Journal of Mathematical Economics 46 (2010) 56–70

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Finite order implications of common priors in infinite models

Barton L. Lipman*

Boston University, Economics, 270 Bay State Road, Boston, MA 02215, United States

ARTICLE INFO

Article history:
Received 20 May 2008
Received in revised form 9 July 2009
Accepted 12 July 2009
Available online 18 July 2009

Keywords:
Common priors

ABSTRACT

Lipman [Lipman, B., 2003. Finite order implications of common priors, Econometrica, 71 (July), 1255–1267] shows that in a finite model, the common prior assumption has weak implications for finite orders of beliefs about beliefs. In particular, the only such implications are those stemming from the weaker assumption of a common support. To explore the role of the finite model assumption in generating this conclusion, this paper considers the finite order implications of common priors in the simplest possible infinite model, namely, a countable model. I show that in countable models, the common prior assumption also implies a tail consistency condition regarding beliefs. More specifically, I show that in a countable model, the finite order implications of the common prior assumption are the same as those stemming from the assumption that priors have a common support and have tail probabilities converging to zero at the same rate.

1. Introduction

While controversial, the common prior assumption (CPA) continues to be used in the vast majority of work in incomplete information game theory and information economics. It is a key ingredient in many of the standard results in the growing literature on epistemology, such as Aumann's classic 1976 "agreeing-to-disagree" theorem or the no-trade theorem (Milgrom and Stokey, 1982). I will be more precise later, but for the moment, I define the common prior assumption to be that the beliefs of all agents are generated from a single prior updated by Bayes' rule. Given the traditional view that tastes can differ arbitrarily across agents, this consistency seems at odds with the view that beliefs are simply part of a representation of preferences as in Savage (1954). For more discussion of the controversy, see Morris (1995).

The goal of this paper is to contribute toward understanding the meaning of the common prior assumption. One way to understand the CPA is to characterize the sets of beliefs agents could have and still satisfy the assumption. Several authors have given various versions of such a characterization—see, for example, Bonnano and Nehring (1999), Feinberg (2000); Halpern (2002); Heifetz (2006), and Samet (1998a, b). In Lipman (2003), I take a different approach, focusing on the implications the CPA has for *finite* orders of beliefs. That is, I characterized the implications of the CPA on beliefs about the beliefs about... the beliefs of others, where "beliefs about" is repeated only a finite number of times and analogously for knowledge. I showed that the common prior assumption has weak finite order implications and that these implications are those stemming from the much weaker assumption that priors have a common support.

One limitation of that analysis was that the result held for finite models—that is, models with a finite set of possible states of the world. Here I explore the importance of this finiteness assumption by giving an analogous characterization result for the simplest possible class of infinite models, namely countable models. What is shown here is that in countable models, the CPA has an additional finite order implication which I call *tail consistency*. The main result of this paper can be summarized

* Corresponding author.
E-mail address: blipman@bu.edu.

0304-4068/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmateco.2009.07.002

by saying that the finite order implications of the CPA in countable models are the same as those stemming from the weaker assumption that priors have a common support and that the tail probabilities converge to zero at the same rate. I also show that in certain information structures, such as in Rubinstein's (1989) email game, the tail consistency condition is not needed.

To state the result a little more precisely, fix a model where the knowledge and beliefs of agents are characterized by a state space, partitions, and conditional probability beliefs. At a state in such a model, we can say which facts about the world are true, what the agents believe and know about these facts (first-order beliefs and knowledge), what they believe and know about the beliefs and knowledge of others about facts (second-order beliefs and knowledge), etc. Suppose the model has the properties that the priors of the agents have a common support and tails which converge to zero at the same rate. Then for any finite N, there is another model which satisfies common priors and a state in that model at which all the same facts about the world are true and all the same statements about beliefs and knowledge of order less than N are true.

An implication of this result is that any theorem about what can happen at a state which requires the common prior assumption must depend on infinitely many orders of beliefs or knowledge. Similarly, if a result does not depend on infinitely many orders of beliefs, it cannot depend on common priors. For example, Weinstein and Yildiz (2007) consider when an equilibrium is robust to higher order beliefs in the sense that we can make a prediction based only on beliefs up to some finite order. My result implies that as long as the priors at least have a common support and tails converging at the same rate, the answer to this question cannot depend on whether priors are common or not since finite orders of beliefs cannot distinguish whether priors are common.[1]

Somewhat loosely, this implies that if the common prior assumption is needed, one cannot drop common knowledge assumptions. For example, the fact that the agreeing to disagree result and no-trade theorems do not hold without common priors indicates that, as is well-known, these results rely on their common knowledge assumptions. Intuitively, this is true because results at a state which make nontrivial use of common knowledge assumptions typically use common knowledge to translate the local statement (something is true at a state) to a global one (something is true at every state). For example, the no-trade and agreeing to disagree theorems convert a statement about what is true at a state to the conclusion that this must be true through an event of the state space which can be taken to be the entire space. The new model I construct to match beliefs at a state cannot match the original model globally. Hence it should not be surprising that results, local or not, whose proofs rely on global statements will not carry over to noncommon priors.

There are three reasons why I focus here on countable models. First, obviously, this is the simplest class of infinite models and hence the natural class to begin with in determining whether the finiteness assumption of Lipman (2003) was restrictive. Second, the proofs used here and in Lipman (2003) are very constructive and exploit the discrete structure of the model heavily. Thus very different techniques seem necessary for considering uncountable models.

Finally, my definition of the common prior assumption requires that all partitions be countable. Countability of partitions does not imply that the model is countable, of course. The point is simply that the difference between the case I analyze here and the most general that could be considered is smaller than it may appear. I explain this point further, including the motivation for my approach, in Remark 1 below.

In the next section, I give basic definitions. Section 3 contains the main result. Most proofs are in Appendix A.

2. Definitions

The set of players is $\mathcal{I} = \{1, \ldots, I\}$. There is a parameter, θ, which may be unknown to these players and the beliefs in question are beliefs about θ and beliefs about the beliefs of others. Θ is the set of possible parameter values. I use *partitions models* to describe the beliefs and knowledge of the players.

Formally, I define a partitions model to consist of the following objects:

1. A finite or countably infinite set of *states*, S.
2. A function $f : S \to \Theta$ where $f(s)$ is the value of the unknown parameter in state s.
3. For each player $i \in \mathcal{I}$, a partition Π_i of S, where $\pi_i(s)$ denotes the event of Π_i containing s.
4. For each player $i \in \mathcal{I}$ and each event $\pi \in \Pi_i$, a conditional belief $\mu_i(\cdot|\pi)$ over π.

Without essential loss of generality, I assume that the partitions model cannot be decomposed into smaller submodels. Formally, say that s' is reachable from s in L steps if there is a sequence $s_0, \ldots, s_L$ of states in $\mathcal{M}$ and a sequence of agents $i_1, \ldots, i_L$, with $s_0 = s$, $s_L = s'$, and

$$s_\ell \in \pi_{i_\ell}(s_{\ell-1}), \quad \ell = 1, \ldots, L.$$

I say that s' is reachable from s if it is reachable in some number of steps. Finally, given any players i and j, I say that s' is reachable from s in L steps through i and j if we can take the sequence of agents $i_1, \ldots, i_L$ to have the property that $i_\ell \in \{i, j\}$ for $\ell = 1, \ldots, L$.

[1] I thank Muhamet Yildiz for pointing this out to me. To be more precise, this statement holds when we restrict attention to finite or countable models or can use an approximation argument for uncountable models, as in Weinstein and Yildiz.

I require a partitions model to have the property that for all s and s' in the model, s' is reachable from s. While the definition is below, readers familiar already with the definition of common knowledge may find it useful to note that this assumption is equivalent to saying that the smallest common knowledge event is S. Without this assumption, two "disconnected" models could be lumped together as one, making the analysis unnecessarily awkward. I use $\mathcal{M}$ to denote a typical partitions model $(S, f, \{(\Pi_i, \mu_i)\}_{i=1,\ldots,I})$ satisfying this requirement.

I define the common prior assumption to mean that every agent's beliefs are generated by a common prior. More specifically, I will say a model $\mathcal{M}$ satisfies the common prior assumption (CPA) if there exists a probability distribution μ on S with $\mu(\pi_i) > 0$ for all $\pi_i \in \cup_i \Pi_i$ such that for all i and all $\pi_i \in \Pi_i$,

$$\mu_i(E|\pi_i) = \frac{\mu(E \cap \pi_i)}{\mu(\pi_i)}$$

for every event E. More generally, I say that a probability distribution on S, $\hat{\mu}_i$, is a *prior for i* if $\hat{\mu}_i(\pi_i) > 0$ for all $\pi_i \in \Pi_i$ and

$$\mu_i(E|\pi_i) = \frac{\hat{\mu}_i(E \cap \pi_i)}{\hat{\mu}_i(\pi_i)}.$$

The common prior assumption is then the statement that there exists single probability distribution over S which is a prior for each i.

