Queen’s Economics Department Working Paper No. 974

Common Knowledge With Monotone Statistics

James Bergin
Queen’s University

Department of Economics
Queen’s University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

9-1998

COMMON KNOWLEDGE WITH MONOTONE STATISTICS

by

James Bergin

Department of Economics
Queen's University
Kingston, ON K7L 3N6
Canada

18 September 1998

Abstract

When individual statistics are aggregated through a strictly monotone function to an aggregate statistic, common knowledge of the value of the aggregate statistic does not imply, in general, constancy of the individual statistics. This paper discusses two circumstances where it does occur. The first case arises when partitions are independently drawn: in this case common knowledge of the value of the aggregator function implies (with probability one) constancy of the individual statistics. The second case is where private statistics are related: affiliation of individual statistics and a lattice condition imply constancy of the individual statistics when the value of the aggregate statistic is common knowledge.

1 Introduction.

A well known paper of Aumann (1976) formalized the notion of common knowledge and provided a mathematically tractable way of modeling hierarchies of knowledge. In addition he gave an equivalence theorem on the value of posterior distributions when a common knowledge condition is satisfied: when the values of every agent's posterior distribution is common knowledge, then the posteriors are equal. This result was strengthened substantially by McKelvey and Page (1986) who showed that common knowledge of an aggregate statistic of the posteriors was sufficient to imply equality of the posteriors. The key requirement of the aggregating function was stochastic monotonicity, or equivalently, additive separability into strictly monotone components. This paper considers the case where the aggregator function is strictly monotone, but not necessarily additively separable. In this case (strict monotonicity without additive separability), common knowledge of the value of the aggregating function does not imply constancy of the posteriors (see McKelvey and Page for a counter-example). Nevertheless, additive separability is a significant restriction and it is natural to consider circumstances under which it may be relaxed. Here, we consider two such situations: (a) when agents' partitional information is independently drawn and (b) when agents' signals are affiliated.

The two main results in the paper are (roughly) the following. If agent's partitions are drawn randomly and with probability 1 the cardinality of each partition is small relative to the cardinality of the state space, then common knowledge of the value of a strictly monotone aggregator function implies constancy of each of its arguments. In the special case where the individual statistics are expectations of some random variable conditional on private information, this implies equivalence of all the arguments. The second result is that if that if individual agents statistics are affiliated, then, under a lattice condition, common knowledge of the value of the aggregator function implies constancy of the arguments of the function.

2 The Framework.

A probability space, $(\Omega, \mathcal{F}, p)$ is given. Each player $i \in N = \{1, \ldots, n\}$ is defined by a signal (a random variable) X_i, and an information partition $\mathcal{P}_i$ on Ω. Let $P_i(\omega)$ denote the element of $\mathcal{P}_i$ containing ω. Write $\mathcal{M} = \wedge_j \mathcal{P}_i$, to denote the meet of $\{\mathcal{P}_i\}_{i=1}^n$ or finest common coarsening with $M(\omega)$ the element of $\mathcal{M}$ containing ω. Similarly, let $\mathcal{J} = \vee_j \mathcal{P}_j$ denote the join of $\{\mathcal{P}_i\}_{i=1}^n$ or coarsest common refinement. Finally, let $\mathcal{G}$ denote the set of all partitions of Ω. With this notation, common knowledge of an event is defined:

Definition 1 *An event Q is common knowledge at ω^* if $M(\omega^*) \subseteq Q$.*

Given an event $A \subset \Omega$, let $X_i(\omega) = q_i(\omega) = p(A \mid P_i(\omega)) \stackrel{\text{def}}{=} E\{\chi_A \mid P_i(\omega)\}$, where χ_A is the characteristic function of A and $p(A \mid P_i(\omega))$ posterior distribution on some event A. Fix $\bar{q} = (\bar{q}_1, \ldots, \bar{q}_n)$ and let $E = \{\omega \mid \forall i,\ q(\omega) = \bar{q}_i\}$, where $q(\omega) = (q_1(\omega), \ldots, q_n(\omega))$. The result of Aumann is that common knowledge of posteriors implies that they are equal:

Theorem 1 *If the event E is common knowledge at ω^*, then for all i and j, $\bar{q}_i = \bar{q}_j$.*

Because $M(\omega^*) \subseteq E$, common knowledge of E requires that *each* agent's posterior distribution q_i is constant on the meet, $M(\omega^*)$. A substantial improvement on this theorem was obtained by McKelvey and Page (1986) who considered the case where posteriors are aggregated according to some function f. Consider the event $E' = \{\omega \mid f(q(\omega)) = c\}$, where c is a constant. McKelvey and Page proved that when f is stochastically monotone then common knowledge of E' implies that the posteriors are constant and equal.

