Sequential Persuasion

Fei Li† Peter Norman ‡
April 16, 2017

Abstract

This paper considers a general class of multi-sender Bayesian persuasion games in which senders move sequentially. We prove that a Markov perfect equilibrium exists, which is non-obvious because all senders’ payoffs are discontinuous at points of indifference for the decision maker. We also show that for a set of sender payoff functions with Lebesgue measure one the equilibrium is essentially unique in terms of the joint distribution over outcomes and states. Finally, we establish some comparative statics. Adding a sender who makes the first move cannot reduce the informativeness of the equilibrium. Also, sequential persuasion generates less informative equilibria than simultaneous persuasion.

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

JEL Classification Codes: D82, D83

*We thank Laurent Mathevet 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

It is often taken for granted that competition increases the information revealed. Specifically, second, or more generally, multiple opinions have been advocated as a mechanism to reduce bias when advisors have conflicts of interest. For example, a patient whose doctor recommends a drug might visit another doctor if the primary doctor has a consulting relationship with the company producing the drug in question. Home sellers using a realtor may consult a second realtor to obtain a separate appraisal as the primary real estate agent may have an incentive to price properties low to get faster sales. In fact, businesses specializing in selling second opinion services are common in medicine, finance, marketing, and law (see Sarvary (2002) and the references therein).

In many examples when multiple experts are consulted, the process is sequential in nature. The provider of the second opinion knows that the initial opinion triggered a second opinion, and often also the exact nature of the initial advice. For example, medical second opinions tend to be (strategically) sequential when triggered by an internal rule or by an insurance provider. Also, the realtor that is providing a secondary appraisal is at least aware of the existence of another appraisal. More generally, in competition between different news organizations, it is clear that players do respond to information provided by other outlets.

In this paper we consider a sequential model of information disclosure building on the recent advances in multi-sender Bayesian persuasion games studies in Gentzkow and Kamenica (2016, 2017). An uninformed decision maker seeks to maximize her payoff, which, to avoid trivialities, is state dependent. There is also a number of senders that to make things interesting are biased from the point of view of the decision maker.\footnote{A sender sharing the preferences with the decision maker could and would in equilibrium fully reveal the state. This cannot be undone by any other sender, so this case is uninteresting.} Senders move in sequence, each providing an experiment or signal that provides anything in between no information and full information about the state. Each sender observes the experiments designed by previous senders. Surprisingly, it is irrelevant whether or not they also observe the outcomes of the existing experiments, which is a consequence of assuming that signals can be arbitrarily correlated. We will postpone an exact statement of what this means to the main bulk of the paper, but the crucial implication of this is that each sender can refine the pre-existing information structure realization by realization.

Senders have infinite action spaces and the reduced form payoff functions over decision maker beliefs are discontinuous at points where the decision maker switches from one action to another. Moreover, nothing in the model guarantees that senders agree on how the
decision maker should break ties. Existence of equilibria is therefore an issue, and we are aware of no off the shelf existence theorem that applies. The reader may notice that the existence question is trivially solved in the simultaneous move models considered in Gentzkow and Kamenica (2016, 2017) because full revelation is always an equilibrium in these models. That is, 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 case on the other hand, all senders understand that they have influence unless a fully revealing signal has already been played. It follows that full revelation is not guaranteed to be an equilibrium, which we view as a desirable property, while it also implies that we have to solve a non-obvious existence issue.

Fortunately, there is a lot of convexity in the model that allows us to prove an existence result. Using linearity in probabilities, the choice rule by the decision maker can be characterized in terms of intersections of upper half spaces, which are 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 edges 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. Existence of a (Markov perfect) equilibrium can then be established by backwards induction, where in each step senders without loss restrict attention to a finite set of beliefs.

Being able to restrict attention to the vertices of decision maker’s “optimal action areas” in belief space also allows us to prove a strong 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 interior of a polytope that defines optimal choices for the decision maker, but such multiplicity is irrelevant in the sense that the joint distribution over states and actions is the same for all such equilibria. Hence, we are without loss of generality focusing on equilibria in which beliefs are on the (finite set of) vertices of the decision maker’s choice areas. We can then show that for a fixed decision maker payoff function essential uniqueness can fail only for a set of 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.

---

2This insight was used recently by Lipnowski and Mathevet (2017) to simplify the “concaTification” in a Bayesian persuasion game with a single sender. We use the same machinery, but mainly to generate qualitative insights with multiple senders.
A one step equilibrium is an equilibrium in which the first sender is the only sender that provide any non-trivial information. Because signals can be arbitrarily correlated every sender can always unilaterally change the information system implied by the previous senders into any Blackwell-more-informative signal. Hence, if we start with an arbitrary equilibrium and change it so that sender 1 plays the 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 get exactly the same payoff as in the initial equilibrium, and any sender who would have an incentive to deviate in the one step profile would have an incentive to deviate under the initial profile as well.

Given a one step equilibrium, the equilibrium path is a distribution of beliefs with the property 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 stationary. 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 Caratheodory’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.

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 (generically unique) equilibrium is more informative in Blackwell’s order or the two 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.

Our paper relates to a large body of work on information disclosure, but is most di-
rectly 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 (2016, 2017), Au and Kawai (2015), and also Ostrovsky and Schwarz (2010)), but none of these papers deal with sequential moves by the senders. There is also a growing body of work embeds persuasion into dynamic models (Ely, Frankel, and Kamenica (2015) and Ely (2017)), but the only other two papers we are aware of that explicitly considers sequential persuasion are Board and Lu (2015) and Glazer and Rubinstein (2001). Board and Lu (2015) incorporates Bayesian persuasion into a search model. Glazer and Rubinstein (2001) studies a finite horizon studies sequential persuasion model, but they consider a very different information structure.3

2 Model

2.1 The Environment

In this section, we study a general sequential persuasion model and examine the effect of adding senders on information revealed. Specifically, we consider an environment with \( n \geq 1 \) senders and one decision maker. Each sender \( i \in \{1, \ldots, n\} \) has payoff function \( u_i : A \times \Omega \rightarrow \mathbb{R} \) where \( a \in A \) is the action taken by the decision maker and \( \omega \in \Omega \) is the state of the world, where both \( A \) and \( \Omega \) is finite. The decision maker has payoff function \( u_D : A \times \Omega \rightarrow \mathbb{R} \) and there is a common prior belief given by \( \mu_0 \in \Delta (\Omega) \). Payoff functions are common knowledge and players evaluate lotteries using expected utilities.