Remark 1. Since my definition of common priors requires $\mu(\pi_i) > 0$ for all π_i, a necessary condition for a model to be consistent with common priors is that the partitions are countable. Hence, as noted in the introduction, it is a "small" (though nontrivial) step from this assumption to my focus on countable models. Of course, one alternative approach, similar to what Mertens and Zamir (1985) call consistency, would be to define common priors to require $\mu_i(E|\pi_i)\mu(\pi_i) = \mu(E \cap \pi_i)$. For partition events π_i such that $\mu(\pi_i) = 0$, this condition is always satisfied. The problem with this approach is that this definition of common priors has virtually no finite order implications for beliefs. See Bonnano and Nehring (1999) and Halpern (2002).

Intuitively, a state is an implicit representation of a set of facts, beliefs and knowledge agents have about those facts (first-order beliefs and knowledge), beliefs and knowledge agents have about the beliefs and knowledge of others about those facts (second-order beliefs and knowledge), etc. To see this more precisely, first note that $f(s)$ gives the parameter value true at state s, representing which facts about the world are true at s. First-order knowledge and beliefs will refer to knowledge and beliefs about the parameter value.

To formalize this, let[2]

$$A_1 = \{E \subseteq S | E = f^{-1}(\hat{\Theta}) \text{ for some } \hat{\Theta} \subseteq \Theta\}.$$

For any collection of subsets $\mathcal{E}$ of S, let $\sigma(\mathcal{E})$ denote the σ-algebra generated by these sets. Let $\mathcal{A}_1 = \sigma(A_1)$. I refer to these events as *1* measurable. First-order beliefs and knowledge then refer to beliefs or knowledge regarding the 1 measurable events. More formally, define:

$$K_i(E) = \{s \in S | \pi_i(s) \subseteq E\}.$$

and

$$B_i^p(E) = \{s \in S | \mu_i(E|\pi_i(s)) = p\}$$

The former is standard: $K_i(E)$ is the set of states where every state which i believes possible is contained in the event E or, more colloquially, where i knows E. The latter is the set of states where i gives probability p to the event E. While the notation I use is the same as theirs, this should not be confused with Monderer and Samet's (1989) notion of p-belief. They define p-belief to mean that i's probability of E is *at least* p, while here i's probability is *exactly* p.

Second-order beliefs and knowledge then refers to the events created by applying the K_i and B_i^p operators to the events corresponding to first-order statements, etc. Formally, for any collection of events $\mathcal{E}$, let:

$$\hat{\mathrm{A}}(\mathcal{E}) = \{E \subseteq S | E \in \mathcal{E} \text{ or } \exists F \in \mathcal{E}, i \in \mathcal{I}, p \in [0, 1] \text{ with } E = B_i^p(F) \text{ or } E = K_i(F)\}.$$

Then define $A_{n+1} = \hat{\mathrm{A}}(\mathcal{A}_n)$ and $\mathcal{A}_{n+1} = \sigma(A_{n+1})$. The events in $\mathcal{A}_n$ are *n measurable events*. An n measurable event corresponds to some statement about facts plus beliefs and knowledge up to order $n-1$. Then n th order beliefs are beliefs over n measurable events.

While these definitions identify subsets of a particular model, there is a sense in which these sets have a model-independent meaning. Intuitively, the meaning of these events can be described entirely in terms of Θ and $\mathcal{I}$, objects which are the same in all models. Similarly, a state in one model corresponds to a set of beliefs and knowledge which could be

[2] One may wish to restrict attention to some class of measurable subsets of Θ. Given that S is assumed to be finite or countable, such a requirement would play no role in the analysis.

held at a state in another model, making these states "equivalent." To identify these notions, I define two sets of equivalence relations.

Suppose we have two models $\mathcal{M}$ and $\bar{\mathcal{M}}$ (where these could be the same model). Recall that a 1 measurable event corresponds simply to a subset of Θ. Hence for 1 measurable events E and $\bar{E}$, it seems natural to call these events "equivalent" if they correspond to the same subset of Θ. That is, I define $E =_0 \bar{E}$ if $f(E) = \bar{f}(\bar{E})$. Similarly, I define $s \sim_0^* \bar{s}$ if $f(s) = \bar{f}(\bar{s})$. Note that I could equivalently start by defining $\sim_0^*$ and then define $=_0$ by $E =_0 \bar{E}$ iff for all $s \in E$, there is a $\bar{s} \in \bar{E}$ with $s \sim_0^* \bar{s}$ and analogously for all $\bar{s} \in \bar{E}$.

A comment on notation may be useful at this point. These relations are defined with respect to a particular pair of models and so should, in principal, be indexed by $\mathcal{M}$ and $\bar{\mathcal{M}}$. For readability, I omit this dependence from the notation. Also, whenever I put a $^-$ on top of a variable, it should be understood to refer to a variable in $\bar{\mathcal{M}}$, while the same object without a $^-$ refers to $\mathcal{M}$.

Returning to these relations, the definition is more complex at higher levels because there are more ways two events or two states could be equivalent. Intuitively, for two states to be equivalent at level n, the same statements about knowledge and beliefs up to level n should hold at each. More precisely, I define $s \sim_n^* \bar{s}$ if $s \sim_{n-1}^* \bar{s}$ and for all n measurable E and $\bar{E}$ with $E =_{n-1} \bar{E}$,

$$s \in K_i(E) \quad \text{iff} \quad \bar{s} \in \bar{K}_i(\bar{E})$$

and

$$s \in B_i^p(E) \quad \text{iff} \quad \bar{s} \in \bar{B}_i^p(\bar{E}).$$

To understand this, recall that the new events that come in at level n are those that involve statements of knowledge or belief regarding events at level $n-1$. Hence if E and $\bar{E}$ represent the same events at $n-1$, equivalence of s and $\bar{s}$ at level n requires knowledge and beliefs regarding these events to be the same. We can then define $=_n$ over $n+1$ measurable events E and $\bar{E}$ by $E =_n \bar{E}$ iff for all $s \in E$, there is a $\bar{s} \in \bar{E}$ with $s \sim_n^* \bar{s}$ and analogously for all $\bar{s} \in \bar{E}$.

Remark 2. The reader may find it easier to follow the notation by noting that n th order beliefs are defined over n measurable events. Hence n measurable events are events referring to $n-1$ order beliefs. In other words, the order of measurability of an event is one higher than the order of the beliefs and knowledge the event is "about." Thus when I refer to events that are the same in terms of beliefs and knowledge up to order n, it is natural to focus on events which are $n+1$ measurable.

It is worth noting for future use that one can define the n measurable sets of a model in terms of $\sim_{n-1}^*$ defined between the model and itself. More specifically, it is not hard to show that

Lemma 1. *E in $\mathcal{M}$ is n measurable iff it is generated by the set of equivalence classes of S under the equivalence relation $\sim_{n-1}^*$ defined on $\mathcal{M} \times \mathcal{M}$.*

I will say that a state s in a model $\mathcal{M}$ s atisfies the common prior assumption to level N if there is a model $\bar{\mathcal{M}}$ satisfying the CPA and a state $\bar{s}$ in that model with $s \sim_N^* \bar{s}$. I will say s is *f* initely consistent with the CPA if it is satisfies the common prior assumption to level N for every finite N. The goal of this paper is to characterize those states which are finitely consistent with the CPA.

Remark 3. To see why I use conditional beliefs rather than priors, consider the following example. There are four states, so $S = \{a, b, c, d\}$, and two players. Player 1 has no information at all, so $\Pi_1 = \{S\}$. Since player 1 has no information, a "conditional" belief for him is simply a distribution $\mu_1(\cdot|S)$ over S. Suppose his beliefs give probability 1/4 to every state. Player 2's partition is $\{\{a, b\}, \{c, d\}\}$. Intuitively, the "prior" probability 2 assigns to $\{a, b\}$ versus $\{c, d\}$ is irrelevant. If 2's conditional belief given $\{a, b\}$ assigns equal probability to a and b and similarly for his conditional belief given $\{c, d\}$, then we can treat 2 as if his prior is the same as 1's. Thus the key is 2's conditional beliefs, not any prior he might have.

3. Results

My approach is to derive necessary conditions for certain states in certain models to be finitely consistent with the common prior assumption. I will then show that these conditions are sufficient for *any* state in *any* model. In this sense, the restriction on $(\mathcal{M}, s)$ pairs for which these conditions are necessary is a restriction to the "hard cases."

Given a model $\mathcal{M}$ and a state s in that model,

Definition 1. *$(\mathcal{M}, s)$ is* difficult *if (a) every $\{s'\}$ is 1 measurable and (b) for every i and j, $i \neq j$, there is a finite L such that any state is reachable from s through i and j within L steps.*

For example, suppose $S = \Theta = \{1, 2, \ldots\}$ and that there are two players. Suppose $f(s) = s$ for all s, so every singleton set is 1 measurable. Suppose $\Pi_1 = \{S\}$, so 1 has no information. Then every state is reachable from every other in one step, so $(\mathcal{M}, s)$ is difficult for all s, regardless of what we assume about Π_2.

