The Essential Equivalence of Pairwise and Mutual Conditional Independence*

Peter J. Hammond† and Yeneng Sun‡

Abstract

For a large collection of random variables, pairwise conditional independence and mutual conditional independence are shown to be essentially equivalent. Unlike in the finite setting, a large collection of random variables remains essentially conditionally independent under further conditioning. The essential equivalence of pairwise and multiple versions of exchangeability also follows as a corollary. Our proof is based on an iterative extension of Bledsoe and Morse's completion of a product measure on a pair of measure spaces.

*Part of this work was done while Peter Hammond was visiting the National University of Singapore in March–April 2004. This version was completed while Yeneng Sun was visiting the University of Illinois at Urbana–Champaign in October 2004 – February 2005.

†Department of Economics, Stanford University, Stanford, CA 94305. e-mail: peter.hammond@stanford.edu

‡Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543; and Department of Economics, National University of Singapore, 1 Arts Link, Singapore 117570. e-mail: matsuny@nus.edu.sg

1 Introduction

Conditional independence is a fundamental concept in probability theory. For example, a Markov process can be defined as a stochastic process in which the past and future are conditionally independent given the present. A classical example due to Bernstein shows that there are three events which are pairwise independent but not mutually independent (see [7, p. 126]). Since independence is a trivial case of conditional independence, this also means that pairwise conditional independence and its multivariate analog is not equivalent for a finite collection of random variables. Despite the fact that independence is a trivial version of conditional independence, very simple examples can be constructed to show that two independent random variables may lose their independence under conditioning (see [3, p. 217]).

The main aim of this paper is to show that if a large collection of random variables is considered, then pairwise conditional independence is essentially equivalent to its multivariate analog. It will also be shown that such a collection of random variables will remain essentially conditionally independent under further conditioning. In particular, a collection of independent random variables will not lose its independence under conditioning (unlike in the finite setting). In addition, the essential equivalence of pairwise and multiple versions of exchangeability follows as a corollary. We prove these results by generalizing Bledsoe and Morse's completion of the product of two measures in [2].

In this paper, a large collection of random variables is formalized as a process indexed by points in an atomless probability space, which is simply called a continuum of random variables. As discussed in [13] and some of the work cited there, such processes occur in many economic models, especially those with essentially independent random variables. In particular, a large literature in macroeconomics has relied on a version of the exact law of large numbers for a continuum of independent random variables/stochastic processes — see [14]. Note that an atomless probability index space provides a convenient idealization for economic models with a large but finite number of agents. From a technical point of view, such an idealization is often necessary for developing the relevant models of important economic phenomena such as competitive markets.

Finally, our exact equivalence results in this paper also correspond to some asymptotic results for a triangular array of random variables. Specifically, following Section 5 of [12] and Section 9 of [13], the routine procedure of lifting, pushing down, and transfer can be applied to processes on a special Loeb product probability space in order to demonstrate an appropriate form of asymptotic equivalence between pairwise and mutual conditional independence for such an array.

The paper is organized as follows. Section 2 generalizes an idea of Bledsoe and Morse on extending the product of two measures to the setting of finite or infinite products of measure spaces. This is done by adding the iterated null sets to the relevant product σ-algebras. Section 3 presents the main results with some discussion of the literature. The proofs are given in Section 4.

2 Extending the finite or infinite products of measure spaces

Let $(T_k, \mathcal{T}_k, \lambda_k)$, $k \in \mathbb{N}$ be a sequence of complete and countably additive probability spaces. Then $(\prod_{k=1}^n T_k, \otimes_{k=1}^n \mathcal{T}_k, \otimes_{k=1}^n \lambda_k)$ is the product of the first n probability spaces, whereas $(\prod_{k=1}^\infty T_k, \otimes_{k=1}^\infty \mathcal{T}_k, \otimes_{k=1}^\infty \lambda_k)$ is the infinite product of the entire sequence of probability spaces.

Though we can always assume that the above product probability spaces are complete in the sense that subsets of measure zero are included as measurable sets with zero measure, this completion is not enough for us to derive our main result — namely, the essential equivalence of pairwise and mutual conditional independence (see Remark 2 below). A stronger form of "iterative" completion will be used for the products $(\prod_{k=1}^n T_k, \otimes_{k=1}^n \mathcal{T}_k, \otimes_{k=1}^n \lambda_k)$ and $(\prod_{k=1}^\infty T_k, \otimes_{k=1}^\infty \mathcal{T}_k, \otimes_{k=1}^\infty \lambda_k)$, involving those "iteratively null" sets whose indicator functions have value zero for iterated integrals of all orders. The following definition extends what Bledsoe and Morse [2] suggested for the case of two measure spaces (see also [4, p. 108]).

Definition 1 *A set $E \subseteq \prod_{k=1}^n T_k$ is said to be **iteratively null** if for every permutation π on $\{1, \ldots, n\}$, the iterated integral*

$$\int_{t_{\pi(1)} \in T_{\pi(1)}} \cdots \int_{t_{\pi(n)} \in T_{\pi(n)}} 1_E \, d\lambda_{\pi(n)}(t_{\pi(n)}) \ldots d\lambda_{\pi(1)}(t_{\pi(1)}) \tag{1}$$

is well-defined with value zero, where 1_E is the indicator function of the set E in $\prod_{k=1}^n T_k$; in other words, for $\lambda_{\pi(1)}$-a.e. $t_{\pi(1)} \in T_{\pi(1)}$, $\lambda_{\pi(2)}$-a.e. $t_{\pi(2)} \in T_{\pi(2)}$, $\ldots$, $\lambda_{\pi(n)}$-a.e. $t_{\pi(n)} \in T_{\pi(n)}$, one has $(t_1, t_2, \ldots, t_n) \notin E$.