Every player is 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 (2016) we formalize signals using the partition representation introduced by McGuire (1959) and Marschak and Miyasawa (1968). A generic signal is a finite partition of \( \Omega \times \mathbb{R} \) such that each set in the partition is Lebesgue measurable.

Given partition signal \( \pi = (s_1, \ldots, s_k) \) we assign probabilities as if a sunspot variable \( Z \) is drawn uniformly from \([0, 1] \): the probability of signal \( s_i \) conditional on state \( \omega \) is then \( p(s_i | \omega) = \int_{s \in I_{\omega}(s)} dz \), where a generic signal realization is on form \( s_i = [\omega_0 \times I_{\omega_0}(s_i)] \cup [\omega_1 \times I_{\omega_1}(s_i)] \), where for each state \( \omega \) the collection \( \{I_{\omega_i}(s_i)\}_{i=1}^k \) are disjoint sets such that

3There are also papers in the cheap talk literature that asks what the implications of multiple senders are. See Ambrus and Takahashi (2008), Battaglini (2002), Kawai (2015), Krishna and Morgan (2001), and Milgrom and Roberts (1986)
The posterior probability of state $\omega \in \Omega$ is thus

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

(1)

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

$$
\sum_{s \in S} \mu(\omega|s) p(s) = \mu_0(\omega).
$$

(2)

We let $\Pi$ denote the set of admissible signals.

Signals are partially ordered as one can partition each set in the partition.

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

The space $(\Pi, \succeq)$ is a lattice and we can define the join (and also meet) of two signals in the usual way: for any two signals $\pi, \pi'$, we denote $\pi \vee \pi'$ as 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 $\pi$ and $\pi'$ for any pair of signals $(\pi, \pi')$, which intuitively suggests that combining two signals by taking intersections generates a more informative new signal. Indeed, Green and Stokey (1978) show something even more powerful. Given a signal on partition form we may just let $S$ be any set with the same cardinality as $\pi$ and represent the signal as the pair $(S, p)$ where $p(\cdot|\omega)$ is defined above. We call this the standard representation of a signal. It then follows:

**Lemma 1** (Green and Stokey (1978)). Consider a signal $\pi$ and let $(S', p')$ be associated standard representation. Also, let $(S, p)$ be any signal that is more informative than $(S, p)$ in the sense of Blackwell (1953). Then, $\pi$ can be further partitioned into some $\pi'$ with standard representation $(S', p')$.

That is, not only does a finer partition correspond with a more informative signal in the sense of the less informative being a garbling of the other, but this also says that for anything that is more informative in Blackwell’s sense there exists a a way to partition the coarser into a signal that provides the information of the Blackwell dominating signal. This fact is a central building block in the analysis below.

---

4Some authors replace the unit interval with an arbitrary infinite set together with a probability measure. This generalization does not add anything for our analysis, so we stick with the simpler framework.
2.2 The Extensive Form

Senders move sequentially with sender 1 moving first posting signal \( \pi_1 \), player 2 moving second posting signal \( \pi_2 \), ..., and player \( n \) being the penultimate mover posting signal \( \pi_n \). Then nature draws \( \omega \) 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 \( a \in A \). To summarize, we write \( \Gamma^n = \{\{u_i\}_{i=1}^n, u_D, \Omega, A\} \) for a generic sequential persuasion game.

Each sender observes previous senders’ signals, so player 1 posts a signal following the trivial null-history \( H_1 = \{h_0\} \) and the set of admissible histories for sender \( i \geq 2 \) is \( H_i = \Pi^{i-1} \). A pure strategy for a sender is this a map \( \sigma_i : H_i \to \Pi \). In general, a generic history for the decision maker is a vector \((\pi_1, \ldots, \pi_n, s)\) with \( s = \cap_{i=1}^n s_i \) and \( s_i \in \pi_i \) for each \( i \). However, we will restrict attention to Markov perfect equilibria, meaning that the decision maker’s choice only depends on the posterior belief induced by the signals and the realization, so \( \sigma_D : \Delta(\Omega) \to A \). This is in general restrictive, and we have an example in Appendix A where the decision maker can use a non-Markovian decision rule as carrots and sticks, so that a qualitatively different equilibrium can be supported with non-Markov strategies.\(^5\)

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, and we will use this equilibrium concept to rule out equilibria that are supported by non-credible threats. In addition the restriction to Markov equilibria is convenient, because the payoff relevant history consists of the implied joint signal that the decision maker would use if there is no further information provided. That is, each history \( h_i = \{\pi_j\}_{j=1}^{i-1} \) induces a joint signal, \( \pi^{i-1} = \vee_{j=1}^{i-1} \pi_j \), and the information that sender \( i \) can provide depends only on \( \pi^{i-1} \) and not on the individual signals. Moreover, in a Markov perfect equilibrium all future senders and the decision makers also only care about the implied joint information structure, so we may think of a Markovian strategy as a map \( \sigma_i : \Pi \to \Pi \) with the interpretation that the domain is the set of admissible joint signals created by \( \{1, \ldots, i-1\} \) and that \( \sigma_i (\pi^{i-1}) \) is the joint signal created by taking the intersection of the sets in \( \pi^{i-1} \) and the sets in the individual signal posted by \( i \).