To see why what I have called the difficult case is indeed the case where finite consistency with the CPA is most restrictive, consider a pair $(\mathcal{M}, s)$ which is not difficult. Suppose we wish to show that this is consistent with common priors to some level N. If condition (a) does not hold, then we may be able to ignore some events which are only measurable at levels higher than N, thus making it easier to prove consistency. Alternatively, if condition (b) fails, then we may be able to ignore some states which are more than N steps away from s, at least as far as beliefs of i and j are concerned. The problem with considering necessary conditions when (a) or (b) fail is that the statements will involve tedious considerations such as at what point various events become measurable and how many steps it takes to get to some state from s for various players.

As this intuition suggests, the necessity proofs would be essentially unchanged if we replace the condition that every $\{s\}$ is 1 measurable with the weaker statement that there is a finite n such that every $\{s\}$ is n measurable. It is not hard to show that if $\mathcal{M}$ has a finite state space, then for all s, $(\mathcal{M}, s)$ is difficult under this weakened definition. Hence the conditions I develop are necessary for all finite models.

There are two necessary conditions for finite consistency with the CPA in the difficult case. To state the first requires additional notation. Fix a model $\mathcal{M}$. For any $E \subseteq S$, let $K_G(E)$ denote the event that "everyone knows" E —that is, that E is group knowledge in the sense that:

$$K_G(E) = \bigcap_{i=1}^{I} K_i(E).$$

Recursively define order n group knowledge, K_G^n, by $K_G^{n+1}(E) = K_G(K_G^n(E))$ where $K_G^1 = K_G$. Finally, $CK(E)$ is the event that E is common knowledge:

$$CK(E) = \bigcap_{n=1}^{\infty} K_G^n(E).$$

Given my assumption that every state is reachable from every other state, the only common knowledge event in a model is S itself.

Given a partitions model $\mathcal{M}$, let τ denote the event that each agent gives zero probability to the occurrence of an event at the same time another agent gives that event zero probability. That is,

$$\tau = \{s' \in S | \mu_i(E \cap B_j^0(E) | \pi_i(s')) = 0, \forall i, j, E\}.$$

Definition 2. *State s in partitions model $\mathcal{M}$ satisfies* support consistency *if $s \in CK(\tau)$.*

Given my assumption that S itself is the smallest common knowledge event in S, support consistency requires that for all states s', players i and j, and events E,

$$\mu_i(E \cap B_j^0(E) | \pi_i(s')) = 0.$$

Thus even though support consistency is a property of a state, the use of common knowledge turns it into a property of the model.

One way to understand support consistency and to see where the name comes from is to note that it is equivalent to the condition that there are priors for the agents with the same support.

Lemma 2. *State s in partitions model $\mathcal{M}$ satisfies support consistency if and only if for all probability distributions $\hat{\mu}_i$, $i \in \mathcal{I}$, such that $\hat{\mu}_i$ is a prior for i, we have $\hat{\mu}_i(s') > 0$ for any i iff $\hat{\mu}_i(s') > 0$ for all i.*

Proof. s satisfies support consistency if and only if for every s', for every i and j, and every event E,

$$\mu_i(E \cap B_j^0(E) | \pi_i(s')) = 0.$$

I claim this holds if and only if for any s' and any i and j,

$$\mu_i(s' | \pi_i(s')) = 0 \quad \text{iff} \quad \mu_j(s' | \pi_j(s')) = 0.$$

To see this, first suppose s satisfies support consistency and that s' is a state such that there is an i and j with:

$$\mu_i(s' | \pi_i(s')) > 0$$

but

$$\mu_j(s' | \pi_j(s')) = 0.$$

Let $E = \{s'\}$. Then $E \cap B_j^0(E) = \{s'\}$. But $\mu_i(s' | \pi_i(s')) > 0$, contradicting support consistency.

For the converse, suppose that for all s' and all i and j, $\mu_i(s' | \pi_i(s')) = 0$ iff $\mu_j(s' | \pi_j(s')) = 0$. Suppose, though, that s is not support consistent. So there is some s', some E, and some i and j with:

$$\mu_i(E \cap B_j^0(E) | \pi_i(s')) > 0.$$

Fix any $s'' \in E \cap B_j^0(E) \cap \pi_i(s')$ such that $\mu_i(s''|\pi_i(s')) > 0$. (Such an s'' must exist.) Since $s'' \in \pi_i(s')$, we have $\pi_i(s'') = \pi_i(s')$ so $\mu_i(s''|\pi_i(s'')) > 0$. Also, by definition, $s'' \in B_j^0(E)$ implies $\mu_j(E|\pi_j(s'')) = 0$. But since $s'' \in E$, we have $\mu_j(s''|\pi_j(s'')) = 0$, a contradiction.

To complete the proof, then, note that if we have $\hat{\mu}_i, i \in \mathcal{I}$, which are priors for the players, then there must exist strictly positive probability distributions β_i on Π_i for each i such that $\hat{\mu}_i(s) = \mu_i(s|\pi_i(s))\beta_i(\pi_i(s))$. Obviously, $\hat{\mu}_i(s') > 0$ iff $\mu_i(s'|\pi_i(s')) > 0$. Hence s satisfies support consistency if and only if for all s' and any i and j, we have $\hat{\mu}_i(s') > 0$ iff $\hat{\mu}_j(s') > 0$. □

Definition 3. *$\mathcal{M}$ satisfies* tail consistency *if there exists $\mathcal{L} < \infty$ and probability distributions $\hat{\mu}_i$, $i \in \mathcal{I}$, such that $\hat{\mu}_i$ is a prior for i and*

$$\frac{1}{\mathcal{L}} < \frac{\hat{\mu}_i(s)}{\hat{\mu}_j(s)} < \mathcal{L}$$

for all s with $\hat{\mu}_j(s) > 0$.

Note that tail consistency implies support consistency. To be more precise, if a model $\mathcal{M}$ satisfies tail consistency, then every s in that model satisfies support consistency. To see this, note that the constant $\mathcal{L}$ in the definition of tail consistency must be strictly positive. Hence the definition requires that if $\hat{\mu}_j(s) > 0$, we have $\hat{\mu}_i(s) > 0$ for all i. From Lemma 2, we know that this implies that every s in the model satisfies support consistency. What tail consistency adds to support consistency is that the priors have tail probabilities converging to zero at the same rate.

The main result is

Theorem 1. *If $(\mathcal{M}, s)$ is difficult and s is finitely consistent with the common prior assumption, then s satisfies support consistency and $\mathcal{M}$ satisfies tail consistency. For any model $\mathcal{M}$ satisfying tail consistency and any state s in that model which satisfies support consistency, s is finitely consistent with the common prior assumption.*

That is, support consistency and tail consistency is necessary for finite consistency with the CPA in difficult models and are sufficient in *any* model.

Remark 4. It is well-known that for any type space, there is a partitions model generating the same beliefs at all levels and conversely. Using the translation between these models as given in, *e.g.*, Brandenburger and Dekel (1993), it is tedious but not difficult to translate Theorem 1 into the language of type spaces. I briefly sketch the idea here. Define a type space to consist of finite or countably infinite sets $T_1, \ldots, T_I$, and belief functions $b_i : T_i \to \Delta(\Theta \times \prod_{j \neq i} T_j)$, $i = 1, \ldots, I$. Let $T = \prod_i T_i$, define t_{-i} and T_{-i} as usual, and say that $\hat{b}_i \in \Delta(\Theta \times T)$ is a prior for i if

$$\sum_{(\theta, t_{-i}) \in \Theta \times T_{-i}} \hat{b}_i(\theta, t_i, t_{-i}) > 0, \quad \forall t_i \in T_i$$

and

$$b_i(\theta, t_{-i})(t_i) = \frac{\hat{b}_i(\theta, t_i, t_{-i})}{\sum_{(\theta', t'_{-i}) \in \Theta \times T_{-i}} \hat{b}_i(\theta', t_i, t'_{-i})}, \quad \forall (\theta, t_i, t_{-i}) \in S \times T.$$

From here, it is straightforward to define the common prior assumption, support consistency, and tail consistency for type spaces. Essentially, (θ, t) replaces s and t_i replaces π_i everywhere. Orders of beliefs can be defined in an analogous way to define the notion of a state (θ, t) being finitely consistent with the CPA as can the notion of a difficult type space. With these adaptations, Theorem 1 carries over to type spaces.

The proof of this result is easily adapted to show that tail consistency is not needed for a broad class of models.

Definition 4. *Given a model $\mathcal{M}$ and a state s in that model, s is* essentially finite *if for every finite L, there are only finitely many s' such that s' is reachable from s in L steps or fewer.*

Corollary 1. *If s satisfies support consistency and is essentially finite, then it is finitely consistent with the common prior assumption.*

To see why this follows, suppose s satisfies support consistency and is essentially finite. To show that s satisfies the common prior assumption to level N requires us to consider only those states s' which are reachable from s in N steps or fewer. Since the model is essentially finite, there are finitely many such s'. Hence we can treat the model as if it were finite and tail consistency is irrelevant.