As mentioned in [4, p. 113], Sierpiński constructed a subset A of $[0,1]^2$ whose ρ^2-outer measure is one although its intersection with every line consists of at most two points. It is obvious that the set A is iteratively null, implying that its ρ^2-inner measure is zero. Thus, A is not in the usual product Lebesgue σ-algebra. For the special class of atomless Loeb probability spaces constructed in [11], it is shown in [1] that there is a continuum of increasing Loeb product null sets with large gaps, in the sense that their set differences have outer measure one under the usual product of two atomless Loeb probability spaces. Since a Loeb product is an extension of the usual product with the Fubini property, Loeb product null sets must be iteratively null. Hence, there is a large class of iteratively null sets that are not measurable with respect to the completion of the usual product σ-algebra.

The following two propositions show that one can extend both the finite and the infinite product probability spaces $(\prod_{k=1}^n T_k, \otimes_{k=1}^n \mathcal{T}_k, \otimes_{k=1}^n \lambda_k)$ and $(\prod_{k=1}^\infty T_k, \otimes_{k=1}^\infty \mathcal{T}_k, \otimes_{k=1}^\infty \lambda_k)$ respectively by including all the iteratively null sets, and then forming the iterated completion.

Proposition 1 *Given any $n \in \mathbb{N}$, let $\mathcal{E}_n$ denote the family of all iteratively null sets in $\prod_{k=1}^n T_k$. Then there exists a complete and countably additive probability space $(\prod_{k=1}^n T_k, \bar{\otimes}_{k=1}^n \mathcal{T}_k, \bar{\otimes}_{k=1}^n \lambda_k)$ with*

$$\begin{aligned} \bar{\otimes}_{k=1}^n \mathcal{T}_k &:= \sigma([\otimes_{k=1}^n \mathcal{T}_k] \cup \mathcal{E}_n) \\ = [\otimes_{k=1}^n \mathcal{T}_k]\,\Delta \mathcal{E}_n &:= \{\, B \Delta E : B \in \otimes_{k=1}^n \mathcal{T}_k,\ E \in \mathcal{E}_n \,\} \\ \text{and} \quad [\bar{\otimes}_{k=1}^n \lambda_k]\,(B \Delta E) &= [\otimes_{k=1}^n \lambda_k]\,(B) \text{ whenever } B \in \otimes_{k=1}^n \mathcal{T}_k,\ E \in \mathcal{E}_n. \end{aligned}$$

Proposition 2 *There exists a countably additive probability space, denoted by $(\prod_{k=1}^\infty T_k, \bar{\otimes}_{k=1}^\infty \mathcal{T}_k, \bar{\otimes}_{k=1}^\infty \lambda_k)$, in which $\bar{\otimes}_{k=1}^\infty \mathcal{T}_k$ is the σ-algebra generated by the union $\mathcal{G} := \cup_{n=1}^\infty \mathcal{G}_n$ of the families $\mathcal{G}_n$ of cylinder sets taking the form $A \times \prod_{k=n+1}^\infty T_k$ for some $A \in \bar{\otimes}_{k=1}^n \mathcal{T}_k$, whereas $\bar{\otimes}_{k=1}^\infty \lambda_k$ is the unique countably additive extension to this σ-algebra of the set function $\mu : \mathcal{G} \to [0,1]$ defined so that $\mu(A \times \prod_{k=n+1}^\infty T_k) := \bar{\otimes}_{k=1}^n \lambda_k(A)$ for all $A \in \bar{\otimes}_{k=1}^n \mathcal{T}_k$.*

Unlike in the finite product setting, the infinite product measure space $(\prod_{k=1}^\infty T_k, \bar{\otimes}_{k=1}^\infty \mathcal{T}_k, \bar{\otimes}_{k=1}^\infty \lambda_k)$ in Proposition 2 above may not be complete in the usual sense. One can always complete it by the usual procedure (see, for example, [4] pp. 78–79). We still use the same notation to denote that completion.

The countably additive probability spaces $(\prod_{k=1}^n T_k, \bar{\otimes}_{k=1}^n \mathcal{T}_k, \bar{\otimes}_{k=1}^n \lambda_k)$ and $(\prod_{k=1}^\infty T_k, \bar{\otimes}_{k=1}^\infty \mathcal{T}_k, \bar{\otimes}_{k=1}^\infty \lambda_k)$ will be called the *iterated completions* of $(\prod_{k=1}^n T_k, \otimes_{k=1}^n \mathcal{T}_k, \otimes_{k=1}^n \lambda_k)$ and of $(\prod_{k=1}^\infty T_k, \otimes_{k=1}^\infty \mathcal{T}_k, \otimes_{k=1}^\infty \lambda_k)$, respectively. Since the measure $\bar{\otimes}_{k=1}^n \lambda_k$ only differs from the usual product measure $\otimes_{k=1}^n \lambda_k$ up to a set whose indicator functions has value zero for an iterated integral of any order, it is obvious that these iterated completions for the finite products still have the usual Fubini property.

When all the probability spaces $(T_k, \mathcal{T}_k, \lambda_k)$ $(k \in \mathbb{N})$ are copies of $(T, \mathcal{T}, \lambda)$, let $(T^n, \bar{\mathcal{T}}^n, \bar{\lambda}^n)$ and $(T^\infty, \bar{\mathcal{T}}^\infty, \bar{\lambda}^\infty)$ respectively denote the iterated completions of the n-fold and infinite product probability spaces.

3 Main results

The following definition introduces some basic concepts and notation.