One convenient feature of this relabeling of the strategic variables is that \( \sigma_i (\pi^{i-1}) \supseteq \pi^{i-1} \), so that any strategy profile (equilibrium or not) generates a monotone sequence of signals on and off the outcome path in the sense that later partitions are always finer.

---

\(^5\)However, we show that equilibria are generically unique, so such examples are knife edge.
3 Results

3.1 Existence of Equilibria

The sequential persuasion model defined in Section 2 is a game of perfect information, so equilibria can be computed by backwards induction. However, each player has an infinite action space, so off-the-shelves existence results don’t apply. Moreover, the payoff functions of senders as a function of posterior beliefs are generically discontinuous at points of indifference for the decision maker. Since there is no reason for different senders to agree on which direction ties should be broken, it appears that existence cannot be guaranteed.\(^6\) However, we are able to demonstrate that equilibria exist, even if we restrict attention to Markovian strategies:

**Theorem 1.** There exists at least one Markov perfect equilibrium in any sequential persuasion game \(\Gamma^n = \{\{u_i\}_{i=1}^n, u_D, \Omega, A\}\) satisfying the assumptions in Section 2.

The proof of this result combines results from convex analysis, linear programming, and the Zermelo-Kuhn backwards induction algorithm. We begin by noting that we can char-

\(^6\)The same issue occurs in simultaneous persuasion games, but there existence questions can be avoided because full revelation is always an equilibrium. With sequential moves, full revelation is not always an equilibrium.
acterize an optimal choice rule for the decision maker in terms of “decision areas” that are convex polytopes. That is, 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

\[ H(a \succeq a') \equiv \{ \mu \in \Delta(\Omega) | \sum_{\omega \in \Omega} \mu(\omega)[u_D(a, \omega) - u_D(a', \omega)] \geq 0 \}. \] (3)

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

\[ M(a) = \cap_{a' \in A} H(a \succeq a'). \] (4)

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

Next, we consider the last sender’s best response. For each \( a \in A \) define \( v_n(a, \mu) \) as the expected utility for sender \( n \) given action \( a \) and belief \( \omega \)

\[ v_n(a, \mu) \equiv \sum_{\omega \in \Omega} u_n(a, \omega) \mu(\omega). \] (5)

In any equilibrium, ties must be broken in favor of sender \( n \) when the decision maker takes an action. Hence, we consider a decision rule satisfying \( \sigma_D(\mu) = a \) if \( \mu \in M(a) \) and \( v_n(a, \mu) \geq v_n(a', \mu) \) for each \( a' \neq a \). Also, let \( \pi^{n-1} = \bigvee_{j=1}^{n-1} \pi_j \) be the joint signal from combining the signals of senders \( 1, \ldots, n-1 \) and let \( \tau^{n-1} \) be the corresponding distribution of posteriors beliefs whose support is denoted by \( \text{Supp}(\tau^{n-1}) \). By the standard argument in Kamenica and Gentzkow (2011), the best response problem for sender \( n \) can then be expressed in terms of choosing conditional distributions of beliefs that satisfy Bayes plausibility. This is a linear programming problem that can be written as

\[
\max_{\lambda} \sum_{\mu \in \text{Supp}(\tau^{n-1})} \left[ \sum_{\mu' \in \Delta(\Omega)} v_n(\sigma_D(\mu'), \mu') \lambda(\mu', \mu) \right] \tau^{n-1}(\mu)
\]

s.t. \( \sum_{\mu'} \mu' \lambda(\mu', \mu) = \mu \) for each \( \mu \in \text{Supp}(\tau^{n-1}) \),

\( ^7 \)See Grünbaum, Klee, and Ziegler (1967).

\( ^8 \)In the remainder of this paper, we will use best response partition \( \pi_i^* \) and best response distribution of posterior belief \( \tau_i^* \) exchangeable. This is without loss of generality in the construction of a Markov equilibrium: since the decision maker’s best response is Markovian, what matters is the distribution of the posterior belief that he faces.
where \( \lambda(\mu', \mu) \) is a (conditional) distribution of posterior beliefs the decision maker faces if the realization \( s^{n-1} \in \pi^{n-1} \) would generate belief \( \mu \) provided that \( n \) doesn’t provide any extra information. But, (6) can be solved by solving the linear program

\[
V_n(\mu) = \max_{\tau \in \Delta(\Delta(\Omega))} \sum_{\mu'} v_n(\sigma_D(\mu'), \mu') \tau(\mu')
\]

s.t. \( \sum_{\mu'} \mu' \tau(\mu') = \mu \) for every \( \mu \in \text{Supp}(\tau^{n-1}) \).

Notice that \( \sigma_D(\cdot) \) breaks all ties in favor of sender \( n \). Consider a restricted linear program

\[
\tilde{V}_n(\mu) = \max_{\tau \in \Delta(X)} \sum_{\mu' \in X} v_n(\sigma_D(\mu'), \mu') \tau(\mu')
\]

s.t. \( \sum_{\mu'} \mu' \tau(\mu') = \mu \) for every \( \mu \in \text{Supp}(\tau^{n-1}) \),

where \( X = \bigcup_{a \in A} \{ \mu_j \}_{j=1}^{J(a)} \) is the set of all vertices that defines the optimal actions for the decision maker, which is finite because both \( \Omega \) and \( A \) are finite. Hence, the linear program in (8) is well defined as it is a finite dimensional bounded linear program.

**Lemma 2.** For each \( \mu \in \Delta(\Omega) \) the linear program (7) is well defined. Moreover:

1. \( V_n(\mu) = \tilde{V}_n(\mu) \) for each \( \mu \in \Delta(\Omega) \);

2. Every solution to (8) given belief \( \mu \) also solves (7);

3. For every solution \( \tau_n \in \Delta(\Delta(\Omega)) \) to (7) there exists some \( \tilde{\tau}_n \in \Delta(X) \) that solves (8) that generates the same probability distribution over \( A \times \Omega \).

