A Note on Generating Arbitrary Joint Distributions using Partitions

Kai Hao Yang*

October 27, 2020

Abstract

Consider a probability space $(\Theta, \mathcal{F}, \mathbb{P})$, two standard Borel spaces $(V, \mathcal{V})$, $(S, \mathcal{S})$, and a random variable $\mathbf{V} : \Theta \to V$. This note shows that for any probability measure $\mu \in \Delta(V \times S, \mathcal{V} \otimes \mathcal{S})$ with $\text{marg}_V \mu = \mathbb{P} \circ \mathbf{V}^{-1}$, there exists a random variable $\mathbf{S} : \Theta \to S$ such that $(\mathbf{V}, \mathbf{S})$ has law μ, provided that $(\Theta, \mathcal{F})$ is *rich relative to* $\mathbf{V}$. This result has applications in generating market segmentations using consumer characteristics; segmenting the residual demand by only partitioning the consumers according their values in a multi-firm, multi-product setting; and connects back to well known results in information economics.

Keywords: Partition, law, random variable, market segmentation, information structure

Jel classification: C18, D80, D83

*Cowles Foundation for Economic Research, Yale University, kaihao.yang@yale.edu.

1 Introduction

A random variable defined on a probability space induces a law. It is well-known that the converse is true (see, for instance, Durrett (2010), Theorem 1.2.2): Given any law, there exists a probability space and a random variable that has this law. This note explores a related question: Given a probability space and a random variable, as well as an arbitrary joint distribution with the correct marginal (i.e., the marginal distribution equals to the law of this random variable), when is it possible to find another random variable such that the two random variables have the desired joint distribution? This is a more involved question as 1) the underlying probability space is fixed, and 2) the constructed random variable must induce both a marginal and a joint distribution that agree with the targeted probability measure.

Theorem 1 and Theorem 2 show that as long as the underlying probability space is *rich* (relative to the given random variable), one can always construct another random variable such that the two random variables would have the desired joint distribution. The intuition is straightforward. Consider a probability space $(\Theta, \mathcal{F}, \mathbb{P})$ and a random variable $\mathbf{V} : \Theta \to \mathbb{R}$. For any joint distribution $\mu \in \Delta(\mathbb{R}^2)$ with $\text{marg}_{\mathbb{R}}\mu = \mathbb{P} \circ \mathbf{V}^{-1}$, since $\mathbb{R}^2$ is a regular probability space, it can be written as $\mu(\mathrm{d}v, \mathrm{d}s) = \pi(\mathrm{d}s|v)\mathbb{P} \circ \mathbf{V}^{-1}(\mathrm{d}v)$, where $\pi : \mathbb{R} \to \Delta(\mathbb{R})$ is a transitional kernel. As such, if for any $v \in \mathbb{R}$, the probability space $(\Theta, \mathcal{F}, \mathbb{P})$, when restricted on $\mathbf{V}^{-1}(v) = \{\theta \in \Theta | \mathbf{V}(\theta) = v\}$, is isomorphic to $([0, 1], \mathcal{B}([0, 1]), L)$, where L is the Lebesgue measure, then the standard method can be applied to generate a random variable with law $\pi(\cdot|v)$. The proof of Theorem 2 below formalizes this logic.

With Theorem 1 and Theorem 2, it then follows that any joint distribution with a fixed marginal can be thought of as a partition of an underlying probability space. This observation can then be applied to various economic settings, including generating arbitrary market segmentations through consumer characteristics; splitting the residual demand by simply partitioning consumers according to their values in an oligopoly setting with heterogeneous products; as well as connecting the Aumann knowledge model (Aumann, 1976) to information structures described by Harsanyi type spaces (Harsanyi, 1967-68) when there is a common prior.

The rest of this note is organized as follows. Section 2 outlines the set up and derives the main result, Section 3 discusses three applications, and Section 4 concludes.