Definition 2 *Let $(\Omega, \mathcal{A}, P)$ be a complete, countably additive probability space. Let $\mathcal{C}$ be a sub-σ-algebra of the σ-algebra $\mathcal{A}$, and X a complete separable metric space with the Borel σ-algebra $\mathcal{B}$. Let $(T, \mathcal{T}, \lambda)$ be a complete atomless probability space. A process g is a mapping from $T \times \Omega$ to X such that for all $t \in T$, the mapping $g_t(\cdot) = g(t, \cdot)$ is $\mathcal{A}$-measurable (i.e., g_t is a random variable).*

1. *Let $n \in \mathbb{N}$ with $n \geq 2$. A finite sequence $\{f_k\}_{k=1}^n$ of X-valued random variables is said to be* mutually conditionally independent *given $\mathcal{C}$ if, for any Borel sets $B_k \in \mathcal{B}$, $k = 1, \ldots, n$, the conditional probabilities satisfy*
$$P(\cap_{k=1}^n f_k^{-1}(B_k)|\mathcal{C}) = \prod_{k=1}^n P(f_k^{-1}(B_k)|\mathcal{C}). \quad (2)$$
When $n = 2$ the two random variables f_1, f_2 satisfing Equation (2) are said to be pairwise conditionally independent *given $\mathcal{C}$.*

2. *A sequence $\{f_k\}_{k=1}^\infty$ of X-valued random variables is said to be* mutually conditionally independent *given $\mathcal{C}$ if, for every $n \geq 2$, the finite sequence $\{f_k\}_{k=1}^n$ of X-valued random variables is mutually conditionally independent given $\mathcal{C}$.*

3. *Let $n \in \mathbb{N}$ with $n \geq 2$. The process g is said to be* essentially n-wise mutually conditionally independent *given $\mathcal{C}$ if, for $\bar{\lambda}^n$-a.e. $(t_1, \ldots, t_n) \in T^n$, the random variables $g_{t_1}, \ldots g_{t_n}$ are mutually conditionally independent given $\mathcal{C}$. When $n = 2$, g is said to be* essentially pairwise conditionally independent *given $\mathcal{C}$.*

4. *The process g is said to be* essentially mutually conditionally independent *given $\mathcal{C}$ if, for $\bar{\lambda}^\infty$-a.e. $\{t_k\}_{k=1}^\infty \in T^\infty$, the random variables $\{g_{t_k}\}_{k=1}^\infty \in T^\infty$ are mutually conditionally independent given $\mathcal{C}$.*

5. *Given the process $g : T \times \Omega \to X$ and the σ-algebra $\mathcal{C} \subseteq \mathcal{A}$, the $\mathcal{T} \otimes \mathcal{C}$-measurable mapping $\mu : T \times \Omega \to \mathcal{M}(X)$ is said to be an* essentially regular conditional distribution process *if for λ-a.e. $t \in T$ the $\mathcal{C}$-measurable mapping $\omega \mapsto \mu_{t\omega}$ is a regular conditional distribution $P(g_t^{-1}|\mathcal{C})$ of the random variable g_t.*

6. *The σ-algebra $\mathcal{C}$ is said to be* countably generated *if there is a countable family $\mathcal{F} \subseteq \mathcal{C}$ that generates $\mathcal{C}$.*

Let $\mathcal{M}(X)$ be the space of Borel probability measures on X endowed with the topology of weak convergence of measures. It is easy to see that the measurability of a mapping ϕ from a measurable space $(I, \mathcal{I})$ to $\mathcal{M}(X)$ with the Borel σ-algebra generated by the topology of weak convergence of measures is equivalent to the $\mathcal{I}$-measurability of all the mappings $\phi(\cdot)(B)$ for any $B \in \mathcal{B}$ (see, for example, [8, p. 748]). The following theorem shows that essential pairwise conditional independence is equivalent to its finite or infinite multivariate versions.

Theorem 1 *Suppose that the process g from $T \times \Omega$ to X and the σ-algebra $\mathcal{C} \subseteq \mathcal{A}$ admit an essentially regular conditional distribution process μ. Provided that $\mathcal{C}$ is countably generated, the following are equivalent:*

1. *The process g is essentially pairwise conditionally independent given $\mathcal{C}$.*

2. *For each fixed $n \in \mathbb{N}$ with $n \geq 2$, the process g is essentially n-wise mutually conditionally independent given $\mathcal{C}$.*

3. *The process g is essentially mutually conditionally independent given $\mathcal{C}$.*

Because (unconditional) independence is a special case of conditional independence given the trivial σ-algebra, the next result is an obvious implication of Theorem 1.

Corollary 1 *Let g be a process from $T \times \Omega$ to X such that the distribution mapping Pg_t^{-1} from T to $\mathcal{M}(X)$ is $\mathcal{T}$-measurable. Then the following are equivalent:*

1. *For $\bar{\lambda}^2$-a.e. $(t_1, t_2) \in T^2$, the random variables g_{t_1} and g_{t_2} are independent.*

2. *For $\bar{\lambda}^n$-a.e. $(t_1, \ldots, t_n) \in T^n$, the random variables $g_{t_1}, \cdots, g_{t_n}$ are mutually independent.*

3. *For $\bar{\lambda}^\infty$-a.e. $\{t_k\}_{k=1}^\infty \in T^\infty$, the sequence $\{g_{t_k}\}_{k=1}^\infty$ of random variables is mutually independent.*

Remark 1 *Corollary 1 has been shown for processes on Loeb product spaces in Theorem 3 and Proposition 3.4 of [12]. When g is essentially pairwise independent, Proposition 1.1 in [12] shows that g is not measurable with respect to the usual product σ-algebra $\mathcal{T} \otimes \mathcal{A}$ except in some trivial cases; since the iterated completion of $\mathcal{T} \otimes \mathcal{A}$ only involves the addition of some null sets, in general g is not measurable even with respect to the iterated completion. Note that the framework of Loeb product spaces naturally allows the Fubini property for functions of several variables that are not necessarily measurable with respect to the usual product σ-algebras (or even their iterated completions). These generalized Fubini properties are used extensively in the proofs of Theorem 3 and Proposition 3.4 of [12]. Corollary 1 generalizes the corresponding result to the general setting without any assumption on generalized Fubini properties related to the process g. On the other hand, in order to prove an exact law of large numbers in a meaningful analytic framework, as in [14], one needs a suitable extension of the usual product σ-algebras that retains the Fubini property.*

The following proposition shows that essential pairwise conditional independence is preserved under further conditioning (i.e. when the underlying countably generated σ-algebra is enlarged).

