The Pennsylvania State University

The Graduate School

College of the Liberal Arts

ESSAYS IN PURE AND APPLIED GAME THEORY

A Dissertation in

Economics

by

Tai-Wei Hu

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

August 2009

The dissertation of Tai-Wei Hu was reviewed and approved* by the following:

Kalyan Chatterjee
Distinguished Professor of Economics and Management Science
Dissertation Co-Adviser, Co-Chair of Committee

Neil Wallace
Professor of Economics
Dissertation Co-Adviser, Co-Chair of Committee

Edward J. Green
Professor of Economics

Stephen G. Simpson
Professor of Mathematics

Vijay Krishna
Distinguished Professor of Economics
Director of Graduate Studies

*Signatures are on file in the Graduate School.

Abstract

This dissertation consists of three chapters.

Chapter 1 presents an axiomatization of expected utility from the frequentist perspective. It starts with a preference relation on the set of infinite sequences with limit relative frequencies. We consider three axioms parallel to the ones for the vN-M expected utility theory. Limit relative frequencies correspond to probability values in lotteries in the vN-M theory. This correspondence is used to show that each of our axioms is equivalent to the corresponding vN-M axiom in the sense that the former is an exact translation of the latter. As a result, a representation theorem is established: The preference relation is represented by an average of utilities with weights given by the relative frequencies.

Chapter 2 reconsiders solutions to the problem of coexistence of money and higher-return substitute of media-of-exchange using imperfect recognizability. Most of the literature on imperfect recognizability focuses on pooling equilibrium and some of it assumes a zero cost of counterfeiting. We replace the zero-cost assumption with a positive cost of counterfeiting, and place the analysis within a signalling game framework in which the intuitive criterion is invoked. With these assumptions, there is no equilibrium in which only money is traded. In particular, if the cost of counterfeiting the substitutes is small, then there is no monetary equilibrium. Therefore, the counterfeiting of substitutes for money can be a threat to monetary systems. This result provides a new rationale for legal restrictions that prohibit substitute media-of-exchange.

Chapter 3 gives a new theory of mixed strategies in zero-sum games. Unpredictable behavior is central for optimal play in many strategic situations because a predictable pattern leaves a player vulnerable to exploitation. A theory of unpredictable behavior is presented in the context of repeated two-person zero-sum games in which the stage games have no pure strategy equilibrium. Players are endowed with sets of feasible functions to generate their strategies. Two dimensions of complexity of these feasible functions are considered: one considers the computability relation and the other considers Kolmogorov complexity. Equilibrium existence is shown under a sufficient condition called mutual complexity. A close characterization of unpredictable strategies is obtained using the criterion stochasticity. In particular, this characterization implies that the failure of some statistical properties of randomness does not justify rejection of equilibrium play.

Table of Contents

Acknowledgments

I am deeply indebted to my advisors, Kalyan Chatterjee and Neil Wallace, for what I have learned in my doctoral studies. Much time they have spent in patiently teaching me game theory and monetary economics. My understanding of these subjects was very much enriched by their insights and experiences. In various occasions they have provided warm encouragements without which I could not finish my dissertation.

I am grateful to Edward Green for his sharp insights and comments to my works. Those advices have much improved this dissertation. I am indebted to Stephen Simpson for his introductory course that equips me with sufficient knowledge in mathematics for this dissertation, especially for the results reported in Chapters 1 and 3.

I am also very thankful to Mamoru Kaneko, who has helped me since I was a master student. His discussions with me in many occasions about new ideas and new approaches to game theory have much widened my view and some of them have inspired the results in Chapters 1 and 3.

I am very grateful to my wife for her support and understanding during the past three years in Penn State. Her sacrifice has made this dissertation possible and she has brought great comfort during periods of disappointment and stress.

I was much blessed by the State College Chinese Alliance Church during my stay in State College. The solid teachings there make me strong during times of weakness and the warm fellowship with brothers and sisters there has kept me in the love of God, without whose grace I could not finish my doctoral studies in such a short time and with so many difficulties.

Dedication

To my wife, Ya-Chen, and my daughter, Joycelyn.