2 Generating Joint Distribution

Consider a probability space $(\Theta, \mathcal{F}, \mathbb{P})$ and two standard Borel spaces $(V, \mathcal{V})$, $(S, \mathcal{S})$. Let $\mathbf{V} : \Theta \to V$ be a measurable function and let $\mu_0 := \mathbb{P} \circ \mathbf{V}^{-1}$ denote the law of $\mathbf{V}$. For any joint distribution $\mu \in \Delta(V \times S, \mathcal{V} \otimes \mathcal{S})$, the goal is to construct a measurable function

$\mathbf{S} : \Theta \to S$ such that

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap \mathbf{S}^{-1}(B)) = \mu(A \times B),$$

for all measurable sets $A \in \mathcal{V}$ and $B \in \mathcal{S}$.

To this end, we introduce two notions of "richness".

Definition 1. The probability space $(\Theta, \mathcal{F}, \mathbb{P})$ is *nonatomic* if for any $A \in \mathcal{F}$ such that $\mathbb{P}(A) > 0$, there exists $B \in \mathcal{F}$ such that

$$0 < \mathbb{P}(B) < \mathbb{P}(A).$$

Definition 2. The probability space $(\Theta, \mathcal{F}, \mathbb{P})$ is *rich relative to* $\mathbf{V}$ if for any $A \in \mathcal{V}$, $(\mathbf{V}^{-1}(A), \mathcal{F}|_{\mathbf{V}^{-1}(A)}, \tilde{\mathbb{P}}_A)$ is isomorphic to $(I, \mathcal{B}([0,1]), L)$ modulo zero for some interval $I \subseteq [0,1]$, where

$$\mathcal{F}|_{\mathbf{V}^{-1}(A)} := \{F \in \mathcal{F} : F \subseteq \mathbf{V}^{-1}(A)\},$$

and

$$\tilde{\mathbb{P}}_A(F) := \mathbb{P}(F \cap \mathbf{V}^{-1}(A)),$$

for any $F \in \mathcal{F}|_{\mathbf{V}^{-1}(A)}$ and L is the Lebesgue measure.

With these two notions of richness, we now introduce the main results.

Theorem 1. *Suppose that $(\Theta, \mathcal{F}, \mathbb{P})$ is nonatomic. Then for any $\mu \in \Delta(V \times S, \mathcal{V} \otimes \mathcal{S})$ such that the support of $\text{marg}_S \mu$ is countable, there exists a countable partition $\mathcal{H}$ of Θ such that $\mathcal{H} \subseteq \mathcal{F}$ and that for any $s \in \text{supp}(\text{marg}_S \mu)$, there exists $H \in \mathcal{H}$ such that*

$$\mathbb{P}(H \cap \mathbf{V}^{-1}(A)) = \mu(A \times \{s\}),$$

for all $A \in \mathcal{V}$.

To prove Theorem 1, we first introduce two useful lemmas,

Lemma 1. *There exists a countable partition $\mathcal{A}$ of V such that for any $A \in \mathcal{V}$, there exists $A' \in \sigma(\mathcal{A})$ such that*

$$\mu_0(A \Delta A') := \mu_0((A \backslash A') \cup (A' \backslash A)) = 0.$$

Lemma 1 follows directly from the fact that the symmetric difference operator Δ, together with a finite measure, induces a separable metric space (since $(V, \mathcal{V})$ is a standard Borel space), which is commonly used in probability theory and therefore the proof is omitted. Henceforth, fix the partition $\mathcal{A}$ given by Lemma 1.

Lemma 2. *For any finite measure $m \in \Delta(V, \mathcal{V})$ such that $m(A) \leq \mu_0(A)$ for all $A \in \mathcal{V}$ and for any $F \in \mathcal{F}$, if for all $A \in \mathcal{A}$,*

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap F) = m(A), \tag{1}$$

then (1) holds for any $A \in \mathcal{V}$.