A detailed 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 \( \mu \) 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 \( \mu \) 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 (8) generates a utility at least as great as (7). But, the feasible set in (8) is a subset of the feasible set in (7), so the two problems must have the same value.

We can now complete the proof of Theorem 1:
Proof. By Lemma 2, for any \( \mu \) there exists an optimal solution to the program (7) where sender \( n \) puts positive probability on \( \mu' \) only if \( \mu' \in \mathcal{J}(a) \) for some \( a \in A \) and \( v_n(a, \mu_a^a) \geq v_n(a', \mu_a^{a'}) \) for every \( a' \in A \). Denote \( \tau_n^*(\mu) \in \Delta(X) \) as a Markovian best reply by sender \( n \) which is defined as the solution to problem (8) for a given \( \mu \). Define

\[ X_n = \cup_{\mu \in \Delta(\Omega)} \text{Supp}(\tau_n^*(\mu)), \]

which is obviously finite and a subset of \( X \). That is, for each \( \mu \notin X_n \), sender \( n \) finds it optimal to “split” it into beliefs in \( X_n \) s.t. the Bayes-plausible condition.

Using the same argument as when breaking up (6) into separate problems for each \( \mu \), we can then find the best reply for sender \( n - 1 \) by solving the linear program

\[ V_{n-1}(\mu) = \max_{\tau \in \Delta(\Omega)} \sum \mu' \left[ \sum \mu'' v_{n-1}(\sigma_D(\mu''), \mu'' \tau_n^*(\mu''; \mu')) \right] \tau(\mu'), \tag{9} \]

s.t. \( \sum \mu' \tau(\mu') = \mu \),

where \( \{\tau_n^*(\mu''; \mu')\}_{\mu'' \in X_n} \) is conditional probability measure defined by the best response of sender \( n \). But, then for each \( \mu'' \in X_n \) the probability that it is eventually realized is \( \sum_{\mu'} \tau_n^*(\mu''; \mu') \tau(\mu') \). Hence, if \( \mu' \notin X_n \), sender \( n - 1 \) knows exactly how these beliefs will be “split” over \( X_n \) is equilibrium. There, replacing \( \mu' \) with the distribution over \( X_n \) that \( n \) will respond with and leaving everything else the same gives \( n - 1 \) the same payoff. Hence, we may without loss consider the problem for sender \( n - 1 \) a finite dimensional bounded linear program where \( n - 1 \) optimizes over \( \tau \in \Delta(X_n) \). This problem has a solution. Fixing an optimal solution for \( n - 1 \) we can make the same argument for senders \( n - 2, \ldots, 1 \).

It is straitforward to construct senders’ equilibrium strategy \( \{\sigma_i\}_{i=1}^n \) from \( \{\tau_i\}_{i=1}^n \). Hence, existence of an equilibrium has been proved.

As a byproduct of our existence argument we also note that it is irrelevant whether sender \( i \) can observe a joint realization \( s^{i-1} \in \pi^{i-1} \) or not.

Corollary 1. Let \( (\sigma_1, \ldots, \sigma_n, \sigma_D) \) be a Markov perfect equilibrium in the sequential model with coordinated signals defined in Section 2. Then it is also an equilibrium of a game where nature draws the state \( \omega \) 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 broken into a number of problems, one for each belief in the support of the distribution generated by the joint signals of previous senders. This is how we justified using (7) instead of (6), (9) instead of the analogue to (6) for.
sender $n-1$ and so on. This, in turn 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 $\pi_i(s^{i-1})$ define a contingent partition of $s^{i-1}$ and note that the state dependent probability of signal $s^i \in \pi_i(s^{i-1})$ 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 $\pi^{i-1}$ or 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 clearly irrelevant as the way player $i$ maximizes the expected payoff is to maximize each component in the sum.

But then we further note that if we see the signal and the realization and act upon that 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 may be arbitrarily correlated as saying that experiments are conducted sequentially with experimental results revealed in each stage. In Board and Lu (2015) 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.

Corollary 1 also buys us the equivalence between the partition and belief representations of information systems that is central to the analysis in Kamenica and Gentzkow (2011). Because each sender can condition on any $s^{i-1} \in \pi^{i-1}$ the problem for sender $i$ is solved by solving $|\pi^{i-1}|$ separate problems, each of them with the same qualitative features as a single sender persuasion model. One can therefore characterize each senders’ equilibrium signal choice by using the convexification approach used in Aumann and Maschler (1968) and Kamenica and Gentzkow (2011): senders make realization contingent choices, allowing sender $i$ to split the interim belief $\mu(s^{i-1})$ in an arbitrary way for every $s^{i-1} \in \pi^{i-1}$. This fact was used heavily in our existence argument.

### 3.2 Generic Uniqueness

An interesting feature of the sequential structure that makes comparison across equilibria simpler is that we can always focus on equilibria where the first mover immediately implements the equilibrium information structure
Definition 2. A Markov perfect equilibrium is one step if \( \sigma_i (\sigma_1 (h_0)) = \sigma_1 (h_0) \) for each sender \( i \in \{1, \ldots, n\} \).

That is, in a one step equilibrium the information structure facing the decision maker is chosen by the first sender, and, on the equilibrium path, senders add redundant or no information at all. We say that two strategy profiles are outcome equivalent if they generate identical joint distribution of posterior belief and state and they are observationally equivalent if they generate identical joint distribution over \( \Omega \times A \). Apparently, two outcome equivalent strategy profiles must be observationally equivalent.

Lemma 3. Every Markov perfect equilibrium has an outcome equivalent one step Markov perfect equilibrium.