Theorem 2 *If the event E' is common knowledge at ω^*, then for all i and j, $q_i(\omega) = q_j(\omega)$, $\forall \omega \in M(\omega^*)$.*

A simplified proof of McKelvey and Page's theorem is given by Nielsen, Brandenburger, Geanakoplos, McKelvey and Page (1990) (see also Bergin and Brandenburger (1990)). There, it is also observed that $\{q_i\}_{i=1}^n$ may be replaced by $\{X_i\}_{i=1}^n = \{E\{X \mid P_i\}\}_{i=1}^n$, where X is any random variable. Additive separability of the "aggregating" function f is a strong assumption, and it is natural to ask if the theorem remains valid when stochastic monotonicity is replaced by the weaker assumption of strict monotonicity of f. McKelvey and Page give a simple counter example to this conjecture, so in general the answer is no.

In what follows we first develop a model of independent information partitions and in this context give a common knowledge theorem for a set of measure 1 of information partitions. After that, we consider the case where private signals are affiliated and again provide a theorem on constancy of signals under the common knowledge condition. In both cases the aggregating function is assumed to be strictly monotonic but not necessarily additively separable.

3 Partition and Random Partitions.

In this section, we begin by describing agents with private or independent information — modeled by having partitions drawn randomly and independently — and then show that generically, random partitions have no elements (or unions of elements) in common. This result is used to develop a common knowledge theorem for random partitions.

Denote the set of partitions of Ω by $\mathcal{G}$. In what follows, attention is restricted to the case of partitions of a countable state space $\Omega = \{\omega_i\}_{i\in\mathcal{Z}}$ ($\mathcal{Z}$ is the set of positive integers). A partition can be represented as a point $s = (s_{\omega_1}, s_{\omega_2}, \ldots)$ in $S \stackrel{\text{def}}{=} \mathcal{Z}^\Omega$ ($s_\omega \in \mathcal{Z}$, $\forall\omega\in\Omega$), where $s_\omega = s_{\omega'}$ if and only if ω and ω' belong to the same partition. If $s \in S$, the corresponding partition of Ω, $\varphi(s)$, the partition determined by s, is defined $\varphi(s) = \{\{\omega\in\Omega \mid s_\omega = k\}_{k\in\mathcal{Z}}\} \in \mathcal{G}$. Any given partition of Ω is associated with many points in S,[1] but one may use either formulation — here, it is usually more convenient to view a partition as an element of S. Formulating partitions in terms of S simplifies the discussion of randomization on the set of partitions. A partition of Ω into k non-empty sets is called a partition of size k. Let $S^k = \{s\in S \mid s_\omega \in N_k\}$, where $N_k = \{1,\ldots,k\}$. Random partitions of size k are then identified as "draws" from some distribution on S^k.

Definition 2 *Let $p(k) = \{p_j^k\}_{j=1}^k$ with $p_j^k > 0$ for each j. For each $\omega\in\Omega$ put $\mu_\omega^k = p(k)$ and define a random k-partition as a draw from the measure $\mu^k = \otimes_{\omega\in\Omega}\mu_\omega^k$.*

A partition of Ω is finitely generated if it is an element of S^k for some k. Note that $S^k \subset S^{k+1} \subset S$, and one may view μ^k as a measure on S, where μ^k has support on S^k. Call S^* the set of finitely generated partitions of Ω where $S^* = \cup_{k=1}^\infty S^k$.

Definition 3 *Let $\lambda_k \geq 0$, $\sum_k \lambda_k = 1$. A random partition is a draw from the measure μ on S^*, defined: $\mu = \sum_k \lambda_k\mu_k$, where μ_k is a random k-partition.*

Thus a draw from a random partition is an element $s \in S^*$. With probability 1, the partition is finite, but the expected size of the partition may be infinite.[2]

If $\mathcal{A}$ and $\mathcal{B}$ are two partitions, then although they may have no elements in common, it may be that the union of some members of $\mathcal{A}$ coincide with a member (or union of members) of $\mathcal{B}$. In such a case, some strict subset of Ω would be common knowledge at some state. The next theorem asserts that this is not the case (generically), for random independently drawn partitions. (All proofs are in the appendix.)

Theorem 3 *Let $\{\mathcal{P}_1,\ldots,\mathcal{P}_n\}$ be n independent randomly drawn partitions. Then for each i, for any $G\in\sigma(\vee_{j\neq i}\mathcal{P}_j)$, $G\neq\emptyset$, Ω, there is 0 probability that $G\in\mathcal{P}_i$.*

In words, random partitions have no "overlap", with probability 1. Thus, the property that an event is common knowledge at some ω is non-generic. (The intuition is simple. Suppose there are a hundred balls labeled 1 to 100 and an individual randomly distributes the balls between two urns. This gives a collection of balls in each urn (a partition): C_A and C_B. If the experiment is repeated by another agent,

[1] If $s, \hat{s} \in S$ satisfy (a) $s_\omega = s_{\omega'}$ implies $\hat{s}_\omega = \hat{s}_{\omega'}$ and (b) $s_\omega \neq s_{\omega'}$ implies $\hat{s}_\omega \neq \hat{s}_{\omega'}$, they determine the same partition of Ω.