Proof. Let $F \in \mathcal{F}$ be the set such that (1) holds for all $A \in \mathcal{A}$. Consider the σ-algebra generated by $\mathcal{A}$, $\sigma(\mathcal{A})$. Notice that $\mathcal{A} \cup \{\emptyset\}$ is a π-system since $\mathcal{A}$ is a partition. Furthermore, let $\mathcal{L}$ be the collection of subsets A of V such that (1) holds. It is then straightforward to verify that, by using the fact that every $A \in \mathcal{A}$ satisfies (1), $\mathcal{L}$ is a λ-system. By Dynkin's π-λ theorem, since $\mathcal{A} \cup \{\emptyset\} \subseteq \mathcal{L}$, it must be that $\sigma(\mathcal{A}) \subseteq \mathcal{L}$. That is, for any $A \in \sigma(\mathcal{A})$,

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap F) = m(A).$$

Finally, for any $A \in \mathcal{V}$, by Lemma 1, there exists $A' \in \sigma(\mathcal{A})$ such that $\mu_0(A \Delta A') = 0$, which implies that, by countable subadditivity, $\mu_0(A \backslash A') = 0$ and $m(A' \backslash A) \leq \mu_0(A' \backslash A) = 0$. As a result,

$$\begin{aligned}
\mathbb{P}(\mathbf{V}^{-1}(A) \cap F) =& \mathbb{P}([\mathbf{V}^{-1}(A') \cap F] \cup [\mathbf{V}^{-1}(A) \backslash \mathbf{V}^{-1}(A')]) \\
=& \mathbb{P}(\mathbf{V}^{-1}(A') \cap F) + \mathbb{P}(\mathbf{V}^{-1}(A) \backslash \mathbf{V}^{-1}(A')) \\
=& \mathbb{P}(\mathbf{V}^{-1}(A') \cap F) + \mathbb{P}(\mathbf{V}^{-1}(A \backslash A')) \\
=& m(A') + \mu_0(A \backslash A') \\
=& m(A \cup [A' \backslash A]) \\
=& m(A) + m(A' \backslash A) \\
=& m(A)
\end{aligned}$$

where the fourth equality holds because $A' \in \sigma(\mathcal{A})$ and $\mathbb{P} \circ \mathbf{V}^{-1} = \mu_0$ while the fifth and the last equality is due to $\mu_0(A \backslash A') = m(A' \backslash A) = 0$ as implied by Lemma 1. This completes the proof. ∎

With Lemma 1 and Lemma 2, Theorem 1 can readily be proved.

Proof of Theorem 1. Let $q := \text{marg}_S \mu$, since $\text{supp}(q)$ is a countable set, it can be enumerated as $\{s_n\}$. For each $n \in \mathbb{N}$, define $m^{s_n} \in \Delta(V, \mathcal{V})$ as

$$m^{s_n}(A) := \frac{\mu_0(A \times \{s_n\})}{\mu_0(V \times \{s_n\})}.$$

By definition, it then follows that

$$\mu_0(A) = \sum_{n=1}^{\infty} m^{s_n}(A) q(s_n).$$

Notice that for $m := m^{s_1} q(s_1)$,

$$m(A) \leq \mu_0(A)$$

for all $A \in \mathcal{V}$.

We now show that there exists $H_1 \in \mathcal{F}$ such that

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap H_1) = m(A),$$

for all $A \in \mathcal{V}$. To see this, take the countable partition $\mathcal{A}$ of V given by Lemma 1. For any $A \in \mathcal{A}$, since

$$\mathbb{P}(\mathbf{V}^{-1}(A)) = \mu_0(A) \geq m(A),$$

by Sierpinski's intermediate value theorem, there exists $H_A \subseteq \mathbf{V}^{-1}(A)$, $H_A \in \mathcal{F}$ such that

$$\mathbb{P}(H_A) = m(A).$$

Now let

$$H_1 := \bigcup_{A \in \mathcal{A}} H_A.$$

Since $\mathcal{A}$ is countable and $H_A \in \mathcal{F}$ for all $A \in \mathcal{A}$, $H_1 \in \mathcal{F}$. Moreover, since $\mathcal{A}$ is a partition of V, $\mathbf{V}^{-1}(A) \cap \mathbf{V}^{-1}(A') = \emptyset$, and hence $H_A \cap H_{A'} = \emptyset$ for all $A, A' \in \mathcal{A}$, $A \neq A'$. As a result, for any $A \in \mathcal{A}$,