This corollary, together with the necessity part of Theorem 1, generalizes the main result in Lipman (2003). There, I showed that support consistency (which I called weak consistency in that paper) is necessary and sufficient for a state to be finitely consistent with the common prior assumption *i* f the model is finite. The results here are more general in two respects. First, Theorem 1 shows that the necessity of support consistency also carries over to difficult countable models.

Second, Corollary 1 weakens the condition under which I showed support consistency to be sufficient from finitely many states to finitely many states reachable from s in any fixed, finite number of steps.

Note that in the information structure used by Rubinstein (1989) in his email game, every state is essentially finite. Consequently, Rubinstein's results do not require common priors, only priors with a common support. For example, it is easy to see that the players in Rubinstein's game do not have to have the same prior to get his result.[3]

Intuitively, Rubinstein's point is to show that arbitrarily high degrees of mutual knowledge need not yield the same implications as common knowledge. As explained in the introduction, my results indicate that a common prior assumption is only essential if common knowledge assumptions are. Hence the implications of a *f* ailure of common knowledge are unlikely to hinge on whether there is a common prior.

Remark 5. It is natural to wonder if tail consistency is related to the compactness condition in Feinberg (2000). There does not appear to be any link between the two concepts. Perhaps the simplest way to see this is with reference to the example Feinberg gives on page 152 of a model violating his compactness condition. The state space in the example is $\{1, 2, \ldots\}$. The partition of player 1 is $\{1\}, \{2, 3\}, \{4, 5\}, \ldots$, while the partition of player 2 is $\{1, 2\}, \{3, 4\}, \{5, 6\}, \ldots$. Feinberg constructs conditional beliefs such that there is no common prior even though there is no "disagreement in expectation." However, the conditional beliefs he gives do satisfy tail consistency. To see this, simply choose a probability distribution over information sets for each player given by a geometric distribution. That is, number the information sets so that 1's first information set is $\{1\}$ and his k th information set is $\{2(k-1), 2k-1\}$ for $k \geq 2$. Let 2's k th information set be $\{2k-1, 2k\}$. Let the prior over information sets for each player be that the k th information set has prior probability $(1-\delta)\delta^{k-1}$. For any s, let $k_i(s)$ be the number of the information set for i containing s. Define priors for the two players using their conditional beliefs and these priors over events. For any s, then, the ratio of i's prior probability on s to j's is

$$\frac{\mu_i(s|\pi_i(s))}{\mu_j(s|\pi_j(s))}\delta^{k_i(s)-k_j(s)}.$$

In Feinberg's example, $\mu_i(s|\pi_i(s))/\mu_j(s|\pi_j(s))$ is never more than 3/2. Also, it is easy to see that $k_i(s)-k_j(s)$ is never smaller than -1. Hence the ratio of the prior probabilities is bounded above by $3/(2\delta)$, so tail consistency holds. In fact, in this example, tail consistency is irrelevant to determining which states are finitely consistent with the CPA. Note that every state in this model is essentially finite. Hence the fact that support consistency holds implies that every state in this model is finitely consistent with the common prior assumption by Corollary 1.

The proof of Theorem 1 is in Appendix A because some of the details are rather tedious. The basic ideas, however, are simple. To see the intuition for the necessity results, consider the following example. Suppose there are two players, countably many states, and neither player has any information. That is, $\Pi_1 = \Pi_2 = \{S\}$. Also, assume f is one-to-one: that is, the parameter values at different states differ. Finally, let μ_1 and μ_2 denote the beliefs of the two players where $\mu_1 \neq \mu_2$. Because neither player ever has any information, the beliefs of the players do not vary with the state. Hence if any state is finitely consistent with the common prior assumption, all states are.[4]

So suppose all states in this model are finitely consistent with the common prior assumption. By definition, for any state s in this model and any finite N, we can find a model satisfying common priors and a state in that model, say $\bar{s}$, where beliefs at $\bar{s}$ match beliefs at s up to order N. It is not hard to show that a stronger statement is true for this model. In particular, for any N, we can find one model satisfying common priors where *every* state in the original model is matched up to level N by some state in the common priors model.

With this in mind, fix any $N \geq 2$ and let $\bar{\mathcal{M}}$ denote the model with common priors for which every state in $\mathcal{M}$ is matched up to level N by some state in $\bar{\mathcal{M}}$. Let $\bar{S}_1$ denote the set of states in $\bar{S}$ such that first and second-order beliefs of player 1 in $\bar{\mathcal{M}}$ match the first and second-order beliefs he has at every state in $\mathcal{M}$. To state this more precisely, for any state s in the original model, let $\mathcal{E}_s$ denote the set of $\bar{s}$ in the new model with the same parameter value as s. That is, it is the set of $\bar{s}$ such that $\bar{f}(\bar{s}) = f(s)$. Then the statement that first-order beliefs of player 1 at any $\bar{s} \in S_1$ are the same as his first-order beliefs at any state in S says:

$$\frac{\bar{\mu}(\mathcal{E}_s \cap \bar{\pi}_1(\bar{s}))}{\bar{\mu}(\bar{\pi}_1(\bar{s}))} = \mu_1(s), \quad \forall s \in S, \quad \forall \bar{s} \in \bar{S}_1$$

implying

$$\sum_{\bar{s} \in \bar{S}_1} \bar{\mu}(\mathcal{E}_s \cap \bar{\pi}_1(\bar{s})) = \mu_1(s) \sum_{\bar{s} \in \bar{S}_1} \bar{\mu}(\bar{\pi}_1(\bar{s})).$$

[3] Information structures which are similar to Rubinstein's are used in the large literature on global games—see, for example, Frankel et al. (2003). Those information structures typically use a continuum of states and are not essentially finite. It remains to be seen whether those results require common priors.

[4] Once Theorem 1 is established, we see that this is true for all models. That is, while it is immediately true in the model of this example that if any state is finitely consistent with the CPA, then all are, an implication of Theorem 1 is that this is true for all models.

It is not hard to see that $\bar{S}_1$ must be a union of $\bar{\pi}_1 \in \bar{\Pi}_1$, so this implies:

$$\bar{\mu}(\mathcal{E}_s \cap \bar{S}_1) = \mu_1(s)\bar{\mu}(\bar{S}_1). \tag{1}$$

The implications of matching 1's second-order beliefs are slightly more complex. Recall that in the original model, neither player has any information. Hence 2's first-order beliefs do not vary with the state, so 1 knows 2's first-order beliefs. Hence for every $\bar{s} \in \bar{S}_1$, for every state 1 considers possible at $\bar{s}$, 2's first-order beliefs are the same as his first-order beliefs in S. Since $\bar{S}_1$ is a union of information events for 1, this implies:

$$\frac{\bar{\mu}(\mathcal{E}_s \cap \bar{\pi}_2(\bar{s}))}{\bar{\mu}(\bar{\pi}_2(\bar{s}))} = \mu_2(s), \quad \forall s \in S, \quad \forall \bar{s} \in \bar{S}_1$$

implying

$$\sum_{\bar{s} \in \bar{S}_1} \bar{\mu}(\mathcal{E}_s \cap \bar{\pi}_2(\bar{s})) = \mu_2(s) \sum_{\bar{s} \in \bar{S}_1} \bar{\mu}(\bar{\pi}_2(\bar{s})). \tag{2}$$

It could be that for some $\bar{s} \in \bar{S}_1$, $\bar{\pi}_2(\bar{s})$ includes some states not contained in $\bar{S}_1$. However, clearly, $\bar{\pi}_2(\bar{s})$ includes $\bar{s}$. That is,

$$\bigcup_{\bar{s} \in \bar{S}_1} [\mathcal{E}_s \cap \bar{\pi}_2(\bar{s})] \supseteq \mathcal{E}_s \cap \bar{S}_1.$$

Hence the left-hand side of Eq. (2) is at least $\bar{\mu}(\mathcal{E}_s \cap \bar{S}_1)$. From Eq. (1), then, we have:

$$\mu_1(s)\bar{\mu}(\bar{S}_1) \leq \mu_2(s) \sum_{\bar{s} \in \bar{S}_1} \bar{\mu}(\bar{\pi}_2(\bar{s})).$$

To see why support consistency is necessary, note that if $\mu_1(s) > 0$, this equation implies $\mu_2(s) > 0$. (Recall that my formulation of the common prior assumption requires that the prior probability of each of the relevant information events is strictly positive.) Since we could obviously reverse the roles of 1 and 2, we see that the beliefs of the two players must have the same support, just as support consistency requires.

