spaces," *Games and Economic Behavior*, 104, 411–429.

——— (2017b): "Competition in persuasion," *The Review of Economic Studies*, 84(1), 300–322.

GLAZER, J., AND A. RUBINSTEIN (2001): "Debates and decisions: On a rationale of argumentation rules," *Games and Economic Behavior*, 36(2), 158–173.

GREEN, J., AND N. STOKEY (1978): "Two Representations of Information Structures and their Comparisons," *Institute for Mathematical Studies in the Social Sciences, Stanford University, technical report no. 271.*

GRÜNBAUM, BRANKOAND KAIBEL, V., V. KLEE, AND G. M. ZIEGLER (1967): *Convex polytopes*, vol. 1967. Springer.

HOFFMANN, F., R. INDERST, AND M. OTTAVIANI (2014): "Persuasion through Selective Disclosure: Implications for Marketing, Campaigning, and Privacy Regulation," *JW Goethe University Frankfurt, UCL, and Bocconi.*

KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian persuasion," *American Economic Review*, 101(6), 2590–2615.

KARTIK, N., F. X. LEE, AND W. SUEN (2016): "A Theorem on Bayesian Updating and Applications to Signaling Games," *Working Paper, Columbia University.*

——— (2017): "Investment in concealable information by biased experts," *The RAND Journal of Economics*, 48(1), 24–43.

KAWAI, K. (2015): "Sequential cheap talks," *Games and Economic Behavior*, 90, 128–133.

KRISHNA, V., AND J. MORGAN (2001): "A model of expertise," *Quarterly Journal of Economics*, 116(2).

LI, F., AND P. NORMAN (2017): "On Bayesian persuasion with multiple senders," *Working Paper, University of North Carolina Chapel Hill.*

LIPNOWSKI, E., AND L. MATHEVET (2017): "Simplifying Bayesian Persuasion," *Working Paper, University of Chicago and New York University.*

MARSCHAK, J., AND K. MIYASAWA (1968): "Economic comparability of information systems," *International Economic Review*, 9(2), 137–174.

MCGUIRE, C. (1959): "Comparisons of Information Structures," *Cowles Foundation Discussion Paper No. 71.*

MILGROM, P., AND J. ROBERTS (1986): "Relying on the information of interested parties," *The RAND Journal of Economics*, pp. 18–32.

RADON, J. (1921): "Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten," *Mathematische Annalen*, 83(1-2), 113–115.

RAYO, L., AND I. SEGAL (2010): "Optimal information disclosure," *Journal of Political Economy*, 118(5), 949–987.

SARVARY, M. (2002): "Temporal differentiation and the market for second opinions," *Journal of Marketing Research*, 39(1), 129–136.

SOBEL, J. (2010): "Giving and receiving advice," in *Econometric Society 10th World Congress.*

WU, W. (2017): "Coordinated Sequential Bayesian Persuasion," *Working Paper, University of Arizona*, 17-18.