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap H_1) = \mathbb{P}(\mathbf{V}^{-1}(A) \cap H_A) = \mathbb{P}(H_A) = m(A).$$

By Lemma 2, since $m(A) \leq \mu_0(A)$ for all $A \in \mathcal{V}$, it then follows that for any $A \in \mathcal{V}$,

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap H_1) = m(A).$$

Now let

$$\tilde{m} := \sum_{n=2}^{\infty} m^{s_n} q(s_n).$$

and let

$$\Theta_1 := \Theta \backslash H_1.$$

It then follows from $\mu_0 = m + \tilde{m}$ that

$$\begin{aligned}\mathbb{P}(\mathbf{V}^{-1}(A) \cap \Theta_1) =& \mathbb{P}(\mathbf{V}^{-1}(A)) - \mathbb{P}(\mathbf{V}^{-1}(A) \cap H_1) \\ =& \mu_0(A) - m(A) \\ =& \tilde{m}(A),\end{aligned}$$

for all $A \in \mathcal{V}$.

Inductively, there exists a disjoint sequence of sets $\{H_n\} \subseteq \mathcal{F}$ such that

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap H_n) = m^{s_n}(A) q(s_n),$$

for all $A \in \mathcal{V}$. In particular,

$$\begin{aligned}\mathbb{P}(\cup_{n=1}^{\infty} H_n) = \sum_{n=1}^{\infty} \mathbb{P}(H_n) &= \sum_{n=1}^{\infty} \mathbb{P}(\mathbf{V}^{-1}(V) \cap H_n) \\ &= \sum_{n=1}^{\infty} m^{s_n}(V) q(s_n) \\ &= \sum_{n=1}^{\infty} q(s_n) \\ &= 1\end{aligned}$$

and hence $\cup_{n=1}^{\infty} H_n = \Theta \backslash H_0$, for some $H_0 \in \mathcal{F}$ with $\mathbb{P}(H_0) = 0$. Together, the partition $\mathcal{H} := \{H_n\}_{n=0}^{\infty}$ is as desired. ■

Theorem 1 only allows for distributions where $\text{marg}_S \mu$ has countable support. For general distribution $\mu \in \Delta(V \times S, \mathcal{V} \otimes \mathcal{S})$, a stronger version of richness (Definition 2) is needed.

Theorem 2. *Suppose that $(\Theta, \mathcal{F}, \mathbb{P})$ is rich relative to $\mathbf{V}$. Then for any $\mu \in \Delta(V \times S, \mathcal{V} \otimes \mathcal{S})$, there exists a measurable function $\mathbf{S} : \Theta \to \mathcal{S}$ such that*

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap \mathbf{S}^{-1}(B)) = \mu(A \times B),$$

for all $A \in \mathcal{V}$ and for any $B \in \mathcal{S}$.

Proof. Consider the countable partition $\mathcal{A}$ given by Lemma 1. For any $A \in \mathcal{A}$, define a Borel measure $\nu_A \in \Delta(S, \mathcal{S})$ as

$$\nu_A(B) := \mu(A \times B), \forall B \in \mathcal{S}$$

and define a measure $\mathbb{P}_A$ on $(\mathbf{V}^{-1}(A), \mathcal{F}|_{\mathbf{V}^{-1}(A)})$ as

$$\mathbb{P}_A(F) := \mathbb{P}(F),$$

for all $F \in \mathcal{F}$ such that $F \subseteq \mathbf{V}^{-1}(A)$. Then since $(\Theta, \mathcal{F}, \mathbb{P})$ is rich relative to $\mathbf{V}$, and since $(S, \mathcal{S})$ is a standard Borel space, there exists a measurable function $\mathbf{S}_A : \mathbf{V}^{-1}(A) \to S$ such that