Proof. Fix a Markov perfect equilibrium \((\sigma, \sigma_D)\) and let \( \pi_1^* = \sigma_1 (h_0) \) and \( \pi_i^* = \sigma_i (\pi_{i-1}^*) \) for each \( i \in \{2, \ldots, n\} \) so that the equilibrium path signals are \( (\pi_1^*, \ldots, \pi_n^*) \). Consider the alternative strategy profile \((\hat{\sigma}, \hat{\sigma_D})\) where \( \sigma_1 (h_0) = \pi_n^* \) and

\[
\hat{\sigma}_i (\pi) = \begin{cases} 
\pi_n^* & \text{if } \pi = \pi_n^* \\
\sigma_i (\pi) & \text{if } \pi \neq \pi_n^* 
\end{cases}
\]

for \( i = 2, \ldots, n \) and \( \hat{\sigma}_D (\pi) = \sigma_D (\pi) \) for each \( \pi \in \Pi \). We note that the implied payoffs for all players are the same if playing \((\hat{\sigma}, \hat{\sigma_D})\) as when playing \((\sigma, \sigma_D)\). Also, there is no change in the strategy used by the decision maker or the off-the equilibrium path strategies used by the senders, so \((\hat{\sigma}, \hat{\sigma_D})\) is an equilibrium if and only if we can show that there is no deviation on the equilibrium path. Suppose such a deviation exists, so that playing \( \pi' \) is strictly better than \( \pi_n^* \) for some sender \( i \in \{1, \ldots, n\} \). But, for each sender \( i \in \{2, \ldots, n\} \) we have that \( \pi_n^* \) is a finer partition than \( \pi_i^* \), so \( \pi' \) is a feasible deviation in the original equilibrium. Since continuation play following \( \pi' \) is the same under \((\sigma, \sigma_D)\) it must be that \( \pi' \) is a profitable deviation from \( \sigma_i \) for some \( i \) contradiction that \((\sigma, \sigma_D)\) is a (even a Nash) equilibrium. Finally, sender 1 has no constraints on what signals to play, so if sender 1 has a profitable deviation from \( \pi_n^* \) there is a profitable deviation from \( \sigma_1 (h_0) \), again contradicting \((\sigma, \sigma_D)\) being an equilibrium.

\[\square\]

Notice that Lemma 3 is due to the fact that a sender can only provide further information rather than “reverse” the signals provided by previous senders. In addition, we can show that it is without loss to further restrict attention on a smaller set of equilibria.

Corollary 2. Every Markov perfect equilibrium has an observationally equivalent one step Markov perfect equilibrium where the resulting posterior belief is only realized in \( X \).
One of our main result shows that all Markov perfect equilibria are observationally equivalent.

**Definition 3.** We say that the equilibrium is **essentially unique** if all equilibria are observationally equivalent.

**Definition 4.** Given some $\mu \in \Delta(\Omega)$ we say that a set of spanning vectors $Y$ is **minimal** if there is a unique $\lambda \in \Delta(Y)$ such that $\mu = \sum_{\mu' \in Y} \lambda(\mu') \mu'$.

When the spanning vectors are not minimal, one can always reduce some of the “redundant” vectors as the convex combination of other vectors, which obviously makes the expression non-unique.

**Theorem 2.** Fix any $\mu_0$. Then there is an essentially unique equilibrium that induces distribution $\tau$ over a minimal set of spanning vectors $X(\mu_0)$ for any fixed preferences of the decision maker and almost all sender preferences.

To prove Theorem 2, we need more intermediate results.

**Definition 5.** A belief $\mu \in \Delta(\Omega)$ is **stable** if there exists an equilibrium such that no information is provided by any sender given prior belief is $\mu$.

We then have that:

**Lemma 4.** Every Markov equilibrium generates a distribution of posterior beliefs for the decision maker such that

1. each belief in its support is stable, and
2. there is no other distribution of stable beliefs that is preferred by sender 1.

**Proof.** Every equilibrium is equivalent to a one step equilibrium by Lemma 3. Since no sender can further refine each signal in its support the distribution must be over stable beliefs, and sender 1 will obviously pick the distribution optimally.

Recall that there is a finite set of edges $X$ corresponding to the convex polytopes that define the optimal choices for the decision maker. Since we know that it is without loss to consider equilibria where the induced belief distribution is over $X$ only we now restrict attention to the intersection between the set of stable beliefs and $X$. We call this set $X_S$ and note that it is non-empty as it contains the edges of the simplex $\Delta(\Omega)$. We now note that only for rare payoffs is it ever the case that there is a weak incentive for any sender to move away from a stable belief.
**Lemma 5.** Let \( X_S \subset X \) be the (finite) set of stable beliefs contained in the vertices of the decision maker’s and pick any \( \mu \in X_S \). Then,

\[
v_i(\mu) > \sum_{\mu' \in X \setminus \{\mu\}} v_i(\mu') \tau(\mu')
\]

for almost all sender utility functions and every \( i \in \{1, ..., n\} \) and every \( \tau \in \Delta(X_S \setminus \{\mu\}) \) such that \( \mu = \sum_{\mu' \in X \setminus \{\mu\}} \mu' \tau(\mu') \).

The idea is that if some sender has a weak incentive to split a stable belief \( \mu \) into a distribution over \( X_S \setminus \{\mu\} \) then it must be that

\[
v_i(\mu) = \sum_{\mu' \in X \setminus \{\mu\}} v_i(\mu') \tau(\mu')
\]

for some sender \( i \) and some \( \tau \) that satisfies Bayes plausibility. Since \( X_S \) must contain the vertices of \( \Delta(\Omega) \) the dimensionality of the convex hull of \( X_S \setminus \{\mu\} \) must be the same as the dimensionality of the state space. It follows from Carathéodory’s theorem that \( \mu \) can be spanned by \(|\Omega| + 1\) vectors in \( X_S \setminus \{\mu\} \). Depending on whether \( \mu \) is in the interior of \( \text{CO}(X_S \setminus \{\mu\}) \) or on the boundary a set of \(|\Omega| + 1\) spanning vectors may or may not be linearly dependent. However, if there are linear dependencies we can eliminate vectors until we get \( \mu \) spanned by a set of linearly independent vectors in \( X_S \setminus \{\mu\} \).