Chapter 1

Expected Utility Theory from the Frequentist Perspective

1.1 Introduction

We study the von Neumann-Morgenstern (vN-M) expected utility theory from the frequentist perspective of probability. von Neumann and Morgenstern [1944] (page 19) emphasized this perspective for expected utility theory, but they did not explicitly formulate it in their development of expected utility theory. In the literature after them, the notion of probability in expected utility theory is not considered from the frequentist viewpoint [1]. Expected utility theory based on the frequentist theory will be crucial when we take experiences seriously into account, such as in inductive game theory (Kaneko and Kline [2008]). In this chapter, we present a frequentist axiomatization of expected utility. This gives a frequentist foundation for expected utility theory.

Although the frequentist interpretation is traced back to the middle of the nineteenth century

[1] See Barbera et al. [1998] for a survey on expected utility theory

(Gillies [2000], chapter 5), the modern frequentist theory began with von Mises [1981][2]. He attempted to construct a formal system based on infinite sequences of outcomes from repetitive experiments. He gives two requirements for such a sequence:

(i) it has a well-defined limit relative frequency for each outcome;

(ii) it is random in the sense that it has no pattern generated by a finite rule.

A sequence satisfying these two requirements is called a *collective* in von Mises [1981]; however, he did not succeed in finding a rigorous definition of the second requirement. Wald [1938] gave the first rigorous definition of randomness, but Ville [1939] showed that it was not yet satisfactory. In the recent literature, a satisfactory definition is emerging (see Downey et al. [2006] for a recent survey).

We reconsider the vN-M expected utility theory from the frequentist perspective of probability, and adopt collectives as the objects of preferences. Requirement (ii) is inessential for our axiomatization in that the main results are not affected regardless of whether or not we impose requirement (ii) in our system. In this chapter, we focus on a system without requirement (ii); thus, a collective in this paper is an infinite sequence satisfying requirement (i). We will discuss the results when requirement (ii) is incorporated in Section 1.5.2. Thus, our approach is capable of interpreting probability values in expected utility theory as generated by a well-defined random mechanism or as frequencies regardless of such random mechanisms.

Our approach starts with a preference relation over infinite sequences of outcomes satisfying requirement (i), and we propose three axioms on the preference relation. Our main result shows that our axioms correspond to the vN-M axioms. This correspondence is based on the translation mapping a collective to a lottery having the same probability values as the frequencies in that

[2] In the literature, there are alternatives to the frequentist theory (see Weatherford [1982] or Gillies [2000] for a survey of interpretations of probability); namely, the classical theory, the subjective theory, and the logical theory.

collective. With this translation, we show that the two axiomatic systems are equivalent. Here, we emphasize that the underlying structures of the two systems are still different, but that the above translation allows us to compare them.

We use the equivalence result to obtain a representation theorem: It states that the preference relation over collectives is represented by the long-run average criterion. This representation theorem, together with the above equivalence theorem, gives the frequentist foundation for expected utility theory.

The rest of this chapter is organized as follows. A review of the vN-M expected utility theory is given in Section 1.2. Our axioms, the equivalence and the representation theorems are presented in Section 1.3. A variant system based on finite sequences is presented in Section 1.4. We discuss our results and some possible extensions in Section 1.5. Proofs of the main theorems and lemmas are given in Section 1.6.

1.2 Expected Utility Theory

Here, we give a small summary of the vN-M expected utility theory. Consider a finite set of *outcomes* $X = \{x_1, ..., x_n\}$. The set of *lotteries* over X is given as

$$\Delta(X) = \{p \in (K[0,1])^n : \sum_{x \in X} p_x = 1\},$$

where $K[0,1]$ is either the set of real numbers from 0 to 1 ($K = \mathbb{R}$) or the set of rational numbers from 0 to 1 ($K = \mathbb{Q}$). While our main results hold for both cases, the latter case ($K = \mathbb{Q}$) will be used in Section 1.4. An element $p \in \Delta(X)$ is denoted as $p = (p_x)_{x \in X} = (p_{x_1}, p_{x_2}, ..., p_{x_n})$.