[2] With probability λ_k, the partition is in S^k with an expected size of $m_k = \sum_{l=1}^k l p_l^k$, so the expected size of the partition is $\sum_{k=1}^\infty \lambda_k m_k$. If $p_j^k = \frac{1}{k}$ then $m_k = \frac{1}{2}(k+1)$ and $\sum_{k=1}^\infty \lambda_k m_k > \frac{1}{2}\sum_{k=1}^\infty \lambda_k k$. For $\epsilon\in(0,1)$, define $\{k_j\}_{j\geq1}$ as $k_j = \min_{k\geq 1/\epsilon^j} k$, $j = 1,2,\ldots$, let $\lambda_{k_i} = (1-\epsilon)\epsilon^i$, $k_i \in \{k_j\}_{j\geq1}$; and $\lambda_k = 0$ otherwise. For these choices, the sum $\sum_{k=1}^\infty \lambda_k m_k$ diverges.

there is very small probability that the same division (C_A and C_B) will be obtained: and as the number of balls becomes larger, this probability goes to 0.)

If we represent "public information" by a partition $\mathcal{H}$, then, for example, if $\mathcal{H} = \{\emptyset, \Omega\}$, there is no useful public information in the sense that $\sigma(\mathcal{P}_i) = \sigma(\mathcal{P}_i) \vee \mathcal{H}$. In general, one will model public knowledge, $\mathcal{H}$, as being finer than $\{\emptyset, \Omega\}$. (In the appendix we show that, if $\mathcal{H}$ is a random partition then with probability 1 $\mathcal{H}$ consists of a finite number of sets, each with an infinite number of elements.) On each member of $\mathcal{H}$ having an infinite number of elements the partitions induced by the $\{\mathcal{P}_i\}$ have no members (or unions of members) in common: $\mathcal{P}_i$ and $\vee_{j \neq i} \mathcal{P}_j$ have no overlap. In this case, each element of $\mathcal{H}$ is common knowledge, and with probability 1, these are the only events that are common knowledge. In particular, given ω, there is some ω' and i with ω and ω' in different partition members for i and in the same partition member for all $j \neq i$. If $X(\tilde{\omega}) = (X_1(\tilde{\omega}), \ldots X_n(\tilde{\omega}))$ is a R^n-valued random variable with X_i constant on each member of $\mathcal{P}_i$, and f a strictly monotone function on R^n, then $f(X(\omega) = f(X(\omega'))$ implies $X_i(\omega) = X_i(\omega')$. This discussion leads to the main theorem of the section.

Theorem 4 *Let f be a strictly monotone function, $f : R^n \to R$. Let $\{\mathcal{P}_i\}_{i=1}^n$, $\mathcal{H}$ be random independently drawn partitions, and $\mathcal{P}_i^* = \mathcal{P}_i \vee \mathcal{H}$. Let $\{X_i\}_{i=1}^n$ be a collection of random variables on Ω, such that X_i is $\mathcal{P}_i^*$ measurable. Put $g(\omega) = f(X(\omega))$, $C = \{c \mid \exists\ \omega,\ g(\omega) = c\}$ and for $c \in C$, let $E_c = \{\omega \mid g(\omega) = c\}$. With probability 1, if E_c is common knowledge at ω^*, then for each i, X_i is constant on E_c.*

So, if private and public information partitions are independently and randomly drawn, then with probability 1, common knowledge of the value of the aggregating function implies constancy of the individual statistics.

4 Co-varying Signals.

This section provides a common knowledge result in the case where the aggregator function is not additively separable, but restrictions are imposed on the distributions of the random variables, or on the measure over the underlying probability space. Taking the partition structure as given, what conditions of association on the individual statistics lead to common knowledge of the constancy of the aggregator function implying constancy of the individual statistics? Here we show that affiliation plus a lattice condition yield the result.

A key feature of additive separability of the aggregator function is the co-variation of the aggregate function with the individual signals. Consider two points ω and ω' with $x = X(\omega)$ and $x' = X(\omega')$. Even though the vector of changes $\{x_i' - x_i\}_{i=1}^n$ may not all have the same sign, the vector of changes $\{[f(x_i') - f(x_i)][x_i' - x_i]\}_{i=1}^n$ are all non-negative, and strictly positive when for some i, $[x_i' - x_i] \neq 0$ because each f_i is strictly increasing. In this case, constancy of $\sum f_i$ on the range of $X = \{X_i\}_{i=1}^n$, over $M_i(\omega^*)$, is inconsistent with variation of any X_i on the meet. When f is not additively separable, without some restriction on the co-variation on the variables $\{X_i\}_{i=1}^n$, constancy of the aggregator function cannot imply constancy of the individual X_i's: an upward movement in one could be offset by a downward movement of another, compensating to a constant value of f. This suggests that if the co-movements of the variables are restricted, a common knowledge result might continue to hold. The restriction considered here is that of affiliation.