To see why tail consistency is necessary, assume for simplicity that $\mu_i(s) > 0$ for all s, $i = 1, 2$. Then we can rewrite the last equation as

$$\frac{\mu_1(s)}{\mu_2(s)} \leq \frac{\sum_{\bar{s} \in \bar{S}_1} \bar{\mu}(\bar{\pi}_2(\bar{s}))}{\bar{\mu}(\bar{S}_1)}.$$

Note that the right-hand side does not depend on s. A similar argument gives an upper bound on $\mu_2(s)/\mu_1(s)$ which is independent of s. Hence we see that $\mu_j(s)/\mu_i(s)$ must be bounded from above, exactly what tail consistency means when $\Pi_1 = \Pi_2 = \{S\}$.

The intuition for the sufficiency proof can also be illustrated in this example. Let us construct $\bar{\mathcal{M}}$ by letting $\bar{S} = S \times \{1, \ldots, K\}$ for some K. Intuitively, $(s, k) \in \bar{S}$ is the k th copy of $s \in S$. In line with this intuition, let $\bar{f}(s, k) = f(s)$. I will construct $\bar{\mathcal{M}}$ so that $s \sim_N^* (s, 1)$ for every s. To construct player 1's partition containing some $(s, 1)$, let us create an event which consists of all the states $\bar{s}$ of the form $(s, 1)$. To match player 1's first-order beliefs, I set:

$$\frac{\bar{\mu}(s, 1)}{\bar{\mu}(\bar{\pi}_1(s, 1))} = \mu_1(s), \quad \forall s \in S.$$

Once we choose $\bar{\mu}(\bar{\pi}_1(s, 1))$, then, we have determined $\bar{\mu}(s, 1)$ for all s. Given this, how we will make player 2's first-order beliefs work out properly at $(s, 1)$? The simplest way is to construct this event of player 2's partition by setting it equal to the set of $\bar{s}$ of the form $(s, 1)$ *or* $(s, 2)$ for $s \in S$. The presence of the states of the form $(s, 2)$ gives us a degree of freedom to get 2's beliefs right. More specifically, we need:

$$\frac{\bar{\mu}(s, 1) + \bar{\mu}(s, 2)}{\bar{\mu}(\bar{\pi}_2(s, 1))} = \mu_2(s), \quad \forall s \in S.$$

Given $\bar{\mu}(\bar{\pi}_1(s, 1))$, we have already determined $\bar{\mu}(s, 1)$ for each s. Hence given $\bar{\mu}(\bar{\pi}_2(s, 1))$, this equation will determine $\bar{\mu}(s, 2)$ for every s. Because, of course, we require $\bar{\mu}(s, 2) \geq 0$, this will force us to make the total prior probability on the $(s, 2)$ states large enough relative to the total prior probability on the $(s, 1)$ states. As long as μ_1 and μ_2 are not diverging too much, this will be possible. Tail consistency ensures that this step will work.

From here, the pattern is clear: To construct the event of player 1's partition containing states of the form $(s, 2)$, we include with these all the $(s, 3)$ states to give us the critical degree of freedom. Then player 2's event containing the $(s, 3)$ states will also contain the $(s, 4)$ states, etc. By this process, we will have completely determined $\bar{\mu}$ as a function of the prior probabilities assigned to the information sets. The only thing remaining is to ensure that we can choose these probabilities so that $\bar{\mu}$ is a legitimate probability distribution. Support consistency and tail consistency ensure that this can be done.

Acknowledgements

I thank the co-editor and two referees for helpful comments. I also wish to thank SSHRCC and NSF for financial support. Some of the material in this paper was originally part of the working paper version of "Finite order implications of common priors."

Appendix A.

Proof of Theorem 1.

Necessity. First, I prove necessity of support consistency and tail consistency. Some definitions will be useful. For these definitions, fix models $\mathcal{M}$ and $\bar{\mathcal{M}}$.

Definition 5. *$\bar{\pi}_i \in \bar{\Pi}_i$ is a copy of order k of $\pi_i \in \Pi_i$ if*

1. *For all $s \in \pi_i$, there exists $\bar{s} \in \bar{\pi}_i$ such that $s \sim_k^* \bar{s}$.*
2. *For all $\bar{s} \in \bar{\pi}_i$, there exists $s \in \pi_i$ such that $s \sim_k^* \bar{s}$.*

This is simply a generalization of $=_n$ to sets which may not be $n+1$ measurable.

Lemma A.1. *Assume every singleton $\{s\} \subset S$ is 1 m* easurable. Then $\bar{s}$ *in model $\bar{\mathcal{M}}$ s* atisfies $s \sim_n^* \bar{s}$ *if and only if $f(s) = \bar{f}(\bar{s})$ and*

1. *For all i, $\bar{\pi}_i(\bar{s})$ is a copy of order $n-1$ of $\pi_i(s)$.*
2. *For all i, for all $s' \in \pi_i(s)$,*

$$\bar{\mu}(\{\bar{s}' \in \bar{S} | \bar{s}' \sim_{n-1}^* s'\} | \bar{\pi}_i(\bar{s})) = \mu_i(s' | \pi_i(s)).$$

Proof.

(Only if.) Suppose $\bar{s} \sim_n^* s$. Obviously, $\bar{s} \sim_0^* s$, so $\bar{f}(\bar{s}) = f(s)$.

Fix any i and any $s' \in \pi_i(s)$. Suppose there is no $\bar{s}' \in \bar{\pi}_i(\bar{s})$ such that $s' \sim_{n-1}^* \bar{s}'$. Then there is an n measurable event, $\{s'\}$, which is believed possible by i at s which is not believed possible by i at $\bar{s}$, a contradiction. More precisely, $s \in \neg K_i \neg \{s'\}$, so $\bar{s} \in \neg \bar{K}_i \neg \{\bar{s}' | \bar{s}' \sim_{N-1}^* s'\}$. Similarly, suppose $s' \notin \pi_i(s)$. Then $s \in K_i \neg \{s'\}$, so $\bar{s} \in \bar{K}_i \neg \{\bar{s}' | \bar{s}' \sim_{n-1}^* s'\}$. Hence $\bar{\pi}_i(\bar{s})$ is a copy of order $n-1$ of $\pi_i(s)$. The necessity of the second statement is similarly shown.

(If.) Suppose that $\bar{\pi}_i(\bar{s})$ is a copy of order $n-1$ of $\pi_i(s)$ for all i and that:

$$\bar{\mu}(\{\bar{s}' \in \bar{S} | \bar{s}' \sim_{n-1}^* s'\} | \bar{\pi}_i(\bar{s})) = \mu_i(s' | \pi_i(s)).$$

for all i and all $s' \in \pi_i(s)$. By construction, every $n-1$ measurable event known (given probability p) by i at s is known (given probability p) by i at $\bar{s}$. Hence $s \sim_n^* \bar{s}$. □

To begin the necessity proof, fix a difficult $(s^*, \mathcal{M})$ and suppose s^* is consistent with the CPA to level N for all N. It is not hard to see that this implies that for N sufficiently large, there is a model $\bar{\mathcal{M}}$ satisfying common priors with the property that for every s in model $\mathcal{M}$, there is a state $\bar{s}$ in model $\bar{\mathcal{M}}$ such that $\bar{s} \sim_N^* s$. To see this, fix a large N, strictly larger than L, and a state and model $(\bar{\mathcal{M}}, \bar{s}^*)$ such that $\bar{s}^* \sim_N^* s^*$. By Lemma A.1, $\bar{\pi}_i(\bar{s}^*)$ must be a copy of order $N-1$ of $\pi_i(s^*)$ for every i. Hence for every $s \in \pi_i(s^*)$, there is an $\bar{s} \in \bar{\pi}_i(\bar{s}^*)$ with $\bar{s} \sim_{N-1}^* s$. Similarly, for each of these s, for every $s' \in \pi_i(s)$, there must be a state $\bar{s}'$ in $\bar{\mathcal{M}}$ with $\bar{s}' \sim_{N-2} s'$, etc. Since $(\mathcal{M}, s^*)$ is difficult, we can reach every other $s \in S$ within L steps. Hence for every $s \in S$, there is a $\bar{s}$ in model $\bar{\mathcal{M}}$ with $\bar{s} \sim_{N-L}^* s$. Since N is arbitrary, the conclusion follows.