$$\mathbb{P}_A(\mathbf{S}_A^{-1}(B)) = \nu_A(B), \forall B \in \mathcal{S}$$

Moreover, since $\mathcal{A}$ is a partition of V, $\{\mathbf{V}^{-1}(A)\}_{A \in \mathcal{A}}$ is a partition of Θ. Now define $\mathbf{S} : \Theta \to S$ as

$$\mathbf{S}(\theta) := \mathbf{S}_A(\theta), \text{ if } \theta \in \mathbf{V}^{-1}(A).$$

Since $\mathbf{S}_A$ is measurable for all $A \in \mathcal{A}$, $\mathbf{S}$ is measurable. Moreover, notice that for any $A \in \mathcal{A}$ and for any $B \in \mathcal{S}$,

$$\begin{aligned}\mathbb{P}(\mathbf{V}^{-1}(A) \cap \mathbf{S}^{-1}(B)) =& \mathbb{P}(\mathbf{V}^{-1}(A) \cap \mathbf{S}_A^{-1}(B)) \\ =& \mathbb{P}(\mathbf{S}_A^{-1}(B)) \\ =& \nu_A(B) \\ =& \mu(A \times B)\end{aligned}$$

where the second equality follows from the fact that $\mathbf{S}_A^{-1}(V) \subseteq \mathbf{V}^{-1}(A)$. Finally, since for all $B \in \mathcal{S}$ and for all $A \in \mathcal{V}$, $\mu(A \times B) \leq \mu(A \times S) = \mu_0(A)$, Lemma 2 implies that for any $A \in \mathcal{V}$ and for any $B \in \mathcal{S}$,

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap \mathbf{S}^{-1}(B)) = \mu(A \times B),$$

as desired. ■

3 Applications

3.1 Generating Market Segmentations using Consumer Characteristics

Consider the canonical third-degree price discrimination setting. A monopolist faces a unit mass of consumers with unit demand and heterogeneous values. The distribution of consumers' values is summarized by a market demand D_0. The monopolist is able to price discriminate the consumers according to a *market segmentation* by charging different segments of consumers different prices. In these type of settings, a market segmentation is usually defined as $\sigma \in \Delta(\mathcal{D})$ such that

$$\int_{\mathcal{D}} D(p)\sigma(\mathrm{d}D) = D_0(p),\ \forall p \geq 0,$$

where $\mathcal{D}$ denotes the collection of demand functions (i.e., the collection of nondecreasing, upper-semicontinuous functions on $\mathbb{R}_+$ with $0 \leq D(p) \leq 1$ for all $p \geq 0$. In other words, instead of directly partitioning consumers into groups, a market segmentation is defined as a way to split the market demand into several "sub-demands" that average back to the market demand. Since the seminal work of Pigou (1920), this has been a textbook definition of market segmentations.

While traditionally it is not difficult to connect this definition of market segmentation to directly partitioning consumers by their characteristics (e.g., by age, location, purchasing time), a recent literature starts to explore questions that involve arbitrary market segmentations (see, for instance, Bergemann et al. (2015), Haghpanah and Siegel (2020), Yang (2020)). These papers consider *every* market segmentation, and therefore, making it more economically relevant to understand the conditions of the consumers' characteristic space under which *any* market segmentation can be generated by some partition of the consumer characteristics. The results above serve exactly this purpose.