Definition 4 *A random vector X (in $\mathcal{R}^k$) with density ρ is said to be affiliated (or multivariate totally positive of order 2 (MTP$_2$)) if ρ satisfies $\rho(x \vee y)\rho(x \wedge y) \geq \rho(x)\rho(y)$.*

(Here $x \vee y = (\max(x_1, y_1), \ldots, \max(x_n, y_n))$ and $x \wedge y = (\min(x_1, y_1), \ldots, \min(x_n, y_n))$.) Note that if ρ satisfies the affiliation condition, then on any sub-lattice, S, of $\mathcal{R}^k$ $(z, z' \in S \Rightarrow z \vee z' \in S,\ z \wedge z' \in S)$, the affiliation condition is also satisfied: the density, $\bar{\rho}$ of the random vector X, conditional on being in S is an affiliated density: $\bar{\rho}(x) = \frac{1}{\rho(S)}\rho(x),\ x \in S$.

Theorem 5 *Let $X : \Omega \to R^n$ be an affiliated random vector on Ω. Let $f : R^n \to R$ be a strictly monotone function and set $g(\omega) = f(X(\omega))$. Suppose that:*

1. *At ω^* it is common knowledge that $g = c$,*
2. *$S = X(M(\omega^*))$ is a lattice.*

Then $X(\omega) = (\bar{x}_1, \ldots, \bar{x}_n)$, $\forall \omega \in M(\omega^*)$.

Proof: Since the event $E = \{\omega \mid f(X(\omega)) = c\}$, where c is a constant and suppose that E is common knowledge at ω^*, so for any constant k, $f(X(\omega)) - k = c - k$, $\forall \omega \in M(\omega^*)$. For $\omega \in M(\omega^*)$, let $\bar{p}(\omega) = \frac{1}{p(M(\omega^*))}p(\omega)$, $\bar{p}(\omega) = 0, \omega \in M(\omega^*)$, let $\mu_i = \sum_{\omega \in M(\omega^*)} X_i(\omega)\bar{p}(\omega)$ and let $\hat{X}$ be the random vector with distribution determined by $\bar{\rho}$ — so that X conditional on S has the same distribution as $\hat{X}$. Then,

$$\begin{aligned} 0 = &\sum_{\omega \in M(\omega^*)} [f(X(\omega)) - k][X_i(\omega) - \mu_i]\bar{p}(\omega) \\ =&E_\rho\{[f(X) - k][X_i - \mu_i] \mid S\}, \quad S = X(M(\omega^*)) \\ =&E_{\bar{\rho}}\{[f(\hat{X}) - k][\hat{X}_i - \mu_i]\} \\ =&E_{\bar{\rho}}\{E_{\bar{\rho}}\{[f(\hat{X}) - k] \mid \hat{X}_i\}[\hat{X}_i - \mu_i]\} \end{aligned}$$

Suppose that $\hat{x}_i^a < \hat{x}_i^b$, then since f is strictly monotonic, $f(\hat{x}_i^a, \hat{X}_{-i}) < f(\hat{x}_i^b, \hat{X}_{-i})$, $\forall \hat{X}_{-i}$, and so $E_{\bar{\rho}}\{f(\hat{x}_i^a, \hat{X}_{-i}) \mid \hat{X}_i^a = \hat{x}_i^a\} < E_{\bar{\rho}}\{f(\hat{x}_i^b, \hat{X}_{-i}) \mid \hat{X}_i^a = \hat{x}_i^a\}$. Because $\hat{X}$ is affiliated, $E_{\bar{\rho}}\{f(\hat{x}_i^b, \hat{X}_{-i}) \mid \hat{X}_i^a = \hat{x}_i^a\} \leq E_{\bar{\rho}}\{f(\hat{x}_i^b, \hat{X}_{-i}) \mid \hat{X}_i^b = \hat{x}_i^b\}$, so noting that $E_{\bar{\rho}}\{f(\alpha, \hat{X}_{-i}) \mid \hat{X}_i^a = \alpha\} = E_{\bar{\rho}}\{f(\hat{X}_i^a, \hat{X}_{-i}) \mid \hat{X}_i^a = \alpha\}$, and combining inequalities:

$$\hat{X}_i^a < \hat{X}_i^b \quad \Rightarrow \quad E_{\bar{\rho}}\{f(\hat{X}_i^a, \hat{X}_{-i}) \mid \hat{X}_i^a\} < E_{\bar{\rho}}\{f(\hat{X}_i^b, \hat{X}_{-i}) \mid \hat{X}_i^b\}.$$