In order for sender \( i \) to be indifferent between staying at \( \mu \) and some distribution over \( X_S \setminus \{\mu\} \) there must be (this is by linearity in probabilities) some minimal spanning vector for \( \mu \) that creates indifference. But, the indifference condition for any particular minimal spanning vector is satisfied with probability zero as the indifference condition defines a lower dimensional hyperplane in the space of (bounded) utility functions. Since there is a finite number of possible minimal spanning vectors and a finite number of senders the result follows by induction. See Appendix B for details.

**Proof of Theorem 2.** By Lemma 3, fixing any prior \( \mu_0 \in \Delta(\Omega) \) sender 1 has a choice over distributions over \( X_S \) that are Bayes plausible. By Lemma 5 it follows that for generic preferences once sender 1 has chosen a distribution over \( X_S \), then the essentially unique equilibrium conditional on this choice is not to add any information. Consequently, if there are multiple equilibria, sender 1 must be indifferent. But, if two Bayes plausible distributions over \( X_S \) generates the same expected payoff for 1 then there exists minimal spanning vectors \( X(\mu_0) \) and \( \bar{X}(\mu_0) \) and (relative to the set of spanning vectors) unique Bayes plausible distributions \( \tau \) (for \( X(\mu_0) \)) and \( \bar{\tau} \) (for \( \bar{X}(\mu_0) \)) such that

\[
\sum_{\mu' \in X(\mu_0)} v_1(\mu') \tau(\mu') = \sum_{\mu' \in \bar{X}(\mu_0)} v_1(\mu') \bar{\tau}(\mu') ,
\]
which can only hold for a set of (bounded) utility functions of measure zero because the support of both $X(\mu_0)$ and $\tilde{X}(\mu_0)$ is minimal. Since there is a finite set of vectors in $X_S$ the result follows.

Our final characterization result says that if prior $\mu_0$ generates a unique equilibrium with minimal support $X(\mu_0)$ (which is generically true), then the unique equilibrium given a prior $\tilde{\mu}_0$ in the convex hull of $X(\mu_0)$ is the unique Bayes plausible distribution over $X(\mu_0)$ given prior $\tilde{\mu}_0$.

**Proposition 1.** Suppose that the essentially unique equilibrium given prior $\mu_0$ is (the unique) distribution $\tau$ over (minimal) support $X(\mu_0)$. Then, for every $\tilde{\mu}_0 \in \text{Co}(X(\mu_0))$ every equilibrium outcome is outcome equivalent with the unique Bayesian plausible belief distribution $\lambda$ over $X(\mu_0)$ for prior $\tilde{\mu}_0$.

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

$$
\mu_0 = \sum_{\mu \in X(\mu_0)} \mu \tau(\mu)
$$

$$
\tilde{\mu}_0 = \sum_{\mu \in X(\mu_0)} \mu \lambda(\mu).
$$

Hence, for any $\beta$

$$
\mu_0 = \sum_{\mu \in X(\mu_0)} \mu (\tau(\mu) - \beta \lambda(\mu)) + \beta \tilde{\mu}_0,
$$

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

$$
\tilde{\mu}_0 = \sum_{\mu \in X(\tilde{\mu}_0)} \mu \tilde{\tau}(\mu).
$$

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

$$\{\tau(\mu) - \beta \lambda(\mu) + \beta \tilde{\tau}(\mu)\}_{\mu \in X(\mu_0) \cup X(\tilde{\mu}_0)}$$
provided that $\beta$ small enough. But, this is suboptimal so
\[
\sum_{\mu \in X(\mu_0)} v(\mu) \tau(\mu) > \sum_{\mu \in X(\mu_0) \cup X(\tilde{\mu}_0)} v(\mu) [\tau(\mu) - \beta \lambda(\mu) + \beta \bar{\tau}(\mu)]
\]
\[
= \sum_{\mu \in X(\mu_0)} v(\mu) \tau(\mu) + \beta \left[ \sum_{\mu \in X(\tilde{\mu}_0)} v(\mu) \bar{\tau}(\mu) - \sum_{\mu \in X(\mu_0)} v(\mu) \lambda(\mu) \right]
\]
Hence,
\[
\sum_{\mu \in X(\tilde{\mu}_0)} v(\mu) \bar{\tau}(\mu) < \sum_{\mu \in X(\mu_0)} v(\mu) \lambda(\mu),
\]
which contradicts that $\bar{\tau}$ is better than $\lambda$ for prior belief $\tilde{\mu}_0$. \qed

4 Comparative Statics

4.1 Adding Additional Experts

Consider a generic sequential persuasion game with $n$ senders, $\Gamma^n = \{\{u_i\}_{i=1}^n, u_D, \Omega, A\}$. We now consider the effects of adding an additional sender, who moves before senders $1, ..., n$, while keeping everything else as in $\Gamma^n$. To avoid introducing additional notation we do this by considering a persuasion game with $n+1$ senders given by $\Gamma^{n+1} = \{\{u_i\}_{i=0}^n, u_D, \Omega, A\}$ where we add sender $i = 0$ who moves before sender 1, while everything else is as in game $\Gamma^n$. Using the one step equilibrium characterization we can show:

**Proposition 2.** Consider the generic case with unique equilibria. Then, if a sender is added who moves before all other senders, either the equilibria are not comparable under Blackwell’s ordering or the equilibrium with the extra sender is more informative.

**Proof.** Suppose not. Let $X_n(\mu_0)$ be the set of spanning vectore of the unique equilibrium with $n$ senders. Then:

1. there exist at least one belief $\mu'$ in the support of the equilibrium with $n+1$ senders that can be split into $X_n(\mu_0)$.

2. there exists no belief $\mu''$ in the support of the equilibrium with $n+1$ senders that is not in the convex hull of $X_n(\mu_0)$.