Proposition 3 *Suppose that the process g from $T \times \Omega$ to X and the countably generated σ-algebra $\mathcal{C} \subseteq \mathcal{A}$ admit an essentially regular conditional distribution process μ. For any given countably generated sub-σ-algebra $\mathcal{C}'$ of $\mathcal{A}$ that contains $\mathcal{C}$, if the process g is essentially pairwise conditionally independent given $\mathcal{C}$, then it is essentially pairwise conditionally independent given $\mathcal{C}'$.*

As noted in the introduction, two independent random variables may lose their independence under conditioning (see [3, p. 217]). However, the following result, which is an obvious corollary of Proposition 3, shows that a large collection of essentially independent random variables will not lose their essential independence under conditioning.

Corollary 2 *Let g be a process from $T \times \Omega$ to X such that the distribution mapping Pg_t^{-1} from T to $\mathcal{M}(X)$ is $\mathcal{T}$-measurable. Suppose that for $\bar{\lambda}^2$-a.e. $(t_1, t_2) \in T^2$, the random variables g_{t_1} and g_{t_2} are independent. Then, given any countably generated sub-σ-algebra $\mathcal{C}$ of $\mathcal{A}$, the process g is essentially pairwise conditionally independent given $\mathcal{C}$.*

Exchangeability is another fundamental concept in probability theory with many applications; see, for example, the book [3] and the survey papers [9] and [10]. Some work such as [6] also uses the weaker concept of pairwise exchangeability. Theorem 4 and Proposition 3.5 of [12] show that essential pairwise exchangeability is equivalent to its finite or infinite multivariate versions for processes on Loeb product spaces. The following corollary extends these results to the general setting without assuming that the process g satisfies any generalized Fubini property. The proof here is based on Theorem 1 and on a de Finetti type result that essential pairwise exchangeability is equivalent to essential pairwise conditional independence with identical distributions.

Corollary 3 *Let g be a process from $T \times \Omega$ to X. Then the following are equivalent.*

1. *The random variables g_t are essentially pairwise exchangeable — i.e., there is a symmetric distribution ν_2 on X^2 such that for $\bar{\lambda}^2$-almost all $(t_1, t_2) \in T^2$, g_{t_1} and g_{t_2} have a joint distribution ν_2.*

2. *For $n \geq 2$, the random variables g_t are essentially n-wise exchangeable — i.e., there is a symmetric distribution ν_n on X^n such that the joint distribution of $g_{t_1}, \ldots, g_{t_n}$ is ν_n for $\bar{\lambda}^n$-almost all $(t_1, \ldots, t_n) \in T^n$.*

3. *For $\bar{\lambda}^\infty$-a.e. $\{t_k\}_{k=1}^\infty \in T^\infty$, $\{g_{t_k}\}_{k=1}^\infty$ is an exchangeable sequence of random variables.*

4 The Proofs

4.1 Proof of the results in Section 2

Proof of Proposition 1: It is easy to see that $\mathcal{E}_n$ is a *hereditary σ-ring* in the sense that $D \in \mathcal{E}_n$ whenever $D \subseteq E \in \mathcal{E}_n$, whereas $\cup_{m=1}^\infty D_m \in \mathcal{E}_n$ for any sequence D_m ($m \in \mathbb{N}$) of sets in $\mathcal{E}_n$. The usual Fubini theorem also implies that if $E \in \mathcal{E}_n \cap \otimes_{k=1}^n \mathcal{T}_k$, then $\otimes_{k=1}^n \lambda_k(E) = 0$. By the usual argument on completing a measure (see, for example, [4] pp. 78–79), the σ-algebra generated by $[\otimes_{k=1}^n \mathcal{T}_k] \cup \mathcal{E}_n$, which will be denoted by $\bar{\otimes}_{k=1}^n \mathcal{T}_k$, is the collection of all sets A of the form $A = B \Delta E$ for $B \in \otimes_{k=1}^n \mathcal{T}_k$ and $E \in \mathcal{E}_n$.

Now extend the product measure $\otimes_{k=1}^n \lambda_k$ to a measure $\bar{\otimes}_{k=1}^n \lambda_k$ on $\bar{\otimes}_{k=1}^n \mathcal{T}_k$ by letting $[\bar{\otimes}_{k=1}^n \lambda_k](A) = \otimes_{k=1}^n \lambda_k(B)$, which is a well-defined countably additive measure on $\bar{\otimes}_{k=1}^n \mathcal{T}_k$.

If $[\bar{\otimes}_{k=1}^n \lambda_k](A) = 0$, then $\otimes_{k=1}^n \lambda_k(B) = 0$. By the usual Fubini theorem, B is also an iteratively null set, which means that A is also iteratively null. For any subset D of A, since all the measures $\lambda_1, \ldots, \lambda_n$ are complete, D is also iteratively null and thus belongs to $\bar{\otimes}_{k=1}^n \mathcal{T}_k$. Hence $(\prod_{k=1}^n T_k, \bar{\otimes}_{k=1}^n \mathcal{T}_k, \bar{\otimes}_{k=1}^n \lambda_k)$ is a complete measure space. ∎

Proof of Proposition 2: It is obvious that μ is a finitely additive measure on the algebra $\mathcal{G}$. The usual proof that an infinite-dimensional product measure exists (see [3], p. 193) can be used here to show that μ is actually countably additive. Note that this proof uses the Fubini property. But any iterative completion of a finite product probability space also has the Fubini property, so this does not create any problem. Therefore μ can be extended to a countably additive measure on $\sigma(\mathcal{G})$ by the Carathéodory Extension Theorem. ∎

4.2 Proof of the results in Section 3

Lemma 1 *Let g be a process from $T \times \Omega$ to X. Let $\mathcal{C} \subseteq \mathcal{A}$ be a countably generated σ-algebra on Ω and μ a $\mathcal{T} \otimes \mathcal{C}$-measurable mapping from $T \times \Omega$ to $\mathcal{M}(X)$. Suppose that for $\bar{\lambda}^2$-a.e. $(t_1, t_2) \in T^2$, one has*

$$P(g_{t_1}^{-1}(B_1) \cap g_{t_2}^{-1}(B_2) | \mathcal{C}) = \mu_{t_1\omega}(B_1)\mu_{t_2\omega}(B_2) \text{ whenever } B_1, B_2 \in \mathcal{B}. \quad (3)$$

Then for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$ one has $P(A \cap g_t^{-1}(B)) = \int_A \mu_{t\omega}(B) dP$ for λ-a.e. $t \in T$.