Thus, $E_{\bar{\rho}}\{f(\hat{X}_i, \hat{X}_{-i}) \mid \hat{X}_i\}$ is strictly increasing in $\hat{X}_i$. Choose so that $E_{\bar{\rho}}\{[f(\hat{X}) - k] \mid \hat{X}_i\} > 0$ when $\hat{X}_i > \mu_i$ and $E_{\bar{\rho}}\{[f(\hat{X}) - k] \mid \hat{X}_i\} < 0$ when $\hat{X}_i < \mu_i$. Thus, $E_{\bar{\rho}}\{E_{\bar{\rho}}\{[f(\hat{X}) - k] \mid \hat{X}_i\}[\hat{X}_i - \mu_i] \mid\} \geq 0$, with strict inequality unless $E_{\bar{\rho}}\{\ \mid [\hat{X}_i - \mu_i] \mid\ \} = 0$.

∎

The result can be strengthened substantially by "piecing" sub-lattices together. Call $\{L_j\}_{j=1}^J$ an overlapping paving of the set (not necessarily lattice) S if (a) each $L_j \subset S$ is a lattice, (b) for any $x,\ x' \in S$, $\exists x = x_1, \ldots, x_J = x'$ such that $x_j \in L_j \cap L_{j+1}$, $j = 1, \ldots, J - 1$. In this case say that S has an overlapping paving.

Corollary 1 *Condition 2 in the previous theorem can be replace by the requirement that S has an overlapping paving.*

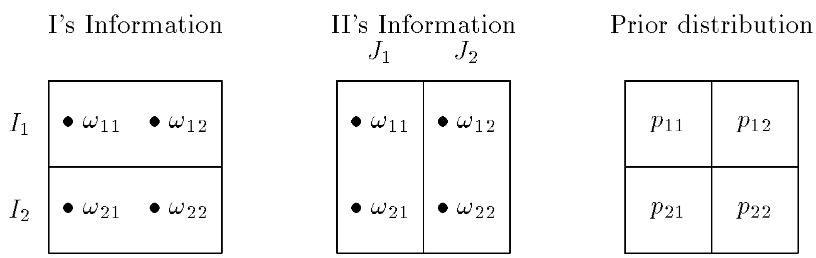
Proof: On each sub-lattice L_j, proceed as in the theorem. On overlapping sub-lattices, equivalence of the values of X at the intersections with constancy on each sub-lattice implies a constant value of on any such pair of sub-lattices. Since all the sub-lattices are connected, this implies constancy over the union of the sub-lattices.

∎

A similar result may be obtained when the distributional restriction is placed on the underlying probability space (so $\Omega = \times_{i=1}^{n}\Omega_i$, where Ω_i is totally ordered, the distribution on Ω affiliated and $X_i : \Omega_i \to R$.)

4.1 Affiliation and Posterior Distributions.

Finally, we conclude with an example showing that, in general, posterior distributions on the same event are not affiliated. The state space is $\Omega = \{\omega_{11}, \omega_{12}, \omega_{21}, \omega_{22}\}$ and ω_{ij} has prior probability p_{ij}. The information structures and probabilities are given by the following figure.



Let $A = \{\omega_{11}, \omega_{22}\}$, so $p_i = \text{prob}(A \mid I_i)$ and $q_i = \text{prob}(A \mid J_i)$ are the posterior distributions conditional on the information. Take $\text{prob}(I_i)\text{prob}(J_j) > 0$ for $i, j \in \{1, 2\}$, so the posterior distributions are unambiguously defined. Thus,

$$p_1 = \frac{p_{11}}{p_{11} + p_{12}}, \quad p_2 = \frac{p_{22}}{p_{21} + p_{22}}, \quad q_1 = \frac{p_{11}}{p_{11} + p_{21}}, \quad q_2 = \frac{p_{22}}{p_{12} + p_{22}}$$

The distribution of the posterior distributions is: $\rho(p_i, q_j) = p_{ij}$, $i, j \in \{1, 2\}$. Suppose that $p_2 \geq p_1$ and $q_2 \geq q_1$. In this case, there is one affiliation inequality to be satisfied: $\rho(p_1, q_1)\rho(p_2, q_2) \geq \rho(p_1, q_2)\rho(p_2, q_1)$, or $p_{11}p_{22} \geq p_{12}p_{21}$ ($(p_1, q_2) \vee (p_2, q_1) = (p_2, q_2)$ and $(p_1, q_2) \wedge (p_2, q_1) = (p_1, q_1)$). So, the posteriors are not affiliated when, for example, $p_{11} = .1$, $p_{12} = p_{22} = p_{21} = .3$ (while $p_2 \geq p_1$ and $q_2 \geq q_1$ at these values).