A *preference relation* $\precsim^P$ is a binary relation over $\Delta(X)$. Given this relation, the *indifference* *relation* $\sim^P$ and the *strict preference relation* $\prec^P$ are defined as follows: For any $p, q \in \Delta(X)$,

$$p \sim^P q \text{ if and only if } p \precsim^P q \text{ and } q \precsim^P p; \tag{1.1}$$

$$p \prec^P q \text{ if and only if } p \precsim^P q \text{ and not } q \precsim^P p. \tag{1.2}$$

In the vN-M theory, the concept of a *compound lottery* plays an important role; it is the convex combination of the form $ap + (1-a)q$, where $p,\ q \in \Delta(X)$ are two lotteries and $a \in K[0,1]$ is a number. For a comparison with the corresponding concept in our frequentist approach, we present the operation of taking a convex combination as

$$(p, q, a) \mapsto ap + (1-a)q. \tag{1.3}$$

We will refer to this mapping when we introduce the corresponding operation in our approach.

The vN-M theory has the following three axioms:

EU1 $\precsim^P$ is a complete and transitive binary relation.

EU2 For all $p,\ q,\ r \in \Delta(X)$, if $p \prec^P q \prec^P r$, then there is some $a \in K[0,1]$ such that $q \sim^P ap + (1-a)r$.

EU3 For all $p,\ q,\ r \in \Delta(X)$ and $a \in K[0,1]$ with $a \neq 0$,

$$ap + (1-a)r \precsim^P aq + (1-a)r \text{ if and only if } p \precsim^P q.$$

The representation theorem for expected utility is given in Theorem 1.2.1, where K can be either $\mathbb{R}$ or $\mathbb{Q}$. The proof is a small variant of the standard expected utility theorem, which can be found in Fishburn [1970], p.112-115.

Theorem 1.2.1. *(Expected utility) A preference relation $\precsim^P$ satisfies EU1-EU3 if and only if there exists a function $h : X \to K$ such that for all $p, q \in \Delta(X)$,*

$$p \precsim^P q \;\Leftrightarrow\; \sum_{x \in X} p_x h(x) \leq \sum_{x \in X} q_x h(x).$$

1.3 A Frequentist Axiomatization

In this section, we give an axiomatization of expected utility from the frequentist perspective. We consider a preference relation over infinite sequences over X satisfying the requirement (i) stated in Section 1.1. We will give three axioms for expected utility in our frequentist approach, and show that those axioms are equivalent to the vN-M axioms under some translation between the two approaches. Using this equivalence result, we obtain a representation of the preference relation by the long-run average criterion.

Let $\mathbb{N} = \{0, 1, 2, ...\}$. For any $p = (p_x)_{x \in X} \in \Delta(X)$, a sequence $\xi = (\xi_0, \xi_1, ...) \in X^{\mathbb{N}} = \prod_{t \in \mathbb{N}} X$ is called a *p-sequence* iff

$$\lim_{T \to \infty} \frac{|\{t : 0 \leq t \leq T-1,\ \xi_t = x\}|}{T} = p_x \text{ for each } x \in X. \tag{1.4}$$

That is, the limit relative frequency of x in a p-sequence is p_x for each outcome x. Then, the set of *collectives* over X is defined to be

$$\Omega_X = \{\xi \in X^{\mathbb{N}} : \xi \text{ is a } p\text{-sequence for some } p \in \Delta(X)\}.$$

Collectives here are defined using only requirement (i) mentioned in Section 1.1. The case with requirement (ii) will be discussed in Section 1.5.2.

We consider a preference relation $\precsim$ over the set of collectives Ω_X. The indifference and strict parts, denoted by $\sim$ and $\prec$, respectively, of $\precsim$ are defined in the same manner as in (1.1) and (1.2). We will formulate three axioms on the preference relation, using similar ideas to those behind the vN-M axioms.