Proof: Consider the case when B_2 in Equation (3) is the whole space Ω. Then, for $\bar{\lambda}^2$-a.e. $(t', t) \in T^2$, one has

$$P(g_{t'}^{-1}(B)|\mathcal{C}) = P(g_{t'}^{-1}(B) \cap g_t^{-1}(\Omega)|\mathcal{C}) = \mu_{t'\omega}(B)\mu_{t\omega}(\Omega) = \mu_{t'\omega}(B)$$

for all $B \in \mathcal{B}$, which implies that $\mu_{t'\omega}$ is a version of the regular conditional distribution $P(g_{t'}^{-1}(\cdot)|\mathcal{C})$ of $g_{t'}$ conditioned on $\mathcal{C}$.

Now fix $B \in \mathcal{B}$. By hypothesis, there exists a fixed set T_1 with $\lambda(T_1) = 1$ such that, for each $t' \in T_1$, one has

$$P(g_{t'}^{-1}(B) \cap g_t^{-1}(B)|\mathcal{C}) = \mathbb{E}(1_{g_{t'}^{-1}(B)} 1_{g_t^{-1}(B)}|\mathcal{C}) = \mu_{t'\omega}(B)\mu_{t\omega}(B) \quad (4)$$

for λ-a.e. $t \in T$, and also

$$P(g_{t'}^{-1}(B)|\mathcal{C}) = \mathbb{E}(1_{g_{t'}^{-1}(B)}|\mathcal{C}) = \mu_{t'\omega}(B). \quad (5)$$

Equations (4) and (5) imply in particular that for any $t', t \in T_1$, one has

$$\begin{aligned} P(g_{t'}^{-1}(B) \cap g_t^{-1}(B)) &= \int_\Omega P(g_{t'}^{-1}(B) \cap g_t^{-1}(B)|\mathcal{C})dP \\ = \int_\Omega \mu_{t'\omega}(B)\mu_{t\omega}(B)dP &= \int_\Omega \mathbb{E}(1_{g_{t'}^{-1}(B)}(\omega)\mu_{t\omega}(B)|\mathcal{C})dP \\ = \int_\Omega 1_{g_{t'}^{-1}(B)}(\omega)\mu_{t\omega}(B)dP. \end{aligned} \tag{6}$$

Fix any $A \in \mathcal{A}$. Consider the Hilbert space $L_2(\Omega, \mathcal{A}, P)$, and let L be the smallest closed linear subspace which contains both the family of $\mathcal{C}$-measurable functions $\{ \mu_{t'\omega}(B) \mid t \in T_1 \}$ and the family of indicator functions $\{ 1_{g_t^{-1}(B)} \mid t \in T_1 \}$. Let the function $h : \Omega \to \mathbb{R}$ be the orthogonal projection of the indicator function 1_A onto L, with $h^\perp$ as its orthogonal complement. By definition, $1_A = h + h^\perp$ where $h^\perp$ is orthogonal to each member of L. So for all $t \in T_1$, one has $0 = \mathbb{E}(h^\perp \mu_{t\omega}(B)) = \int_\Omega h^\perp \mu_{t\omega}(B)dP$ and $0 = \mathbb{E}(h^\perp 1_{g_t^{-1}(B)}) = \int_\Omega h^\perp 1_{g_t^{-1}}(B)dP$. Because $1_A = h + h^\perp$, it follows that for all $t \in T_1$,

$$\mathbb{E}(1_A\, \mu_{t\omega}(B)) = \mathbb{E}(h\, \mu_{t\omega}(B)) \text{ and } \mathbb{E}(1_A\, 1_{g_t^{-1}(B)}) = \mathbb{E}(h\, 1_{g_t^{-1}(B)}). \tag{7}$$

Next, because $h \in L$, there exists a sequence of functions

$$h_n(\omega) = \sum_{k=1}^{i_n} \left[\alpha_{kn}\, \mu_{t_{kn}\omega}(B) + \beta_{kn} 1_{g_{t_{kn}}^{-1}(B)}(\omega) \right] \ (n = 1, 2, \ldots)$$

with $t_{kn} \in T_1$, as well as α_{kn} and β_{kn} $(k = 1, \ldots, i_n)$ all real constants, such that $h_n \to h$ in the norm of $L_2(\Omega, \mathcal{A}, P)$ — that is, $\int_\Omega (h_n - h)^2 dP \to 0$.

Let T_{kn} be the intersection of the set of t for which Equation (4) holds when $t' = t_{kn}$ with the set of t' for which Equation (5) holds. By hypothesis, $\lambda(T_{kn}) = 1$ because each $t_{kn} \in T_1$. Define $T^* := T_1 \cap \left(\cap_{n=1}^\infty \cap_{k=1}^{i_n} T_{kn} \right)$. Because T^* is the intersection of a countable family of sets all having measure 1 w.r.t. λ, it follows that $\lambda(T^*) = 1$. Also, for any $t \in T^*$, Equation (7) and the limiting property of the sequence h_n imply that

$$P(A \cap g_t^{-1}(B)) = \mathbb{E}(1_A 1_{g_t^{-1}(B)}) = \mathbb{E}(h\, 1_{g_t^{-1}(B)}) = \lim_{n\to\infty} \mathbb{E}(h_n 1_{g_t^{-1}(B)}) \tag{8}$$