To see this, notice that for any $D \in \mathcal{D}$, there is a (uniquely) associated probability measure $m^D \in \Delta(\mathbb{R}_+)$. As a result, a market segmentation σ can be thought of as a joint distribution on $\mathbb{R}_+ \times \mathcal{D}$ such that the marginal on $\mathbb{R}_+$ equals to m^{D_0}. That is, by letting

$$\mu(A \times B) := \int_B m^D(A)\sigma(\mathrm{d}D)$$

for any measurable sets $A \subseteq \mathbb{R}_+$ and $B \subseteq \mathcal{D}$, a market segmentation σ can be thought of as a joint distribution with a certain marginal on $\mathbb{R}_+$. Meanwhile, the probability space $(\Theta, \mathcal{F}, \mathbb{P})$ can be thought of as the consumers' characteristic space, so that each consumer is identified with a characteristic $\theta \in \Theta$, and a consumer with characteristic θ has value $\mathbf{V}(\theta)$. Together with Theorem 1 and Theorem 2, we have following corollaries.

Corollary 1. *Suppose that the characteristic space* $(\Theta, \mathcal{F}, \mathbb{P})$ *is nonatomic. Then for any market segmentation* σ *with* $\text{supp}(\sigma)$ *being countable, there exists a countable partition* $\mathcal{H}$ *of* Θ *such that* $\mathcal{H} \subseteq \mathcal{F}$ *and that for any* $D \in \text{supp}(s)$*, there exists* $H \in \mathcal{H}$ *such that*

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap H) = m^D(A)s(D),$$

for all measurable set $A \subseteq \mathbb{R}_+$*.*

In other words, given this partition $\mathcal{H}$, for any demand function $D \in \text{supp}(\sigma)$, there exists $H \in \mathcal{H}$ such that among the consumers with $\theta \in H$, their demand is given by D. Similar conclusion can also be drawn even when σ is uncountable.

Corollary 2. *Suppose that* $(\Theta, \mathcal{F}, \mathbb{P})$ *is rich relative to* $\mathbf{V}$*. Then for any market segmentation* σ*, there exists a* σ*-algebra* $\mathcal{H} \subseteq \mathcal{F}$ *such that for any measurable* $B \subseteq \mathcal{D}$*, there exists* $H \in \mathcal{H}$ *such that*

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap H) = \int_B m^D(A)\sigma(\mathrm{d}D),$$

for all measurable set $A \subseteq \mathbb{R}_+$*.*

For example, $(\Theta, \mathcal{F}, \mathbb{P})$ is rich relative to $\mathbf{V}$ if $\Theta = \Theta_1 \times \Theta_2 \times \dots \Theta_K$ with $\Theta_k \subset \mathbb{R}$ for all k (e,g., θ contains K different consumer characteristics that are described numerically , such as income, tax record, age, time spent on a certain website); $\mathbb{P}$ is absolutely continuous with respect to the Lebesgue measure; and if $\mathbf{V}^{-1}(v)$ has Hausdorff dimension greater than 2 for all $v \in \mathbf{V}(\Theta)$.

3.2 Splitting the Residual Demand by Partitioning Consumers' Values

Consider the following generalization of the environment described above. There are $n \geq 2$ firms, each firm produces a different product. There is a unit mass consumers with unit demand and heterogeneous values. Across the consumers, their values $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n_+$ are distributed according to a probability measure μ_0. A market segmentation, analogous to what is defined above, is a probability measure $\sigma \in \Delta(\mathbb{R}^n_+)$ such that

$$\int_{\Delta(\mathbb{R}^n_+)} m(A)\sigma(\mathrm{d}m) = \mu_0(A),$$

for all measurable set $A \subseteq \mathbb{R}^n_+$. The firms compete on the price margin, and the consumers choose to buy from the firm for which $v_i - p_i$ is the largest.