But, using Lemma 5 every $i \in \{1, ..., n\}$ is strictly better off to split belief $\mu'$ into a distribution over $X_n(\mu_0)$ for generic preferences. The proposition follows because we may always restrict attention to one step equilibria. \qed
It is important to remember that the order of moves matter. If an additional sender is placed in any other position than first, the result may fail, as is demonstrated explicitly but the example in Section 2. However, a decision maker that can control what experts to consult could also control the order, and would therefore avoid putting experts that would reduce the information from earlier movers at the end. We also note that if equilibria are unique, then Proposition 2 guarantees that adding senders at the root cannot lead to unambiguously worse information when adding a sender at the root. At worst, the information systems would not be comparable using Blackwell’s order.

4.2 Simultaneous vs Sequential Persuasion

Finally we note that an equilibrium in a simultaneous persuasion game with \( n \) senders has the exact same equilibrium characterization, except that each sender can refine any \( \mu \) in the support of the equilibrium into any Bayes plausible distribution. In contrast, the sequential characterization only allow players earlier in the game to refine beliefs into Bayes plausible distribution that senders later in the game don’t want to further refine. We this conclude:

Proposition 3. Consider the generic case in which the sequential game has a unique equilibrium. Then, there exists no equilibrium in the simultaneous game that is less informative than the equilibrium in the sequential game.

Proof. If for contradiction an equilibrium in the simultaneous game that is less informative would exist, then each sender could unilaterally deviate to the outcome generated in the equilibrium of the sequential game. With generic preferences the sequential game has a unique minimal set of spanning vectors, so using the same argument as in Proposition 1 any belief that is spanned by the minimal set of spanning vectors should be strictly worse than the unique Bayes plausible distribution over the spanning vectors.

This result is in stark contrast to Glazer and Rubinstein (2001) who, in a related model, find that a sequential debate is more informative that a simultaneous debate. The crucial difference between the models is that the senders in Glazer and Rubinstein (2001) have more constraints on how to generate information than in our model.
A Appendix: Non-Markovian Equilibrium

In this section, we consider an example which has a non-Markovian equilibrium that is qualitatively different from the Markov Equilibrium. Suppose that $\Omega = \{\omega_0, \omega_1\}$ and the optimal choice correspondence for the decision maker is

$$\sigma (\mu) = \begin{cases} a_1 \text{ or } a_2 & \text{if } \mu \leq 0.1 \\ a_3 & \text{if } 0.1 \leq \mu \leq 0.9 \\ a_4 \text{ or } a_5 & \text{if } \mu \geq 0.9 \end{cases}$$

Also suppose that two sender have state independent preferences

$$u_1 (a, \omega) = \begin{cases} 3 & \text{if } a \in \{a_1, a_4\} \\ 1 & \text{if } a = a_3 \\ 0 & \text{if } a \in \{a_2, a_5\} \end{cases}$$

$$u_2 (a, \omega) = \begin{cases} 3 & \text{if } a \in \{a_2, a_5\} \\ 1 & \text{if } a = a_3 \\ 0 & \text{if } a \in \{a_1, a_4\} \end{cases}$$

Consider a Markovian 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 are

$$\frac{3\sigma_1 (0) + \sigma_4 (1)}{2}$$

for sender 1

$$\frac{3\sigma_1 (0) - \sigma_4 (1)}{2}$$

for sender 2,

so the payoff is greater than or equal to $\frac{3}{2}$ for at least one sender. Hence, beliefs in $[0.1, 0.9]$ 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 $[0, 0.1]$ or $[0.9, 1]$ each sender may as well not provide any information and qualitatively different equilibria with action $a_3$ can be supported by such non-Markovian strategies.
B Appendix: Omitted Proofs

B.1 Proof of Lemma 2

\textit{Proof.} Pick any feasible solution \( \tau \) to program (7). For each \( a \in A \) write \( \tau^a (\mu') \) for \( \mu' \) such that \( \sigma_D (\mu') = a \) so that we may write \( \tau = \{ \{ \tau^a (\mu') \}_{\mu' \in \tilde{M}_a} \}_{a \in A} \) where \( \tilde{M}_a = \{ \mu \in \Omega | \sigma_D (\mu) = a \} \) is the “decision area” of action \( a \) defined by \( \sigma_D (\cdot) \). Obviously, \( \tilde{M}_a \subseteq M_a, \forall a \).

Since every \( \mu' \in M_a \) there exists \( \lambda' \in \Delta \left( \{ \mu_j^a \}_{j=1}^{J(a)} \right) \) such that \( \mu' = \sum_{j=1}^{J(a)} \lambda_j^a \mu_j^a \). For every \( a \in A \) and \( \mu_j^a \) spanning \( M_a \) let \( \tilde{\tau} (\mu_j^a) = \sum_{\mu' \in \tilde{M}_a} \tau (\mu') \lambda_j^a \) so that

\[
\sum_{j=1}^{J(a)} \tilde{\tau} (\mu_j^a) = \sum_{\mu' \in \tilde{M}_a} \tau (\mu') \sum_{j=1}^{J(a)} \lambda_j^a = \sum_{\mu' \in \tilde{M}_a} \tau (\mu').
\]

Since it is possible that \( v_n (a, \mu_j^a) < v_n (a', \mu_j^a) \) for some \( \mu_j^a \in M_a \) (and \( \mu_j^a \notin \tilde{M}_a \), because breaking the tie in favor of \( a' \) may be better than \( a \)) it thus follows that the optimal solution to (8) satisfies

\[
\tilde{V}_n (\mu) \geq \sum_{a \in A} \sum_{j=1}^{J(a)} v_n (a, \mu_j^a) \tilde{\tau} (\mu_j^a) = \sum_{a \in A} \sum_{j=1}^{J(a)} \sum_{\omega \in \Omega} u_n (a, \omega) \mu_j^a (\omega) \tilde{\tau} (\mu_j^a)
\]

\[
= \sum_{a \in A} \sum_{\omega \in \Omega} u_n (a, \omega) \left[ \sum_{j=1}^{J(a)} \tilde{\tau} (\mu_j^a) \lambda_j^a \right] \left[ \sum_{\mu' \in \tilde{M}_a} \tau (\mu') \right] = \sum_{a \in A} \sum_{\omega \in \Omega} u_n (a, \omega) \mu' \left[ \sum_{\mu' \in \tilde{M}_a} \tau (\mu') \right]
\]

\[
= \sum_{\mu'} v_n (\sigma_D (\mu'), \mu') \tau (\mu').
\]