It follows from Equations (4)–(6) that

$$\mathbb{E}(h_n 1_{g_t^{-1}(B)}) \quad = \quad \sum_{k=1}^{i_n} \left[\alpha_{kn}\, \mathbb{E}(\mu_{t_{kn}\omega}(B)\, 1_{g_t^{-1}(B)}) + \beta_{kn} \mathbb{E}(1_{g_{t_{kn}}^{-1}(B)}\, 1_{g_t^{-1}(B)}) \right]$$

$$
\begin{aligned}
&= \sum_{k=1}^{i_n} \left[\alpha_{kn}\, \mathbb{E}(\mu_{t_{kn}\omega}(B)\, \mu_{t\omega}(B)) + \beta_{kn} \mathbb{E}(1_{g_{t_{kn}}^{-1}(B)}\, \mu_{t\omega}(B)) \right] \\
&= \mathbb{E}(h_n(\omega) \mu_{t\omega}(B)).
\end{aligned}
$$

Taking the limit as $n \to \infty$ and using Equation (8), one infers that

$$
P(A \cap g_t^{-1}(B)) = \mathbb{E}(h 1_{g_t^{-1}(B)}) = \mathbb{E}(h(\omega)\mu_{t\omega}(B)) = \int_A \mu_{t\omega}(B) dP,
$$

where the last equality follows from Equation (7). Since $\lambda(T^*) = 1$, this completes the proof. ∎

Lemma 2 *Let g be a process from $T \times \Omega$ to X. Let $\mathcal{C} \subseteq \mathcal{A}$ be a countably generated σ-algebra on Ω and μ a $\mathcal{T} \otimes \mathcal{C}$-measurable mapping from $T \times \Omega$ to $\mathcal{M}(X)$. Assume that for each fixed $A \in \mathcal{A}$ and $B \in \mathcal{B}$, one has*

$$
P(A \cap g_t^{-1}(B)) = \int_A \mu_{t\omega}(B)\, dP \tag{9}
$$

for λ-a.e. $t \in T$. Then:

1. *for λ-a.e. $t \in T$, $\mu_{t\omega}(\cdot)$ is a regular conditional distribution $P(g_t^{-1}|\mathcal{C})$ of the random variable g_t;*

2. *for any fixed $k \in \mathbb{N}$ with $k \geq 2$, the process g is essentially k-wise mutually conditionally independent given $\mathcal{C}$.*

Proof: Let $\mathcal{C}^\pi = \{C_n\}_{n=1}^\infty$ and $\mathcal{B}^\pi = \{B_m^\pi\}_{m=1}^\infty$ be countable π-systems for $\mathcal{C}$ and $\mathcal{B}$ respectively. For each pair (m, n), there exists a set T_{mn} with $\lambda(T_{mn}) = 1$ such that for all $t \in T_{mn}$, Equation (9) holds with $A = C_n$ and $B = B_m^\pi$. So for any $t \in T^* := \cap_{m=1}^\infty \cap_{n=1}^\infty T_{mn}$, Equation (9) holds whenever $A = C_n$ and $B = B_m^\pi$, for all pairs (m, n) simultaneously. Because $\mathcal{C}^\pi$ is a π-system that generates $\mathcal{C}$, Dynkin's π–λ theorem (see [5], p. 404) implies that Equation (9) must hold whenever $t \in T^*$, for all $A \in \mathcal{C}$ and all $B \in \mathcal{B}^\pi$ simultaneously. Finally, because $\mathcal{B}^\pi$ is a π-system that generates $\mathcal{B}$, Equation (9) must hold whenever $t \in T^*$, $A \in \mathcal{C}$, and $B \in \mathcal{B}$. In particular, $\mu_{t\omega}$ must be a version of the regular conditional distribution $P(g_t^{-1}|\mathcal{C})$, for all $t \in T^*$.

After this preliminary step, we prove by induction on k that for $\bar{\lambda}^k$-a.e. $(t_1, \ldots, t_k) \in T^k$, one has

$$
\mathbb{E}\left(\prod_{i=1}^k 1_{g_{t_i}^{-1}(B_i)} \middle| \mathcal{C} \right) = \prod_{i=1}^k \mu_{t_i\omega}(B_i) \tag{10}
$$

for all $B_i \in \mathcal{B}^\pi$ ($i = 1$ to k). When $k = 1$, this is shown in the last paragraph.

As the induction hypothesis, suppose that (10) holds for $k-1$ (where $k \geq 2$). That is, for $\bar{\lambda}^{k-1}$-a.e. $(t_1, \ldots, t_{k-1}) \in T^{k-1}$, one has

$$\mathbb{E}\left(\prod_{i=1}^{k-1} 1_{g_{t_i}^{-1}(B_i)} \middle| \mathcal{C}\right) = \prod_{i=1}^{k-1} \mu_{t_i\omega}(B_i) \tag{11}$$

for all $B_i \in \mathcal{B}^\pi$ ($i = 1$ to $k-1$). Take any $(t_1, \ldots, t_{k-1}) \in T^{k-1}$ with the above property.

Fix any $C \in \mathcal{C}^\pi$, and any family B_i ($i = 1, \ldots, k$) from the countable π-system $\mathcal{B}^\pi$. Because (9) holds in particular when $A = C \cap [\cap_{i=1}^{k-1} g_{t_i}^{-1}(B_i)]$, it follows that for λ-a.e. $t_k \in T^*$,

$$\begin{aligned} \int_C \prod_{i=1}^{k} 1_{g_{t_i}^{-1}(B_i)} dP \;&=\; P\left(C \cap \left[\cap_{i=1}^{k} g_{t_i}^{-1}(B_i)\right]\right) \\ =\; \int_{C \cap \left[\cap_{i=1}^{k-1} g_{t_i}^{-1}(B_i)\right]} \mu_{t_k\omega}(B_k) dP \;&=\; \int_C \prod_{i=1}^{k-1} 1_{g_{t_i}^{-1}(B_i)}\, \mu_{t_k\omega}(B_k) dP. \end{aligned} \tag{12}$$