So fix any large enough N and the model $\bar{\mathcal{M}}$ with the property that for every s in $\mathcal{M}$, there is a $\bar{s}$ in $\bar{\mathcal{M}}$ with $\bar{s} \sim_N^* s$. Define a sequence of functions as follows. For each i, $\pi_i \in \Pi_i$, and $k = 0, \ldots, N$, let

$$\bar{\Pi}_i^k(\pi_i) = \{\bar{\pi}_i \in \bar{\Pi}_i | \bar{\pi}_i \text{ is a copy of order } k \text{ of } \pi_i\}.$$

For each $s \in S$, let

$$\mathcal{E}_s = \{\bar{s} \in \bar{S} | \bar{f}(\bar{s}) = f(s)\}.$$

Lemma A.2. *For all i and all $k \geq 0$, if $\bar{s} \in \mathcal{E}_s \cap \bar{\pi}_i$ for some $\bar{\pi}_i \in \bar{\Pi}_i^k(\pi_i(s))$, then $\bar{s} \sim_k^* s$.*

Proof. By definition, if $\bar{s} \in \bar{\pi}_i \in \bar{\Pi}_i^k(\pi_i(s))$, then $\bar{s} \sim_k^* s'$ for some $s' \in \pi_i(s)$. But $\bar{s} \in \mathcal{E}_s$ implies $\bar{f}(\bar{s}) = f(s)$. Since every $\{s'\}$ is 1 measurable, $f(s) \neq f(s')$ for any $s' \neq s$. Hence $\bar{s} \not\sim_k^* s'$ for any $s' \neq s$. So $\bar{s} \sim_k^* s$. □

Fix any s and consider $\bar{\pi}_i \in \bar{\Pi}_i^k(\pi_i(s))$ for $k \geq 1$. Since $k \geq 1$, first-order beliefs at $\bar{\pi}_i$ and $\pi_i(s)$ must be the same. Hence the fact that every $\{s'\}$ is 1 measurable implies:

$$\bar{\mu}(\mathcal{E}_s \cap \bar{\pi}_i) = \bar{\mu}(\bar{\pi}_i)\mu_i(s|\pi_i(s)).$$

Fix any $\bar{s}$ contained in $\bar{\pi}_j \cap \mathcal{E}_s$ for any $\bar{\pi}_j \in \bar{\Pi}_j^{k+1}(\pi_j(s))$. By Lemma A.2, $\bar{s} \sim_{k+1}^* s$. Hence by Lemma A.1, $\bar{\pi}_i(\bar{s})$ must be a copy of order k of $\pi_i(s)$ for any i. Hence:

$$\mathcal{E}_s \cap \left[\bigcup \bar{\Pi}_j^{k+1}(\pi_j(s))\right] \subseteq \mathcal{E}_s \cap \left[\bigcup \bar{\Pi}_i^{k}(\pi_i(s))\right].$$

So

$$\sum_{\bar{\pi}_j \in \bar{\Pi}_j^{k+1}(\pi_j(s))} \bar{\mu}(\mathcal{E}_s \cap \bar{\pi}_j) \leq \sum_{\bar{\pi}_i \in \bar{\Pi}_i^{k}(\pi_i(s))} \bar{\mu}(\mathcal{E}_s \cap \bar{\pi}_i).$$

From the above, this implies:

$$\mu_j(s|\pi_j(s))\bar{\mu}(\cup\bar{\Pi}_j^{k+1}(\pi_j(s))) \leq \mu_i(s|\pi_i(s))\bar{\mu}(\cup\bar{\Pi}_i^{k}(s)). \tag{3}$$

This equation implies that s must satisfy support consistency: if $\mu_j(s|\pi_j(s)) > 0$ for any j, then $\mu_i(s|\pi_i(s)) > 0$ for all i. Henceforth, I assume support consistency holds.

To show the necessity of tail consistency, note that Eq. (3) implies that for all π_i and π_j such that $\pi_i \cap \pi_j \neq \varnothing$ and all $s \in \pi_i \cap \pi_j$,

$$\mu_j(s|\pi_j)\bar{\mu}(\cup\bar{\Pi}_j^{k+1}(\pi_j)) \leq \mu_i(s|\pi_i)\bar{\mu}(\cup\bar{\Pi}_i^{k}(\pi_i))$$

for $k = 1, \ldots, N-1$. For each i, $\pi_i \in \Pi_i$, and $k = 1, \ldots, N$, let:

$$\alpha_{ik} = \sum_{\pi_i \in \Pi_i} \bar{\mu}(\cup\bar{\Pi}_i^{k}(\pi_i))$$

and

$$\beta_{ik}(\pi_i) = \frac{\bar{\mu}(\cup\bar{\Pi}_i^{k}(\pi_i))}{\alpha_{ik}}.$$

For every i and k, β_{ik} is a strictly positive probability distribution over Π_i. Hence we obtain a family of priors for i, $\hat{\mu}_{ik}$, $k = 1, \ldots, N$, defined by

$$\hat{\mu}_{ik}(s) = \beta_{ik}(\pi_i(s))\mu_i(s|\pi_i(s)).$$

From the above, we see that:

$$\alpha_{j,k+1}\hat{\mu}_{j,k+1}(s) \leq \alpha_{ik}\hat{\mu}_{ik}(s)$$

for all s, for all i and j, and for $k = 1, \ldots, N$. Letting $\hat{\mathcal{L}} = 1 + \max_{i,j,k}\alpha_{ik}/\alpha_{j,k+1}$, we obtain:

$$\hat{\mu}_{j,k+1}(s) < \hat{\mathcal{L}}\hat{\mu}_{ik}(s)$$

for all s, i, and j, for $k = 1, \ldots, N-1$.

I now show by induction that there exists n, m, and a finite $\mathcal{L}$ such that:

$$\frac{1}{\mathcal{L}} < \frac{\hat{\mu}_{in}(s)}{\hat{\mu}_{jm}(s)} < \mathcal{L}$$

for all s such that $\mu_j(s|\pi_j) > 0$. This conclusion will establish that tail consistency holds. To do so, I first simplify the notation. Because support consistency holds, we can assume $\mu_i(s|\pi_i(s)) > 0$ for all i and all $s \in S$. Otherwise, simply restrict all statements regarding s below to apply to states with $\mu_i(s|\pi_i(s)) > 0$ for all i. Second, I replace i and j with 1 and 2.

Fix any odd N sufficiently large (in a sense made precise shortly). From the above, we know that there is a finite $\hat{\mathcal{L}} > 1$ such that:

$$\hat{\mu}_{2N}(s) < \hat{\mathcal{L}}\hat{\mu}_{1,N-1}(s) < \hat{\mathcal{L}}^2\hat{\mu}_{2,N-2}(s) < \ldots < \hat{\mathcal{L}}^{N-2}\hat{\mu}_{12}(s) < \hat{\mathcal{L}}^{N-1}\hat{\mu}_{21}(s) \tag{4}$$

for all $s \in S$. Fix any $\pi_2^* \in \Pi_2$. Recursively define sets $S_0 = \pi_2^*$. For $j = 0, 1, \ldots$,

$$S_{2j+1} = \bigcup_{s \in S_{2j}} \pi_1(s).$$

For $j = 1, 2, \ldots,$

$$S_{2j} = \bigcup_{s \in S_{2j-1}} \pi_2(s).$$

The assumption that $\mathcal{M}$ is difficult implies that there is a finite L such that $S_\ell = S$ for all $\ell \geq L$. Without loss of generality, assume $N \geq 4L + 1$.

I now show by induction that for every $j = 1, \ldots, L$, there exists a $\mathcal{L}_j < \infty$ such that each of the following holds.

I1 For every $s \in S_j$, all even n with $2j \leq n \leq N - 2j + 1$, and all odd m with $2j + 1 \leq m \leq N - 2j$,

$$\frac{1}{\mathcal{L}_j} < \frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2m}(s)} < \mathcal{L}_j.$$

I2 For every $s \in S_{j-2}$ and every even k and n with $2j \leq k, n \leq N - 2j + 1$,

$$\frac{1}{\mathcal{L}_j} < \frac{\beta_{1k}(\pi_1(s))}{\beta_{1n}(\pi_1(s))} < \mathcal{L}_j.$$

I3 For every $s \in S_{j-1}$ and every odd ℓ, m with $2j + 1 \leq \ell, m \leq N - 2j$,

$$\frac{1}{\mathcal{L}_j} < \frac{\beta_{2\ell}(\pi_2(s))}{\beta_{2m}(\pi_2(s))} < \mathcal{L}_j.$$

Statement **I1** for $j = L$ yields the claim. Note that $N \geq 4L + 1$ ensures that the set of n and m referred to in **I1** for $j = L$ is nonempty.

The proof is by induction. So consider $j = 1$. By Eq. (4), for any $s \in \pi_2^*$ and any even k with $2 \leq k \leq N - 1$,

$$\frac{1}{\hat{\mathcal{L}}^{N-k}} \beta_{2N}(\pi_2^*) < \frac{\hat{\mu}_{1k}(s)}{\mu_2(s|\pi_2^*)} < \hat{\mathcal{L}}^{k-1} \beta_{21}(\pi_2^*).$$

Since $\hat{\mathcal{L}} > 1$, this implies:

$$\frac{1}{\hat{\mathcal{L}}^{N}} \beta_{2N}(\pi_2^*) < \frac{\hat{\mu}_{1k}(s)}{\mu_2(s|\pi_2^*)} < \hat{\mathcal{L}}^{N} \beta_{21}(\pi_2^*).$$

Let:

$$\bar{\mathcal{L}} = 1 + \frac{\hat{\mathcal{L}}^N}{\beta_{2N}(\pi_2^*)} > \frac{\hat{\mathcal{L}}^N}{\beta_{2N}(\pi_2^*)} \geq \hat{\mathcal{L}}^N \geq \hat{\mathcal{L}}^N \beta_{21}(\pi_2^*).$$

Hence we have a finite $\bar{\mathcal{L}}$ satisfying:

$$\frac{1}{\bar{\mathcal{L}}} < \frac{\mu_1(s|\pi_1(s))\alpha_{1k}(\pi_1(s))}{\mu_2(s|\pi_2^*)} < \bar{\mathcal{L}}, \tag{5}$$

for all $s \in \pi_2^*$ for all even k with $2 \leq k \leq N - 1$.