Suppose that each firm can further price discriminate according to (firm-specific) market segmentations. A crucial object for understanding the firms' pricing incentives is the *residual demand* (see, for instance, (Li, 2020)). That is, for firm i, given other firms' prices p_{-i}, the share of consumers that will buy from firm i if firm i sets the price at p_i. Specifically,

$$D_i(p_i) := \{\mathbf{v} \in V | v_i - p_i \geq \max_{j \neq i}(v_j - p_j)\}.$$

Thus, for firm i' pricing incentives, it would be convenient to work with $\hat{\sigma} \in \Delta(\mathcal{D})$ such that

$$\int_{\mathcal{D}} D(p_i)\hat{\sigma}(\mathrm{d}D) = D_i(p_i) \tag{2}$$

for all $p_i \geq 0$. Furthermore, it would be even more convenient if it is without loss to restrict attention to partitions of $\mathrm{supp}(\mu_0)$ so that any $\hat{\sigma} \in \Delta(\mathcal{D})$ satisfying (2) can be induced by some partition of $\mathrm{supp}(\mu_0)$. By Theorem 2, this is possible if μ_0 is absolutely continuous with respect to the Lebesgue measure and has full support on a connected set.

Corollary 3. *Given any $p_{-i} \in \mathbb{R}^{n-1}_+$, suppose that μ_0 is absolutely continuous and has full support on a connected set. Then for any $\hat{\sigma} \in \Delta(\mathcal{D})$ such that*

$$\int_{\mathcal{D}} D(p_i)\hat{\sigma}(\mathrm{d}D) = D_i(p_i),$$

there exists a σ-algebra $\mathcal{H} \subseteq \mathcal{B}(\mathbb{R}^n_+)$ so that for any measurable $B \subseteq \mathcal{D}$, there exists $H \in \mathcal{H}$ such that

$$\mu_0(\{\mathbf{v} | v_i - p_i \geq \max_{j \neq i}\{v_j - p_j\}\} \cap H) = \int_B D(p_i)\sigma(\mathrm{d}D),$$

for all $p_i \geq 0$

3.3 Aumann Meets Harsanyi

In seminal papers such as Aumann (1976) and Aumann (1987), information is modeled as partitions of the state space. While this type of partitional models has provided countless profound insights, sometimes it is more convenient to model information through joint distributions (see, for example, Bergemann and Morris (2016)). The results above formalize a straightforward connection between these two approaches when there is a common prior.

More specifically, suppose that there are $n \geq 1$ agents with a common prior. In Aumann's model, information is defined as partitions of an underlying state of the world $(\Theta, \mathcal{F}, \mathbb{P})$. That is, the information player i has is given by a σ-algebra $\mathcal{H}_i$ on Θ. Meanwhile, another way to define an *information structure* is through Harsanyi's approach (Harsanyi, 1967-68). That is, suppose that $(V, \mathcal{V})$ is the payoff-relevant state space. An information structure (when there is a common prior) is modeled by a joint distribution $\mu \in \Delta(V \times S, \mathcal{V} \otimes \mathcal{S})$, where $S = S_1 \times \ldots \times S_n$ and $\mathcal{S} = \mathcal{S}_1 \otimes \ldots \otimes \mathcal{S}_n$ and $(S_i, \mathcal{S}_i)$ stands for the signal space for agent i.[1]

Suppose that $\mathbf{V} : \Theta \to V$ is a measurable function. Then any informational partitions $\{\mathcal{H}_i\}_{i=1}^N$ define an information structure $\mu \in \Delta(V \times S, \mathcal{V} \otimes \mathcal{S})$.[2] The results above ensure that the converse is true, as long as $(\Theta, \mathcal{F}, \mathbb{P})$ is rich relative to $\mathbf{V}$.

Corollary 4. *Suppose that $(\Theta, \mathcal{F}, \mathbb{P})$ is rich relative to $\mathbf{V}$. Then for any information structure $\mu \in \Delta(V \times S, \mathcal{V} \otimes \mathcal{S})$ with $\text{marg}_V \mu = \mathbb{P} \circ \mathbf{V}^{-1}$, there exists a collection of σ-algebras $\{\mathcal{H}_i\}_{i=1}^n$ such that $\mathcal{H}_i \subseteq \mathcal{F}$ for all i and that for any measurable sets $B_1 \subseteq S_1, \ldots, B_n \subseteq S_n$, there exists $H_1 \in \mathcal{H}_1, \ldots, H_n \in \mathcal{H}_n$ so that*

$$\mathbb{P}(\mathbf{V}^{-1}(A) \cap H_1 \cap \ldots \cap H_n) = \mu(A \times B_1 \times \ldots \times B_n), \forall A \in \mathcal{V}$$