In our approach, we need an operation corresponding to the compound lottery operation (1.3). For this purpose, we introduce the *shuffle operator*. Let $\xi = (\xi_0, \xi_1, ...)$ and $\zeta = (\zeta_0, \zeta_1, ...)$ be two collectives in Ω_X, and let $\nu = (\nu_0, \nu_1, ...)$ be an infinite binary sequence in $\{0,1\}^{\mathbb{N}}$. The collectives ξ, ζ correspond to the lotteries p, q, and ν corresponds to the probability weight a in (1.3). The shuffle operator is expressed as

$$(\xi, \zeta, \nu) \mapsto \xi \oslash_\nu \zeta. \tag{1.5}$$

Formally, it is defined by the following equations:

$$(\xi \oslash_\nu \zeta)_0 = (1-\nu_0)\xi_0 + \nu_0\zeta_0 \text{ for } t = 0, \tag{1.6}$$

$$(\xi \oslash_\nu \zeta)_t = (1-\nu_t)\xi_{t-f^\nu(t)} + \nu_t\zeta_{f^\nu(t)} \text{ for } t > 0, \tag{1.7}$$

where $f^\nu(t) = \sum_{s=0}^{t-1} \nu_s$ is the number of occurrences of 1's in the initial segment of ν with length t.

To illustrate this definition, consider the following table:

ν	0	1	0	0	1	1	1	0	...
ξ	ξ_0		ξ_1	ξ_2				ξ_3	...
ζ		ζ_0			ζ_1	ζ_2	ζ_3		...
$\xi \oslash_\nu \zeta$	ξ_0	ζ_0	ξ_1	ξ_2	ζ_1	ζ_2	ζ_3	ξ_3	...

The first line in the table describes the weight sequence ν; in the second line, the elements from ξ appear in the places where ν has value 0; in the third line, the ones from ζ appear in the places where ν has value 1; and those two lines are combined into the shuffled sequence in the bottom. If we put $\nu = (0, 1, 0, 1, 0, 1, 0, 1, ...)$, $\xi = (x, y, x, y, x, y, ...)$, and $\zeta = (y, x, y, x, y, x, ...)$ in the above table, then the shuffled sequence $\xi \oslash_\nu \zeta$ becomes $(x, y, y, x, x, y, y, x, ...)$.

In the expected utility theory as in Section 1.2, it is assumed that a compound lottery is reduced to a lottery in $\Delta(X)$ using convex combination. In our framework, we need to show a corresponding reduction, which is given in the following lemma. Its proof is given in Section 1.6.

Lemma 1.3.1. *Let $\xi, \zeta \in X^{\mathbb{N}}$, $\nu \in \{0, 1\}^{\mathbb{N}}$, and let $p, q \in \Delta(X)$, $a \in [0, 1]$. Suppose that ξ is a p-sequence, ζ is a q-sequence, and ν is an $(a, 1 - a)$-sequence in $\{0, 1\}^{\mathbb{N}}$. Then, $\xi \oslash_\nu \zeta$ is an $(ap + (1 - a)q)$-sequence.*

Here, an $(a, 1 - a)$-sequence is one with limit relative frequencies a and $1 - a$ for 0 and 1, respectively.

Now we are ready to present our axioms on a preference relation $\precsim$ over Ω_X.

A1 $\precsim$ is a complete and transitive binary relation.

A2 For all ξ, ζ, η in Ω_X, if $\xi \prec \zeta \prec \eta$, then there is a number a in $K[0, 1]$ and an $(a, 1-a)$-sequence ν in $\{0, 1\}^{\mathbb{N}}$ such that $\zeta \sim \xi \oslash_\nu \eta$.

A3 For all ξ, ζ, η in Ω_X and all $(a, 1 - a)$-sequences ν^1, ν^2 in $\{0, 1\}^{\mathbb{N}}$ with $a \in K[0, 1]$ and $a \neq 0$, $\xi \oslash_{\nu^1} \eta \precsim \zeta \oslash_{\nu^2} \eta$ if and only if $\xi \precsim \zeta$.

Axioms A1-A3 are parallel to EU1-EU3. Perhaps, Axiom A3 needs more comments, which will be given after our theorems.