This holds for any feasible solution to (7). Hence, \( \tilde{V}_n (\mu) \geq V_n (\mu) \). Moreover, any optimal solution to (8) is a feasible solution to (7), so \( \tilde{V}_n (\mu) \leq V_n (\mu) \). This establishes that there exists solutions to (7) and that \( \tilde{V}_n (\mu) = V_n (\mu) \) and that every \( \tilde{\tau}_n \in \Delta (X) \) that solves (8) given belief \( \mu \) also solves (7). Finally, if \( \tau \) solves (7) and \( \mu' \) is such that \( \tau_n (\mu') > 0 \) there can be no \( \mu_k^a \in M^a \) such that \( v_n (a, \mu_k^a) < v_n (a', \mu_k^a) \) and \( \lambda_k^a > 0 \) for the weight on vector \( \mu_k^a \) in the convex combination such that \( \mu' = \sum_{j=1}^{J(a)} \lambda_j^a \mu_j^a \). This is seen from noting that this would generate a strict inequality in the first inequality of (11). \( \square \)

B.2 Proof of Corollary 2

TBD
B.3 Proof of Lemma 5

Proof. Suppose that there exists $\mu \in X_S$ and $\tau \in \Delta (X_S \setminus \{\mu\})$ such that $\mu = \sum_{\mu' \in X_S \setminus \{\mu\}} \mu' \tau (\mu')$ such that (10) is violated for sender $n$. Then

$$v_n (\mu) = \sum_{\mu' \in X_S \setminus \{\mu\}} v_n (\mu') \tau (\mu')$$

because staying at $\mu$ must be weakly preferred to any feasible split of beliefs as otherwise sender $n$ has a strict incentive to split the beliefs, which contradicts $\mu$ being a stable belief. Let $CO(X_S \setminus \{\mu\})$ be the convex hull of $X_S \setminus \{\mu\}$ and $M$ be the dimensionality of $CO(X_S \setminus \{\mu\})$. By Carathéodory’s theorem it follows that there is a set of $M + 1$ points $(\mu_1, ..., \mu_{M+1})$ from $X_S \setminus \{\mu\}$ such that $\mu = \sum_{i=1}^{M+1} \alpha_i \mu_i$ for some $\tau \in \Delta^{M+1}$. Since $X_S$ contains the vertices of $\Delta (\Omega)$ and since these vertices cannot be split further it is without loss to assume that $M = |\Omega|$. 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

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

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 $\beta$

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

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

$$\tau_i^* = \tau_i - \frac{\tau_{j^*}}{\alpha_{j^*}} \alpha_i.$$ 

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 $(\mu_1, ..., \mu_{M+1})$ and still find a convex combination that generates $\mu$. By induction we thus know that it is without loss of generality that we have $M + 1 \leq |\Omega| + 1$
vectors \((\mu_1, \ldots, \mu_{M+1})\) such that \((\mu_2 - \mu_1, \ldots, \mu_{M+1} - \mu_1)\) are linearly independent. Then we seek to solve

\[
\mu = \sum_{i=1}^{M+1} \mu_i \tau_i \iff \mu - \mu_1 = \sum_{i=2}^{M+1} (\mu_i - \mu_1) \tau_i,
\]

which is \(M\) equations in \(M\) unknowns with vectors being linearly independent, so it has a unique solution \(\tau_2^*, \ldots, \tau_{M+1}^*\) (existence is because \(\mu\) is the convex hull) satisfying \(\sum_{i=2}^{M+1} \tau_i^* \leq 1\). By letting \(\tau_1^* = 1 - \sum_{i=2}^{M+1} \tau_i^*\) we then have that

\[
v_n(\mu) = \sum_{i=1}^{M+1} v_n(\mu_i) \tau_i^*
\]

can only hold for a measure zero set of utility functions. If there is another set of belief vectors \((\tilde{\mu}_1, \ldots, \tilde{\mu}_{M+1})\) that span \(\mu\) as a convex combination and \((\tilde{\mu}_2 - \tilde{\mu}_1, \ldots, \tilde{\mu}_{M+1} - \tilde{\mu}_1)\) 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 \(\mu\) using \((\tilde{\mu}_1, \ldots, \tilde{\mu}_{M+1})\) such that \(\tilde{\mu}_i \in X_S\) and \((\tilde{\mu}_2 - \tilde{\mu}_1, \ldots, \tilde{\mu}_{M+1} - \tilde{\mu}_1)\) are linearly dependent and since indifference for a mixture of \((\tilde{\mu}_1, \ldots, \tilde{\mu}_{M+1})\) and \((\mu_1, \ldots, \mu_{M+1})\) is only possible if the sender is indifferent between \(\mu\) and the convex combination of \((\tilde{\mu}_1, \ldots, \tilde{\mu}_{M+1})\) and \(\mu\) and the convex combination of \((\mu_1, \ldots, \mu_{M+1})\) we conclude that

\[
v_n(\mu) > \sum_{\mu' \in X_S \setminus \{\mu\}} v_n(\mu') \tau(\mu')
\]

for almost all utility functions of the sender. Finally, since \(\mu \in X_S\), a finite set, we conclude that the same conclusion holds also over all \(\mu \in X_S\). Finally, provided that sender \(n\) has generic preferences the problem for sender \(n - 1\) is qualitatively identical and again the set of preferences for \(n - 1\) with unique (vertex) continuation equilibria has full measure. The result follows by induction. \(\square\)
References