This observation may also be seen in terms of expectations. If q is the posterior function for person 2, $q(\omega_{ij}) = \text{prob}(A \mid J(\omega_{ij}))$ where $J(\omega_{ij})$ is the partition member containing ω_{ij}, then, in terms of expectations, some calculations yield:

$$\begin{aligned} E(q \mid p_2) - E(q \mid p_1) =& (q_2 - q_1)\{\frac{p_{22}p_{11} - p_{21}p_{12}}{(p_{21} + p_{22})(p_{11} + p_{12})}\} \\ =& (q_2 - q_1)\{\frac{\rho(p_1, q_1)\rho(p_2, q_2) - \rho(p_1, q_2)\rho(p_2, q_1)}{[\rho(p_2, q_1) + \rho(p_2, q_2)][\rho(p_1, q_1) + \rho(p_1, q_2)]}\} \end{aligned}$$

Although $p_2 \geq p_1$ and $q_2 \geq q_1$, the term on the left is negative if the affiliation inequality fails.

Appendix

In the appendix we first establish some properties of random partitions that are used in the proofs of theorems 3 and 4.

In the set S^k, one may identify partitions of size less than k. For example, the partition consisting of just one member Ω is identified with $s = \{s_\omega\}_{\omega\in\Omega}$ where for some $j \in \{1, \ldots, k\}$, $s_\omega = j$, $\forall \omega \in \Omega$. However, a random k partition puts probability 0 on partitions of size less than k.

Lemma 1 *Let μ^k be a random k-partition and $\bar{S}^k = \{s \in S^k \mid \exists j \in \{1, \ldots, k\}, \forall \omega \in \Omega, s_\omega \neq j\}$. Then $\mu^k(\bar{S}^k) = 0$. Furthermore, with probability 1, each member of the partition has an infinite number of elements.*

Proof: The proof of the first part of the lemma is immediate: $\mu^k(\bar{S}^k) \leq \sum_{j=1}^k [\times_{\omega\in\Omega}(1 - p_j^k)] = 0$. Thus, in a random k partition, the partition has k members with probability 1, with $\mu^{k+1}(S^k) = 0$, $\mu^{k+1}(S^{k+1} \setminus S^k) = 1$, and since $S^k \subset S^{k+1}$, $\mu^k(S^{k+j}) = 1$, $j \geq 0$. The next observation establishes the second claim. To see that in a random k-partition, with probability 1 each member of the partition has an infinite number of elements, let

$$\xi_{\omega_j}^r = \begin{cases} 1, & \text{if } s_{\omega_j}^r = r \\ 0, & \text{if } s_{\omega_j}^r \neq r \end{cases}$$

and note that (by the law of large numbers), for almost all draws of $s \in S^k$

$$\frac{1}{n}\sum_{j=1}^{n} \xi_j^r \to p_r > 0.$$

This implies that for almost all $s \in S^k$, $s_{\omega_j} = r$ for an infinite number of times.

∎

Thus, with probability 1, each element of a random partition contains an infinite number of elements, and with probability λ_k, the partition has k members.

Partitions $s(1), s(2), \ldots s(r) \in S^*$ are drawn independently if drawn from a distribution $\mu = \otimes_{i=1}^r \mu(i)$ on $(S^*)^r$, with $\mu(i)$ a random partition. In terms of partitions defined directly on Ω, the partitions $s(1), s(2), \ldots s(r) \in S^*$ are interpreted as $(\varphi(s(1)), \varphi(s(2)), \ldots \varphi(s(r))) = (\mathcal{P}_1, \ldots, \mathcal{P}_r)$, and the vector $(\mathcal{P}_1, \ldots, \mathcal{P}_r)$ is a vector of random independently drawn partitions. Given a partition $\mathcal{E}$, let $\sigma(\mathcal{E})$ be the set of subsets of Ω obtained by taking unions of members of $\mathcal{E}$ (the sigma field generated by $\mathcal{E}$). Next we show that if the partitions $\{\mathcal{P}\}_{i=1}^r$ are random independently drawn partitions, then with probability 1, the only elements $\sigma(\mathcal{P}_i)$ and $\sigma(\mathcal{P}_j)$ have in common are $\emptyset$ and Ω; or more generally, the only elements $\sigma(\mathcal{P}_i)$ and $\sigma(\vee_{j\neq i}\mathcal{P}_j)$ have in common are $\emptyset$ and Ω. Consequently, "generically", the only event that is common knowledge is Ω. These observations are formulated in lemmas 2 and 3.