Now we give a comparison between our axiomatic system and the vN-M system. For this comparison, we use the mapping ψ from Ω_X to $\Delta(X)$:

$$\psi(\xi) = (\lim_{T\to\infty} \frac{|\{t : 0 \le t \le T-1,\ \xi_t = x\}|}{T})_{x\in X} \text{ for each } \xi \in \Omega_X. \tag{1.8}$$

The mapping ψ specifies the relative frequencies of outcomes in a collective. We can translate our system $(\Omega_X, \precsim)$ into the vN-M system $(\Delta(X), \precsim^P)$ using the mapping ψ if for all $\xi,\ \zeta \in \Omega_X$,

$$\xi \precsim \zeta \text{ if and only if } \psi(\xi) \precsim^P \psi(\zeta). \tag{1.9}$$

This means that the preference relation $\precsim$ over Ω_X can be reduced to the preference relation $\precsim^P$ over $\Delta(X)$, in which case $\precsim$ depends only on relative frequencies.

The following theorem states, first, that our axioms A1 and A3 guarantee the above reduction. Also, it states that under condition (1.9) each of our axioms corresponds precisely to its parallel axiom in the vN-M theory. Its proof will be given in Section 1.6.

Theorem 1.3.1. *(Frequentist translation) (a) Suppose that the preference relation $\precsim$ over Ω_X satisfies A1 and A3. Then there is a preference relation $\precsim^P$ over $\Delta(X)$ satisfying (1.9).*

(b) Suppose that $\precsim$ over Ω_X and $\precsim^P$ over $\Delta(X)$ satisfy (1.9). Then,

(b1) $\precsim$ satisfies A1 if and only if $\precsim^P$ satisfies EU1.

(b2) $\precsim$ satisfies A2 if and only if $\precsim^P$ satisfies EU2.

(b3) $\precsim$ satisfies A3 if and only if $\precsim^P$ satisfies EU3.

Thus, our axiomatic system is a faithful translation of the vN-M system within our frequentist framework. It justifies the frequentist perspective taken by von Neumann and Morgenstern [1944] (p. 19), while their theory is neutral to interpretations of probability.

Theorem 1.3.1 allows us to obtain our representation theorem. Its proof will be given in Section 1.6.

Theorem 1.3.2. *(Frequentist axiomatization of expected utility) A preference relation $\precsim$ satisfies A1-A3 if and only if there exists a function $h : X \to K$ (which we call a representing utility function*[3] *of $\precsim$) such that for all $\xi,\ \zeta \in \Omega_X$,*

$$\xi \precsim \zeta \ \Leftrightarrow \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} \leq \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\zeta_t)}{T}. \tag{1.10}$$

Theorems 1.3.1 and 1.3.2 manifest the parallelism between our theory and the vN-M theory. By this parallelism, we can regard our theory as a frequentist version of the vN-M expected utility theory.

An additional comment on the particular form of (1.10) is still needed; why does the representation take the long-run average criterion? More precisely, where is the hidden assumption for the equal treatment of every element in a collective $\xi = (\xi_0, \xi_1, \xi_2, ...)$? Axiom A3 is the driving force of this equal treatment: In Axiom A3, ν^1 and ν^2 are required only to have the same relative frequencies, and the orders of elements in those sequences can be changed. In fact, the main cause is hidden in the present definition (1.6) and (1.7) of the shuffle operator. It is crucial in the present proof of Theorem 1.3.2. In Section 1.5.2, an alternative formulation of the shuffle operator is discussed, but some difficulty is pointed out.

[3] As in the vN-M theory, representing utility functions are determined by the preference relation to be unique up to positive linear transformations.

1.4 Axiomatization with Finite Sequences

In this section we show that the framework of Section 1.3 can be restricted to finite sequences, but we can keep the same axiomatization of expected utility with a small modification.

Given a finite set X of outcomes, the set of finite sequences over X is denoted by $X^{<\mathbb{N}}$. Each finite sequence is written as $\sigma = (\sigma_0, ..., \sigma_{t-1})$. The length t of σ is denoted by $|\sigma|$.

Consider a preference relation $\precsim^F$ over $X^{<\mathbb{N}}$. We give four axioms, AF1-AF4, on $\precsim^F$; the first three are parallel to Axioms A1-A3, and the last axiom is additional to deal with finite sequences.