Since $\mu_{t_k\omega}(\cdot) = P(g_{t_k}^{-1}(\cdot)|\mathcal{C})$ for $t_k \in T^*$,

$$\begin{aligned} \mathbb{E}\left(\prod_{i=1}^{k-1} 1_{g_{t_i}^{-1}(B_i)}(\omega)\, \mu_{t_k\omega}(B_k) \middle| \mathcal{C}\right) \;&=\; \mu_{t_k\omega}(B_k)\mathbb{E}\left(\prod_{i=1}^{k-1} 1_{g_{t_i}^{-1}(B_i)} \middle| \mathcal{C}\right) \\ &=\; \prod_{i=1}^{k} \mu_{t_i\omega}(B_i) \end{aligned} \tag{13}$$

by Equation (11). Hence Equations (12) and (13) imply that

$$\int_C \prod_{i=1}^{k} 1_{g_{t_i}^{-1}(B_i)} dP = \int_C \prod_{i=1}^{k} \mu_{t_i\omega}(B_i)\, dP. \tag{14}$$

Summarizing, given any fixed $C \in \mathcal{C}^\pi$ and any fixed family B_i ($i = 1, \ldots, k$) from the countable π-system $\mathcal{B}^\pi$, we have shown that Equation (14) holds for $\bar{\lambda}^{k-1}$-a.e. $(t_1, \ldots, t_{k-1}) \in T^{k-1}$ and for λ-a.e. $t_k \in T$. Let D be the set of all $(t_1, \ldots, t_k) \in T^k$ such that Equation (14) fails. Then, we know that the iterated integral

$$\int_{t_1 \in T_1} \cdots \int_{t_k \in T_k} 1_D \; d\lambda(t_k) \ldots d\lambda(t_1)$$

is zero. But the symmetry of Equation (14) implies that D is symmetric — i.e., for any permutation π on $\{1,\ldots,k\}$ and any $(t_1,\ldots,t_k) \in T^k$, one has $(t_1,\ldots,t_k) \in D$ if and only if $(t_{\pi(1)},\ldots,t_{\pi(q)}) \in T^k$. For this reason, all the iterated integrals of 1_D in any other order are also zero, which means that D is iteratively null. This proves that Equation (14) holds for $\bar{\lambda}^k$-a.e. $(t_1,\ldots,t_k) \in T^k$.

For each $n \in \mathbb{N}$ and each list $m^k = (m_i)_{i=1}^k \in \mathbb{N}^k$, let $T^k(n, m^k)$ denote the set of all $(t_1,\ldots,t_k) \in T^k$ such that Equation (14) holds for the sets C_n and $B^\pi_{m_i}$ ($i = 1$ to k) of the countable π-systems $\mathcal{C}^\pi$ and $\mathcal{B}^\pi$ respectively. We have just proved that $\bar{\lambda}^k(T^k(n, m^k)) = 1$. Define the set

$$\hat{T}^k := (T^*)^k \cap \cap_{n=1}^\infty \cap_{m_1=1}^\infty \cdots \cap_{m_k=1}^\infty T^k(n, m^k)$$

where $(T^*)^k$ denotes the k fold Cartesian product of the set T^*. Then $\bar{\lambda}^k(\hat{T}^k) = 1$. For all $(t_1,\ldots,t_k) \in \hat{T}^k$ and all $B_i \in \mathcal{B}^\pi$ ($i = 1$ to k), Equation (14) holds for all $C \in \mathcal{C}^\pi$. Because $\mathcal{C}^\pi$ is a π-system for $\mathcal{C}$, it follows that Equation (10) holds for all $(t_1,\ldots,t_k) \in \hat{T}^k$, and all $B_i \in \mathcal{B}^\pi$ ($i = 1$ to k). This proves the induction step.

Finally, since $\mathcal{B}^\pi$ is a π-system for $\mathcal{B}$, for each $(t_1,\ldots,t_k) \in \hat{T}^k$ Equation (10) holds for all $B_i \in \mathcal{B}$ ($i = 1$ to k), and also $P(g_{t_i}^{-1}|\mathcal{C}) = \mu_{t_i\omega}$. Hence, the random variables $g_{t_1},\ldots g_{t_k}$ are mutually conditionally independent given $\mathcal{C}$. The rest is clear. ■

Remark 2 *We have only shown that the set D where Equation (14) fails is iteratively null. In general it will not be $\mathcal{T}^k$-measurable. That is why we need the concept of iterated completion.*

Proof of Theorem 1: (1) $\Longrightarrow$ (2) follows from Lemmas 1 and 2.

Next consider (2) $\Longrightarrow$ (3). For $k \geq 2$, let E_k be the collection of all the $(t_1,\ldots,t_k) \in T^k$ such that $g_{t_1},\ldots,g_{t_k}$ are mutually conditionally independent given $\mathcal{C}$, and let $D_k = E_k \times T^\infty$. By (2) $\bar{\lambda}^\infty(D_k) = \bar{\lambda}^k(E_k) = 1$. Let $D = \cap_{k=2}^\infty D_k$. It is clear that $\bar{\lambda}^\infty(D) = 1$, and that for any $(t_1, t_2,\ldots) \in D$, the random variables in the sequence $\{g_{t_k}\}_{k=1}^\infty$ are mutually conditionally independent given $\mathcal{C}$.

(3) $\Longrightarrow$ (1) is obvious. ■

We use the following corollary in the remaining proofs. It is an obvious implication of Lemmas 1 and 2.

Corollary 4 *Let g be a process from $T \times \Omega$ to X. Let $\mathcal{C} \subseteq \mathcal{A}$ be a countably generated σ-algebra on Ω and μ a $\mathcal{T} \otimes \mathcal{C}$-measurable mapping from $T \times \Omega$ to $\mathcal{M}(X)$. Then the following two conditions are equivalent.*

1. *For all $A \in \mathcal{A}$ and $B \in \mathcal{B}$ one has $P(A \cap g_t^{-1}(B)) = \int_A \mu_{t\omega}(B)dP$ for λ-a.e. $t \in T$.*
2. *Given $\mathcal{C}$, the process g has an essentially regular conditional distribution process μ, and is essentially pairwise conditionally independent given $\mathcal{C}$.*

Proof of Proposition 3: Since $\mu_{t\omega}(\cdot)$ is a regular conditional distribution $P(g_t^{-1}|\mathcal{C})$ of the random variable g_t for λ-a.e. $t \in T$, if the process g is essentially pairwise conditionally independent given $\mathcal{C}$, then Part 2 of Corollary 4 is satisfied, which implies Part 1 of Corollary 4.