In light of this, fix any even k and n with $2 \leq k, n \leq N - 1$. Obviously, for any s,

$$\frac{\beta_{1k}(\pi_1(s))}{\beta_{1n}(\pi_1(s))} = \frac{(\hat{\mu}_{1k}(s))/(\mu_2(s|\pi_2^*))}{(\hat{\mu}_{1n}(s))/(\mu_2(s|\pi_2^*))}.$$

From Eq. (5), for all $s \in \pi_2^*$, the numerator and denominator are both bounded between $1/\bar{\mathcal{L}}$ and $\bar{\mathcal{L}}$. Hence for all $s \in \pi_2^* = S_0$,

$$\frac{1}{\bar{\mathcal{L}}^2} < \frac{\beta_{1k}(\pi_1(s))}{\beta_{1n}(\pi_1(s))} < \bar{\mathcal{L}}^2. \tag{6}$$

Fix any $s \in S_1$. By definition, there exists $s' \in S_0$ such that $\pi_1(s) = \pi_1(s')$. Hence Eq. (6) holds for all $s \in S_1$. In light of this, fix even k and n with $2 \le k, n \le N-1$ and odd m with $1 \le m \le N$ and $n > m > k$. Then

$$\frac{\hat{\mu}_{1k}(s)}{\hat{\mu}_{2m}(s)} = \frac{\beta_{1k}(\pi_1(s))}{\beta_{1n}(\pi_1(s))} \frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2m}(s)}. \tag{7}$$

From the above, we see that the first term on the right-hand side is bounded between $1/\bar{\mathcal{L}}^2$ and $\bar{\mathcal{L}}^2$ for all even k and n in the range considered. From Eq. (4), $n > m$ implies the second term on the right-hand side is bounded from above by $\hat{\mathcal{L}}^{n-m}$ and so by $\hat{\mathcal{L}}^N$. Hence:

$$\frac{\hat{\mu}_{1k}(s)}{\hat{\mu}_{2,m}(s)} < \bar{\mathcal{L}}^2 \hat{\mathcal{L}}^N.$$

Also, (4) and $m > k$ implies:

$$\frac{1}{\hat{\mathcal{L}}^N} < \frac{\hat{\mu}_{1k}(s)}{\hat{\mu}_{2,m}(s)},$$

implying from Eq. (7) that:

$$\frac{1}{\bar{\mathcal{L}}^2 \hat{\mathcal{L}}^N} < \frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2m}(s)} < \hat{\mathcal{L}}^N$$

where the second inequality comes from (4). Since $\bar{\mathcal{L}}^2 > 1$, these inequalities imply that:

$$\frac{1}{\bar{\mathcal{L}}^2 \hat{\mathcal{L}}^N} < \frac{\hat{\mu}_{1k}(s')}{\hat{\mu}_{2m}(s')}, \qquad \frac{\hat{\mu}_{1n}(s')}{\hat{\mu}_{2m}(s')} < \bar{\mathcal{L}}^2 \hat{\mathcal{L}}^N \tag{8}$$

for all $s' \in S_1$ for all even k and n with $2 \le k, n \le N-1$ and all odd m with $1 \le m \le N$ such that $n > m > k$.

I claim that this implies:

$$\frac{1}{\bar{\mathcal{L}}^2 \hat{\mathcal{L}}^N} < \frac{\hat{\mu}_{1n}(s')}{\hat{\mu}_{2m}(s')} < \bar{\mathcal{L}}^2 \hat{\mathcal{L}}^N$$

for all $s' \in S_1$ for all even n with $2 \le n \le N-1$ and all odd m with $3 \le m \le N-2$. To see this, fix any such n and m. First, suppose $n > m$. Let $k = 2$. Since $m \ge 3$, we have $m > k$. Since $n > m > k$, Eq. (8) implies the result. So suppose $m > n$. In this case, change variables by replacing n with k. So we have k is even with $2 \le k \le N-1$, m odd with $3 \le m \le N-2$, and $m > k$. Let $n = N-1$. Since $m \le N-2$, we have $n > m > k$ and, again, Eq. (8) yields the result.

Given this, we can choose any $\mathcal{L}_1 \ge \bar{\mathcal{L}}^2 \hat{\mathcal{L}}^N$ and **I1** will be satisfied for $j = 1$. Recall that $\hat{\mathcal{L}} > 1$ so any such $\mathcal{L}_1$ will satisfy $\mathcal{L}_1 > \bar{\mathcal{L}}^2$. Hence Eq. (6) implies that for all $s \in S_0$, any such $\mathcal{L}_1$ will satisfy:

$$\frac{1}{\mathcal{L}_1} < \frac{\beta_{1k}(\pi_1(s))}{\beta_{1n}(\pi_1(s))} < \mathcal{L}_1$$

for all even k and n with $2 \le k, n \le N-1$. Hence such a $\mathcal{L}_1$ will imply that **I2** holds for $j = 1$.

I3 for $j = 1$ simply states that:

$$\frac{1}{\mathcal{L}_1} < \frac{\beta_{2\ell}(\pi_2^*)}{\beta_{2m}(\pi_2^*)} < \mathcal{L}_1$$

for all odd ℓ and m with $3 \le \ell, m \le N-2$. Because there are only finitely many inequalities here, we can obviously choose $\mathcal{L}_1 \ge \bar{\mathcal{L}}^2 \hat{\mathcal{L}}^N$ to satisfy this. Hence **I3** holds for $j = 1$.

To complete the induction, fix any value of $j \le L$ and suppose we have demonstrated that **I1**, **I2**, and **I3** hold for all smaller values of j. Without loss of generality, suppose j is even. (If j is odd, reverse the roles of players 1 and 2.) By the induction hypothesis, we know that there is $\mathcal{L}_{j-1}$ such that:

$$\frac{1}{\mathcal{L}_{j-1}} < \frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2m}(s)} < \mathcal{L}_{j-1}$$

for all $s \in S_{j-1}$, every even n with $2(j-1) \le n \le N-2j+3$, and every odd m with $2j-1 \le m \le N-2j+2$. Also,

$$\frac{1}{\mathcal{L}_{j-1}} < \frac{\beta_{1k}(\pi_1(s))}{\beta_{1n}(\pi_1(s))} < \mathcal{L}_{j-1}$$

for every $s \in S_{j-2}$ for all even k and n with $2(j-1) \le k, n \le N-2j+3$. Finally,

$$\frac{1}{\mathcal{L}_{j-1}} < \frac{\beta_{2\ell}(\pi_2(s))}{\beta_{2m}(\pi_2(s))} < \mathcal{L}_{j-1}$$

for every $s \in S_{j-2}$ for all odd ℓ and m with $2j-1 \le \ell, m \le N-2j+2$.

Since j is even, $j-1$ is odd. Hence S_{j-1} is a union of events in Π_1. In particular, for any $s \in S_{j-1}$, there is an $s' \in \pi_1(s)$ such that $s' \in S_{j-2}$. Hence for any $s \in S_{j-1}$, **I2** at $j-1$ implies:

$$\frac{1}{\mathcal{L}_{j-1}} < \frac{\beta_{1k}(\pi_1(s))}{\beta_{1n}(\pi_1(s))} < \mathcal{L}_{j-1} \tag{9}$$

for every $s \in S_{j-1}$ for all even k and n with $2(j-1) \le k, n \le N-2j+3$. Hence **I2** will hold for j if we choose any $\mathcal{L}_j \ge \mathcal{L}_{j-1}$.

With this in mind, fix any $s \in S_{j-1}$ and let $\pi_2 = \pi_2(s)$ and $\pi_1 = \pi_1(s)$. Fix any even k and n with $2(j-1) \le k, n \le N-2j+3$. Fix any odd ℓ and m with $2j-1 \le \ell, m \le N-2j+2$. Then:

$$\frac{\beta_{2\ell}(\pi_2)}{\beta_{2m}(\pi_2)} = \frac{\beta_{1k}(\pi_1)}{\beta_{1n}(\pi_1)} \frac{\hat{\mu}_{1n}(s)/\hat{\mu}_{2m}(s)}{\hat{\mu}_{1k}(s)/\hat{\mu}_{2\ell}(s)}.$$

From **I1** at $j-1$ and Eq. (9), the right-hand side is larger than $1/\mathcal{L}^3_{j-1}$ and less than $\mathcal{L}^3_{j-1}$. Hence **I3** will hold at j if we take any $\mathcal{L}_j \ge \mathcal{L}^3_{j-1}$.