The main difference from the previous axiomatization is that the shuffle operator needs to be defined for finite sequences. Another difference is the preference comparison between finite sequences of different lengths, which is taken care of by the new axiom, AF4. Here, the lengths of finite sequences may be different, and this fact causes difficulties with the shuffle operator and preference comparisons. To avoid these difficulties, we replicate input sequences of the shuffle operator so that they have same lengths. For this purpose, we give one definition: For finite sequence σ, $\sigma^t = (\sigma, \sigma, ..., \sigma)$ is the sequence obtained from replicating σ t times. This new sequence preserves the relative frequencies of the original sequence σ.

The finite shuffle operator is formally defined as follows: Given two sequences $\sigma,\ \tau \in X^{<\mathbb{N}}$ and a finite binary sequence $\rho \in \{0,1\}^{<\mathbb{N}}$, they are replicated into $\sigma' = \sigma^{|\tau||\rho|}$, $\tau' = \tau^{|\sigma||\rho|}$, and $\rho' = \rho^{|\tau||\sigma|}$, respectively. Then, those replicated sequences have the same length $|\tau| \cdot |\sigma| \cdot |\rho|$ and preserve the relative frequencies of the original sequences. The *finite-shuffle* of σ and τ with ρ, denoted by $\sigma \oslash_\rho \tau$, has the length $|\tau| \cdot |\sigma| \cdot |\rho|$ and is defined by

$$(\sigma \oslash_\rho \tau)_0 = (1-\rho'_0)\sigma'_0 + \rho'_0\tau'_0 \text{ for } t = 0,$$

$$(\sigma \oslash_\rho \tau)_t = (1-\rho'_t)\sigma'_{t-f^{\rho'}(t)} + \rho'_t\tau'_{f^{\rho'}(t)} \text{ for } t = 1, ..., |\tau| \cdot |\sigma| \cdot |\rho| - 1,$$

where $f^{\rho'}(t) = \sum_{s=0}^{t-1} \rho'_s$.

The finite-shuffle operator is closely related to the shuffle operator defined in Section 1.3: For any finite sequences σ, $\tau \in X^{<\mathbb{N}}$, and $\rho \in \{0,1\}^{<\mathbb{N}}$, if $\xi = (\sigma, \sigma, \sigma, ...)$, $\zeta = (\tau, \tau, \tau, ...)$, and $\nu = (\rho, \rho, \rho, ...)$, then $\xi \oslash_\nu \zeta = (\sigma \oslash_\rho \tau, \sigma \oslash_\rho \tau, \sigma \oslash_\rho \tau, ...)$. Moreover, Lemma 1.3.1 can be modified to the present finite framework.

The four axioms on the preference relation $\precsim^F$ are as follows.

AF1: $\precsim^F$ is a complete and transitive binary relation.

AF2: For all σ, τ, $\pi \in X^{<\mathbb{N}}$, if $\sigma \prec^F \tau \prec^F \pi$, then there is a binary sequence $\rho \in \{0,1\}^{<\mathbb{N}}$ such that $\tau \sim^F \sigma \oslash_\rho \pi$.

AF3: For all σ, τ, $\pi \in X^{<\mathbb{N}}$ and $\rho^1, \rho^2 \in \{0,1\}^{<\mathbb{N}}$ that have the same positive relative frequency of outcome 0,

$$\sigma \oslash_{\rho^1} \pi \precsim^F \tau \oslash_{\rho^2} \pi \text{ if and only if } \sigma \precsim^F \tau.$$

AF4: For all $\sigma \in X^{<\mathbb{N}}$ and for all $t > 0$, $\sigma \sim^F \sigma^t$.

Axioms AF1-AF3 are very parallel to A1-A3, but AF4 is new. It states that a finite sequence σ is indifferent to any replication of itself. It is the essence of this axiom that the preference relation counts only the relative frequencies but not the lengths. This enables us to compare sequences with different lengths.

The finite sequence version of the mapping ψ in (1.8) becomes now the mapping $\phi : X^{<\mathbb{N}} \to \Delta(X)$:

$$\phi(\sigma) = (\frac{|\{t : 0 \le t \le |\sigma| - 1,\ \sigma_t = x\}|}{|\sigma|})_{x \in X}. \tag{1.11}$$

Translation (1.9) then becomes for all $\sigma,\ \tau \in X^{<\mathbb{N}}$,

$$\sigma \precsim^F \tau \text{ if and only if } \phi(\sigma) \precsim^P \phi(\tau), \tag{1.12}$$

where the preference relation $\precsim^P$ is a binary relation over $\Delta(X)$ with $K = \mathbb{Q}$.