The next lemma asserts that given any draw, $s \in S^k$, from a random k-partition, the probability that this partition has any overlap with any other partition (common element or union of elements) is 0. Given $s \in S$, let $r(s) = \{j \mid \exists\, \omega \in \Omega, s_\omega = j\}$, the "range" of s, and $\Omega_j(s) = \{\omega \mid s_\omega = j\}$, those points ω that s "assigns" the value j. For $I \subset N_k$, $\Omega_I(s) = \cup_{i\in I}\Omega_i(s)$. Elements of a partition may be combined to form additional sets. Let $s \in S^k$. Say that $\hat{\Omega}$ *is generated* by s if there is some $I \subset N_k$,

$s_\omega \in I$, $\omega \in \hat{\Omega}$ and $s_\omega \in I^c$, $\omega \in \hat{\Omega}^c$. Thus, $\hat{\Omega}$ may be written as the union of members of the partition determined by s if and only if this condition is satisfied. Let $s \in S^k$

$$\phi^k(s) = \{s' \in S^k \mid \exists I, J \subset N_k,\ \Omega_J(s') = \Omega_I(s)\}.$$

So, $\phi^k(s)$ consists of those partitions $s' \in S^k$ generating some Ω' that is also generated by s.

Lemma 2 *In a random k-partition, for any $s \in S^k$,*

$$\mu^k(\{\tilde{s} \in S^k \mid \tilde{s} \in \phi^k(s)\}) = 0.$$

Proof: Since s generates a finite number of subsets of Ω, it is sufficient to show that the measure of the set of partitions generation any set Ω is 0. Let I be a subset of K with $\emptyset \neq I \neq K$. Let $E^I_{\hat{\Omega}} = \{s \mid s_\omega \in I, \forall \omega \in \hat{\Omega}\}$ and $N^I_{\hat{\Omega}} = \{s \mid s_\omega \notin I, \forall \omega \in \hat{\Omega}^c\}$, where $\hat{\Omega} \subset \Omega$. Put $S^I_{\hat{\Omega}} = E^I_{\hat{\Omega}} \cap N^I_{\hat{\Omega}^c}$. Thus, $\mu^k(S^I_{\hat{\Omega}}) \leq \max\{\mu^k(E^I_{\hat{\Omega}}), \mu^k(N^I_{\hat{\Omega}})\}$. At least one of the sets $\hat{\Omega}$ and $\hat{\Omega}^c$ is infinite, and since $\mu^k(E^I_{\hat{\Omega}}) = \times_{\omega \in \hat{\Omega}} \mu_\omega(I) = \times_{\omega \in \hat{\Omega}}[\sum_{k \in I} p_k]$ and $\mu^k(N^I_{\hat{\Omega}^c}) = \times_{\omega \in \hat{\Omega}^c} \mu_\omega(I^c) = \times_{\omega \in \hat{\Omega}^c}[\sum_{k \in I^c} p_k]$. Since $0 < \sum_{k \in I} p_k = 1 - \sum_{k \in I^c} p_k < 1$, $\mu^k(S^I_{\hat{\Omega}}) = 0$. Since I is finite, there are a finite number of S^I, $I \subset k$, so that:

$$\mu^k(S_{\hat{\Omega}}) = \sum_{\substack{I \neq \emptyset \\ I \neq \{1,\ldots,k\}}} \mu^k(S^I_{\hat{\Omega}}) = 0.$$

∎

Let $\psi^k(s) = \{\Omega' \subset \Omega \mid \exists R \subset N_k, \Omega' = \Omega_R(s)\}$: $\psi^k(s)$ is the set of subsets of Ω that can be identified with unions of members of the partition (determined by) s. Given two collections of subsets of Ω, $\mathcal{Q}$ and $\mathcal{Q}'$, write $\mathcal{Q} \Delta \mathcal{Q}'$ to denote the set of subsets of Ω common to both $\mathcal{Q}$ and $\mathcal{Q}'$.

Lemma 3 *Let $\mathcal{Q} = \{\Omega_j\}$ be a finite or countable collection of subsets of Ω. Then:*

$$\mu^k(\{s \in S^k \mid \mathcal{Q} \Delta \psi^k(s) \neq \emptyset\}) = 0.$$

Proof: With this notation, the previous lemma asserts that for any $\hat{\Omega}$, $\mu^k(\{s \in S^k \mid \hat{\Omega} \in \psi^k(s)\}) = 0$ or $\mu^k(\{s \in S^k \mid \hat{\Omega} \notin \psi^k(s)\}) = 1$. Since

$$\mu^k(\{s \in S^k \mid \mathcal{Q} \Delta \psi^k(s) \neq \emptyset\}) \leq \sum_j \mu^k(\{s \in S^k \mid \Omega_j \in \psi^k(s)\}) = 0$$

this completes the proof.

∎

For $s \in S^*$, $\exists k$ such that $s \in S^k$, let $k(s) = \min\{k \in K \mid s \in S^k\}$ and define $\psi^*(s) = \psi^{k(s)}(s)$. If $\mathcal{Q} = \{\Omega_j\}$ be a finite or countable collection of sets, then $\mu(\{s \in S^* \mid \mathcal{Q} \Delta \psi^*(s)\} = \emptyset) = 1$, since $\mu(\{s \in S^* \mid \mathcal{Q} \Delta \psi^*(s) = \emptyset\}) = \sum \lambda_k \mu^k(\{s \in S^k \mid \mathcal{Q} \Delta \psi^k(s) = \emptyset\}) = 0$.