As a result, for any information structure μ, by constructing a large enough underlying state space $(\Theta, \mathcal{F}, \mathbb{P})$, one can always represent this information structure using an information partition. In fact, this method is commonly used in the information design literature as it gives a convenient way to define the information generated by multiple Blackwell experiments and to capture arbitrary correlations among multiple signals (see, for instance, Green and Stokey (1978), Gentzkow and Kamenica (2017), Frankel and Kamenica (2019), Brooks et al. (2020)).[3]

4 Conclusion

This note shows that it is always possible to generate a joint distribution through a partition of an underlying probability space, as long as this underlying space is rich enough. Several

[1]For technical reasons, assume that $(V, \mathcal{V})$ and $\{(S_i, \mathcal{S}_i)\}_{i=1}^n$ are standard Borel spaces.

[2]For each i, let $\mathbf{S}_i := \mathbb{E}[\mathbf{V}|\mathcal{H}_i]$. Then the law of $(\mathbf{V}, \mathbf{S}_1, \ldots, \mathbf{S}_n)$ is a probability measure on $V \times S$.

[3]The typical way to construct the underlying state space is to define $\Theta := V \times [0, 1]$, $\mathcal{F} = \mathcal{V} \otimes \mathcal{B}([0, 1])$, $\mathbb{P} = \mu_0 \otimes L$ and $\mathbf{V}(v, x) := v$ for all $(v, x) \in \Theta$. By construction, $(\Theta, \mathcal{F}, \mathbb{P})$ is clearly rich relative to $\mathbf{V}$.

applications are discussed, including generating market segmentations by partitioning consumers' characteristics; splitting a residual demand by partitioning consumers according to their value vectors; and representing an information structure by partitioning an underlying state space. While richness of the underlying probability space is only a sufficient condition for these results, characterizing the necessary and sufficient condition, as well as exploring the connection between joint distributions and partitions with limited underlying probability space, remain open questions.

References

AUMANN, R. J. (1976): "Agreeing to Disagree," *Annals of Statistics*, 4, 1236–1239.

——— (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

BERGEMANN, D., B. BROOKS, AND S. MORRIS (2015): "The Limits of Price Discrimination," *American Economic Review*, 105, 921–957.

BERGEMANN, D. AND S. MORRIS (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487–522.

BROOKS, B., A. FRANKEL, AND E. KAMENICA (2020): "Information Hierarchies," Working Paper.

DURRETT, R. (2010): *Probability: Theory and Examples (4th-ed)*, New York, NY: Cambridge University Press.

FRANKEL, A. AND E. KAMENICA (2019): "Quantifying Information and Uncertainty," *American Economic Review*, 109, 3650–3680.

GENTZKOW, M. AND E. KAMENICA (2017): "Bayesian Persuasion with Multiple Senders and Rich Signal Spaces," *Games and Economic Behavior*, 104, 411–429.

GREEN, J. R. AND N. STOKEY (1978): "Two Representations of Information Structures and their Comparisons," Working Paper.

HAGHPANAH, N. AND R. SIEGEL (2020): "Pareto Improving Segmentation of Multi-product Markets," Working Paper.

HARSANYI, J. C. (1967-68): "Games of Incomplete Information Played by Bayesian Players, Part I, II, III," *Management Science*, 14, 159–182, 320–334, 486–502.

LI, W. (2020): "Using Information to Amplify Competition," Working Paper.

Pigou, A. C. (1920): *The Economics of Welfare*, London: Macmillan.

Yang, K. H. (2020): "Selling Consumer Data for Profit: Optimal Market-Segmentation Design and its Consequences," Discussion Paper 2258, Cowles Foundation for Research in Economics, Yale University.