Now, Theorem 1.3.1 becomes the following.

Theorem 1.4.1. *(a) Suppose that the preference relation $\precsim^F$ over $X^{<\mathbb{N}}$ satisfies AF1, AF3, and AF4. Then there is a preference relation $\precsim^P$ over $\Delta(X)$ $(K = \mathbb{Q})$ satisfying (1.12).*

(b) Suppose that $\precsim^F$ over $X^{<\mathbb{N}}$ and $\precsim^P$ over $\Delta(X)$ satisfy (1.12). Then,

(b.1) $\precsim^F$ satisfies AF1 if and only if $\precsim^P$ satisfies EU1.

(b.2) $\precsim^F$ satisfies AF2 if and only if $\precsim^P$ satisfies EU2.

(b.3) $\precsim^F$ satisfies AF3 if and only if $\precsim^P$ satisfies EU3.

The proof of Theorem 1.4.1 is quite parallel to that of Theorem 1.3.1. The only difference is to use Axiom AF4 to compare finite sequences with different lengths. We do not give the proof in this dissertation, but it is available upon request.

Then, Theorem 1.3.2 becomes the following.

Theorem 1.4.2. *A preference relation $\precsim^F$ satisfies AF1-AF4 if and only if there exists a function $h : X \to \mathbb{Q}$ such that for all $\sigma,\ \tau \in X^{<\mathbb{N}}$,*

$$\sigma \precsim \tau \ \Leftrightarrow \sum_{t=0}^{|\sigma|-1} \frac{h(\sigma_t)}{|\sigma|} \leq \sum_{t=0}^{|\tau|-1} \frac{h(\tau_t)}{|\tau|}.$$

The proof follows exactly the same arguments as that of Theorem 1.3.2, and is omitted.

One difference between the finite approach and the infinite approach is that the randomness requirement (ii) mentioned in Section 1.1 cannot be incorporated in the finite approach. This may be regarded as a demerit from the frequentist perspective of probability. However, from the viewpoint of expected utility theory, this is rather a merit in the sense that it can be applied to finite sequences.

1.5 Conclusions and Remarks

1.5.1 Conclusions

We gave two axiomatic approaches to expected utility, based on infinite sequences and finite sequences. Theorems 1.3.1 and 1.4.1 describe the correspondences between our axiomatic systems and the vN-M system. Theorems 1.3.2 and 1.4.2 are frequentist expected utility theorems. As mentioned in Section 1.1, von Neumann and Morgenstern [1944], p. 19, emphasized the frequentist interpretation of probability for their expected utility theory. Our theorems give a justification of their interpretation. Moreover, Theorems 1.4.1 and 1.4.2 widened the frequentist perspective to accommodate finite sequences for expected utility theory.

This approach can be used to study game theory, especially for research programs that take into account players' *ex post* experiences. In particular, this approach has the potential to serve a foundation for expected utility in inductive game theory (cf. Kaneko and Kline [2009]).

1.5.2 Remarks

In this subsection, we first give some comments on the technical side of our approach. Then, we discuss a possible extension of our framework and one application.

1. **Randomness requirement:** The randomness requirement can be incorporated in our formulation. For this, we define collectives to be infinite sequences satisfying requirements (i) and (ii) in Section 1.1; specifically, we use the definition of random sequences given by Martin-Löf [1966] to formulate requirement (ii). The axioms are formulated in the same way as in Section 1.3, but we restrict the application of the shuffle operator to "independent" random sequences. We can show that Theorem 1.3.1 and Theorem 1.3.2 hold in this framework as well[4].