Since $\mathcal{C} \subseteq \mathcal{C}'$ and the mapping μ is $\mathcal{T} \otimes \mathcal{C}$-measurable, it is automatically $\mathcal{T} \otimes \mathcal{C}'$-measurable. Because μ satisfies Part 1 of Corollary 4, we can apply Corollary 4 with $\mathcal{C}$ replaced by $\mathcal{C}'$ to show that for $\bar{\lambda}^2$-a.e. $(t_1, t_2) \in T^2$, g_{t_1} and g_{t_1} are conditionally independent given $\mathcal{C}'$. ■

Proof of Corollary 3: It is obvious that (3) $\Longrightarrow$ (1). The proof that (2) $\Longrightarrow$ (3) is similar to the corresponding part in the proof of Theorem 1. It remains only to prove that (1) $\Longrightarrow$ (2).

If the random variables g_t are essentially pairwise exchangeable, then Lemmas 4 and 5 in [8] imply that there is a measurable mapping $\omega \mapsto \mu_\omega$ from $(\Omega, \mathcal{A})$ to $\mathcal{M}(X)$ such that for each $A \in \mathcal{A}$ and $B \in \mathcal{B}$, and for λ-a.e. $t \in T$, one has $P(A \cap g_t^{-1}(B)) = \int_A \mu_\omega(B)\, dP$. (Note that the proofs of Lemmas 4 and 5 in [8] still carry through without the additional assumption of pairwise measurable probabilities that is made in [8]. Nor does it matter that the earlier paper uses the classical rather than the iterative definition of a product null set.) Let $\mathcal{C}$ be the sub-σ-algebra of $\mathcal{A}$ generated by μ_ω, which is countably generated. This means that Part 1 of Corollary 4 is satisfied. By the equivalence in Corollary 4, we obtain that for λ-a.e. $t \in T$, $\mu_\omega(\cdot)$ is a regular conditional distribution $P(g_t^{-1}|\mathcal{C})$ of the random variable g_t, and for $\bar{\lambda}^2$-a.e. $(t_1, t_2) \in T^2$, g_{t_1} and g_{t_2} are conditionally independent given $\mathcal{C}$. By Theorem 1, for $\bar{\lambda}^k$-a.e. $(t_1, \ldots, t_k) \in T^k$ the random variables $g_{t_1}, \ldots g_{t_k}$ are mutually conditionally independent given $\mathcal{C}$ with essentially identical regular conditional distributions. In particular, for $\bar{\lambda}^k$-a.e. $(t_1, \ldots, t_k) \in T^k$ and for any $B_i \in \mathcal{B}$ ($i = 1$ to k),

$$P\left(\cap_{i=1}^k g_{t_i}^{-1}(B_i) \middle| \mathcal{C}\right) = \mathbb{E}\left(\prod_{i=1}^k 1_{g_{t_i}^{-1}(B_i)} \middle| \mathcal{C}\right) = \prod_{i=1}^k \mu_\omega^k(B_i),$$

which implies that

$$P\left((g_{t_1},\dots g_{t_k})^{-1}\left(\prod_{i=1}^{k} B_i\right)\right) = \int_\Omega \mu_\omega^k \left(\prod_{i=1}^{k} B_i\right) dP.$$

But the family of k-fold Cartesian product sets $\prod_{i=1}^{k} B_i$ with $B_i \in \mathcal{B}$ $(i = 1,\dots,k)$ is a π-system for the k-fold product σ-algebra $\mathcal{B}^k$. So for $\bar{\lambda}^k$-a.e. $(t_1,\dots,t_k) \in T^k$, it follows that

$$P((g_{t_1},\dots g_{t_k})^{-1}(V)) = \int_\Omega \mu_\omega^k(V)dP$$

for all $V \in \mathcal{B}^k$. Hence, the random variables g_t are essentially k-wise exchangeable. ■

References

[1] J. Berger, H. Osswald, Y. N. Sun, J. L. Wu, On nonstandard product measure spaces, *Illinois Journal of Mathematics* **46** (2002), 319–330.

[2] W. W. Bledsoe and A. P. Morse, Product measures, *Transactions of the American Mathematical Society* **79** (1955), 173–215.

[3] Y. S. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, 3rd. edition, Springer, New York, 1997.

[4] R. M. Dudley, *Real Analysis and Probability*, Chapman & Hall, New York, 1989.

[5] R. Durrett, *Probability: Theory and Examples*, Wadsworth, Belmont, California, 1991.

[6] N. Etemadi, Maximal inequalities for averages of i.i.d. and 2-exchangeable random variables, *Stat. Probabil. Lett.* **44** (1999), 195–200.

[7] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd Edition, Wiley, New York, 1968.

[8] P. J. Hammond and Y. N. Sun, Monte Carlo simulation of macroeconomic risk with a continuum of agents: The symmetric case, *Economic Theory* **21** (2003), 743–766.

[9] O. Kallenberg, From optional skipping to random time change — on some recent advances in exchangeability theory, *Theory of Probability and Applications* **37** (1992), 67–75.

[10] J. F. C. Kingman, Uses of exchangeability, *Annals of Probability* **6** (1978), 183–197.

[11] P. A. Loeb, Conversion from nonstandard to standard measure spaces and applications in probability theory, *Transactions of the Amererican Mathematical Society* **211** (1975), 113–122.

[12] Y. N. Sun, The almost equivalence of pairwise and mutual independence and the duality with exchangeability, *Probability Theory and Related Fields* **112** (1998), 425–456.

[13] Y. N. Sun, A theory of hyperfinite processes: The complete removal of individual uncertainty via exact LLN. *Journal of Mathematical Economics* **29** (1998), 419–503.

[14] Y. N. Sun, The exact law of large numbers via Fubini extension and characterization of insurable risks, *Journal of Economic Theory*, in press; preprint available at
`http://www.math.nus.edu.sg/~matsuny/ELLN.pdf`