To complete the argument, continue to assume $s \in S_{j-1}$ and to let $\pi_i = \pi_i(s)$, $i = 1, 2$. Obviously,

$$\frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2m}(s)} = \frac{\beta_{2\ell}(\pi_2)}{\beta_{2m}(\pi_2)} \frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2\ell}(s)}$$

for any even n and odd m and ℓ. Fix any even n with $2(j-1) \le n \le N-2j+3$ and any odd ℓ and m with $2j-1 \le \ell, m \le N-2j+2$ such that $\ell > n > m$. From Eq. (4), we know that:

$$\frac{1}{\mathcal{L}^N} < \frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2m}(s)}$$
$$\frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2\ell}(s)} < \mathcal{L}^N.$$

Hence for all $s \in S_j$,

$$\frac{1}{\mathcal{L}^N} < \frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2m}(s)}, \qquad \frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2\ell}(s)} < \mathcal{L}^3_{j-1}\mathcal{L}^N \tag{10}$$

I claim that this implies:

$$\frac{1}{\mathcal{L}^N} < \frac{\hat{\mu}_{1n}(s)}{\hat{\mu}_{2m}(s)} < \mathcal{L}^3_{j-1}\mathcal{L}^N$$

for all $s \in S_j$, all even n with $2j \le n \le N-2j+1$, and all odd m with $2j+1 \le m \le N-2j$. To see this, fix any such n and m. First, suppose $n > m$. Let $\ell = N-2j+2$. Since $n \le N-2j+1$, we must have $\ell > n > m$, so Eq. (10) implies the claim. Next, suppose that $m > n$. In this case, change variables to use ℓ in place of m. Let $m = 2j-1$. Since $n \ge 2j$, clearly, $\ell > n > m$. So, again, Eq. (10) implies the claim. In light of this, we can choose any $\mathcal{L}_j \ge \mathcal{L}^3_{j-1}\hat{\mathcal{L}}^N$ and **I1** will hold at j. Since any such $\mathcal{L}_j$ will be larger than $\mathcal{L}_{j-1}$ and $\mathcal{L}^3_{j-1}$, **I2** and **I3** hold at j as well, completing the induction.

Hence support consistency is necessary for a state to be finitely consistent with the common prior assumption.

Sufficiency. To show sufficiency, suppose $\mathcal{M}$ satisfies support consistency and tail consistency. Let $\mathcal{L}$ be the constant tail consistency requires and let $\hat{\mu}_i$, $i \in \mathcal{I}$, denote the priors it requires. I define a new model $\bar{\mathcal{M}}$ as follows. Let $\bar{S} = S \times \{1, \ldots, NI\}$. Let $\bar{f}(s, k) = f(s)$ for all k. To define the partition for player 1, first fix any $\pi \in \Pi_1$. Then $\bar{\Pi}_1$ includes the sets:

$$\{(s, 1) | s \in \pi\}$$
$$\{(s, k) | k = 2, \ldots, I+1, s \in \pi\}$$
$$\{(s, k) | k = I+2, \ldots, 2I+1, s \in \pi\}$$
$$\vdots$$
$$\{(s, k) | k = (N-1)I+1, \ldots, NI, s \in \pi\}$$

The rest of $\bar{\Pi}_1$ is defined analogously. For $\bar{\Pi}_2$, again fix $\pi \in \Pi_2$. $\bar{\Pi}_2$ includes the sets:

$$\{(s,k)|k = 1, 2, s \in \pi\}$$
$$\{(s,k)|k = 3, \ldots, I+2, s \in \pi\}$$
$$\vdots$$
$$\{(s,k)|k = (N-1)I+3, \ldots, NI, s \in \pi\}$$

and so on. In short, $\bar{\Pi}_i$ consists of sets of the form:

$$\{(s,k)|k = 1, \ldots, i, s \in \pi\}$$
$$\{(s,k)|k = i+1, \ldots, I+i, s \in \pi\}$$
$$\vdots$$
$$\{(s,k)|k = (N-1)I+i+1, \ldots, NI\}$$

for each $\pi \in \Pi_i$.

I now define $\bar{\mu}$. For an $\alpha > 0$ to be determined below, let:

$$\bar{\mu}(s,1) = \alpha\hat{\mu}_1(s)$$
$$\bar{\mu}(s,2) = \mathcal{L}\alpha\hat{\mu}_2(s) - \bar{\mu}(s,1)$$
$$\bar{\mu}(s,i) = \mathcal{L}^{i-1}\alpha\hat{\mu}_i(s) - \sum_{j=1}^{i-1}\bar{\mu}(s,j)$$
$$\bar{\mu}(s,I+1) = \mathcal{L}^{I}\alpha\hat{\mu}_1(s) - \sum_{j=1}^{I}\bar{\mu}(s,j)$$
$$\bar{\mu}(s,kI+i) = \mathcal{L}^{kI+i-1}\alpha\hat{\mu}_i(s) - \sum_{j=1}^{kI+i-1}\bar{\mu}(s,j).$$

Note that $\bar{\mu}(s,kI+i) > 0$ iff:

$$\mathcal{L}^{kI+i-1}\alpha\hat{\mu}_i(s) > \sum_{j=1}^{kI+i-1}\bar{\mu}(s,j).$$

For $i \geq 2$, the right-hand side is $\mathcal{L}^{kI+i-2}\alpha\hat{\mu}_{i-1}(s)$. Hence if $\hat{\mu}_i(s) > 0$, the inequality holds iff:

$$\mathcal{L} > \frac{\hat{\mu}_{i-1}(s)}{\hat{\mu}_i(s)},$$

which must hold by tail consistency. Hence if $\mu_j(s|\pi_j(s)) > 0$, we have $\bar{\mu}(s,k) > 0$ for all k, while if $\mu_j(s|\pi_j(s)) = 0$, $\bar{\mu}(s,k) = 0$ for all k. Also,

$$\sum_{s\in S}\sum_{k=1}^{NI}\bar{\mu}(s,k) = \sum_{s\in S}\mathcal{L}^{NI-1}\alpha\hat{\mu}_I(s) = \mathcal{L}^{NI-1}\alpha.$$

Hence by setting $\alpha = (1/\mathcal{L})^{NI-1}$, we ensure that this sum is one.

The following lemma shows that this construction yields a model with the appropriate properties.

Lemma 3. *For all s and n, $(s,k)\sim_n^* s$ for $k = 1, \ldots, (N-n)I+n$.*

The proof of this lemma is identical to the proof of the same lemma in Lipman (2003) and so is omitted.

In light of the lemma, we see that $(s,k)\sim_N^* s$ for $k = 1, \ldots, N$. Hence we see that $(\mathcal{M}, s) \in \mathrm{CP}_N$ for all s. Since N is arbitrary, $(\mathcal{M}, s) \in \mathrm{CP}^*$. □

References

Aumann, R., 1976. Agreeing to disagree. Annals of Statistics 4, 1236–1239.
Brandenburger, A., Dekel, E., 1993. Hierarchies of beliefs and common knowledge. Journal of Economic Theory 59 (February), 189–198.

Bonnano, G., Nehring, K., 1999. How to make sense of the common prior assumption under incomplete information. International Journal of Game Theory 28 (August), 409–434.
Feinberg, Y., 2000. Characterizing common priors in the form of posteriors. Journal of Economic Theory 91 (April), 127–179.
Frankel, D., Morris, S., Pauzner, A., 2003. Equilibrium selection in global games with strategic complementarities. Journal of Economic Theory 108 (January), 1–44.
Halpern, J., 2002. Characterizing the common prior assumption. Journal of Economic Theory 106 (October), 316–355.
Heifetz, A., 2006. The positive foundation of the common prior assumption. Games and Economic Behavior 56 (July), 105–120.
Lipman, B., 2003. finite order implications of common priors. Econometrica 71 (July), 1255–1267.
Mertens, J.-F., Zamir, S., 1985. Formalization of Bayesian analysis for games with incomplete information. International Journal of Game Theory 14, 1–29.
Milgrom, P., Stokey, N., 1982. Information, trade, and common knowledge. Journal of Economic Theory 26, 17–27.
Monderer, D., Samet, D., 1989. Approximating common knowledge with common beliefs. Games and Economic Behavior 1, 170–190.
Morris, S., 1995. The common prior assumption in economic theory. Economics and Philosophy 11, 227–253.
Rubinstein, A., 1989. The electronic mail game: a game with almost common knowledge. American Economic Review 79 (June), 385–391.
Samet, D., 1998a. Iterated expectations and common priors. Games and Economic Behavior 24 (July), 131–141.
Samet, D., 1998b. Common priors and separation of convex sets. Games and Economic Behavior 24 (July), 172–174.
Savage, L.J., 1954. The Foundations of Statistics. Wiley, New York.
Weinstein, J., Yildiz, M., 2007. A structure theorem for rationalizability with application to robust predictions of refinements. Econometrica 75 (March), 365–400.