Theorem 3 *Let $\{\mathcal{P}_1, \ldots, \mathcal{P}_n\}$ be n independent randomly drawn partitions. Then for each i, for any $G \in \sigma(\vee_{j \neq i} \mathcal{P}_j)$, $G \neq \emptyset$, Ω, there is 0 probability that $G \in \mathcal{P}_i$.*

Proof: By assumption, the probability that all of the partitions $\{\mathcal{P}_j\}_{j\neq i}$, have no more than m elements is no less than $\beta_m = (\sum_{r=1}^m \lambda_r)^{(n-1)}$. Since $\beta_m \uparrow 1$, with probability 1, $\sigma(\vee_{j\neq i}\mathcal{P}_j)$ contains a finite number of sets. In this case, by lemma 3, the probability that $\sigma(\mathcal{P}_i)$ has any set in common with $\sigma(\vee_{j\neq i}\mathcal{P}_j)$ is 0. ∎

Theorem 4 *Let f be a strictly monotone function, $f : R^n \to R$. Let $\{\mathcal{P}_i\}_{i=1}^n$, $\mathcal{H}$ be random independently drawn partitions, and $\mathcal{P}_i^* = \mathcal{P}_i \vee \mathcal{H}$. Let $\{X_i\}_{i=1}^n$ be a collection of random variables on Ω, such that X_i is $\mathcal{P}_i^*$ measurable. Put $g(\omega) = f(X(\omega))$, $C = \{c \mid \exists\ \omega,\ g(\omega) = c\}$ and for $c \in C$, let $E_c = \{\omega \mid g(\omega) = c\}$. With probability 1, if E_c is common knowledge at ω^*, then for each i, X_i is constant on E_c.*

Proof: Let $\{\mathcal{P}_1, \ldots, \mathcal{P}_n, \mathcal{H}\}$ be $n+1$ independent randomly drawn partitions, and for $i = 1, \ldots, n$, let $\mathcal{P}_i^* = \mathcal{P}_i \vee \mathcal{H}$. Then for each i, for any $G \in \sigma(\vee_{j\neq i}\mathcal{P}_j^*)$, there is 0 probability that $G \in \mathcal{P}_i^*$. Then, from the lemma 1, with probability 1, each element of $\mathcal{H}$ contains an infinite number of points. Given $H \in \mathcal{H}$, $\#H = \infty$, define $\mu_k^H = \times_{\omega\in H}\mu_\omega^k$, and on this new space, all of the previous results may be applied. Henceforth, assume the reference space is H with the relevant definitions modified accordingly. Let $\mathcal{P}_i^*$ be the partition induced on H. From the previous results, the probability that $\sigma(\mathcal{P}_i^*)$ has any set in common with $\sigma(\vee_{j\neq i}\mathcal{P}_j^*)$ is 0. An implication of this is that with probability 1, for each $Q \in \sigma(\vee_{j\neq i}\mathcal{P}_j^*)$, there is some $G \in \mathcal{P}_i^*$ such that $G \cap Q \neq \emptyset$ and $G \cap Q^c \neq \emptyset$.

Since the partitions $\{\mathcal{P}_i\}_{i=1}^n$ and $\mathcal{H}$ are independently drawn, with probability 1, $\wedge_{i=1}^n\mathcal{P}_i^* = \mathcal{H}$. Suppose that $\omega^* \in H \in \mathcal{H}$. On the member of $Q \in \wedge_{j\neq i}\mathcal{P}_j^*$ containing ω^*, every G intersecting with Q has a constant value for X_i. There is some $G \in \mathcal{P}_i^*$ overlapping Q and some Q' both in $\wedge_{j\neq i}^n\mathcal{P}_j^*$. Since X_{-i} is constant on Q', the value of X_i (determined by G) must be constant on Q'. Thus, X_i is constant on $Q \cup Q'$. Proceed inductively to cover H with X_i constant. ∎

References

[1] Bergin, J. and A. Brandenburger, (1990), "A Simple Characterization of Stochastically Monotone Functions", Econometrica, Vol 58, No 5, 1241-1243.

[2] Nielsen, L.T., Brandenburger, A., Geanakoplos, J., McKelvey, R. and T. Page, (1990), "Common Knowledge of Aggregate Expectations", *Econometrica*, Vol 58, No. 5, 1235-1239.

[3] McKelvey, R.D., and T. Page, (1986), "Common Knowledge, Consensus, and Aggregate Information", *Econometrica*, Vol 54, No. 1, 109-127.