2. **Shuffle operator:** We formulate the shuffle operator to reflect compound lotteries, and our formulation is justified by Lemma 1.3.1. One alternative operator is the following: Let $\xi,\ \zeta \in X^{\mathbb{N}}$, and let $\nu \in \{0,1\}^{\mathbb{N}}$; define $\xi \ominus_\nu \zeta$ by setting

$$(\xi \ominus_\nu \zeta)_t = (1-\nu_t)\xi_t + \nu_t\zeta_t \text{ for all } t.$$

However, this simpler operator does not work for Theorems 1.3.1 and 1.3.2. The choice of ν may affect the frequency of the shuffled sequence: Lemma 1.3.1 is no longer valid with this shuffle operator.

If the randomness requirement is incorporated, Lemma 1.3.1 is recovered, but the present author has not yet succeeded in finding a proof for Theorem 1.3.1 using this shuffle operator. The main difficulty is that with this alternative operator, in general the input sequences cannot be recovered from the resulting sequence and the weight sequence, while one can always recover them with the shuffle operator defined in Section 1.3.

3. **A Frequentist definition of subjective probability:** This chapter gives the frequentist version of vN-M expected utility theory. By extending our framework, it is possible to formulate an axiomatic system for a frequentist definition of subjective probability. The subjective probability values thus obtained are interpreted as the personal estimation of the long-run frequencies in a

[4] The precise definitions and results are available upon request.

random sequence encountered in the decision maker's problem. We suspect that the axioms in Anscombe and Aumann [1963] can be used to formulate such an axiomatic system.

4. **Application to game theory:** Our results are used in Chapter 3, which studies strategic unpredictable behavior from the frequentist perspective. There, a collective game that consists of infinite repetitions of a finite two-person zero-sum game is considered, and a play in the collective game is an infinite sequence of joint actions from the two players. The long-run average criterion is applied to specify the payoffs in that chapter, and our axiomatic system provides a foundation for that application.

1.6 Proofs

We give proofs of the main results in Section 1.3 in this section. In the first subsection, we will provide two more lemmas, together with the proof of Lemma 1.3.1. In the second subsection, we give the proofs of Theorems 1.3.1 and 1.3.2.

1.6.1 Preliminary Lemmas

First we introduce the function $L^{\xi,x} : (\mathbb{N} - \{0\}) \rightarrow \mathbb{N}$ that counts the number of occurrences of outcome $x \in X$ in initial segments of a collective $\xi \in \Omega_X$:

$$L^{\xi,x}(T) = |\{t : 0 \leq t \leq T-1, \xi_t = x\}| \text{ for all } T > 0.$$

This function will simplify the notations in our proofs. Now we give the proof of Lemma 1.3.1.

Proof of Lemma 1.3.1: The key to this proof is to notice the following equation:

$$L^{\xi\oslash_\nu\zeta,x}(T) = L^{\xi,x}(L^{\nu,0}(T)) + L^{\zeta,x}(L^{\nu,1}(T)) \text{ for each } T \text{ and for each } x. \tag{1.13}$$

It states that the number of occurrences of outcome x in an initial segment of the shuffled sequence with length T is the sum of those numbers in the initial segments of the input sequences with lengths specified by the numbers of occurrences of 0 and 1 in the initial segment of the weighting sequence with length T. It is straightforward to check its valid using the definition of the shuffle operator, equations (1.6) and (1.7). This equation can be used directly to prove the lemma for $a \in (0,1)$ as follows: For all T large enough, $L^{\nu,0}(T) > 0$ and $L^{\nu,1}(T) > 0$ (actually, $\lim_{T\to\infty} L^{\nu,0}(T) = \infty = \lim_{T\to\infty} L^{\nu,1}(T)$), and so

$$\begin{aligned}
&\lim_{T\to\infty} \frac{L^{\xi\oslash_\nu\zeta,x}(T)}{T} = \lim_{T\to\infty} \Big(\frac{L^{\xi,x}(L^{\nu,0}(T))}{L^{\nu,0}(T)}\frac{L^{\nu,0}(T)}{T} + \frac{L^{\zeta,x}(L^{\nu,1}(T))}{L^{\nu,1}(T)}\frac{L^{\nu,1}(T)}{T}\Big) \\
&= \lim_{T\to\infty} \frac{L^{\xi,x}(L^{\nu,0}(T))}{L^{\nu,0}(T)} \lim_{T\to\infty} \frac{L^{\nu,0}(T)}{T} + \lim_{T\to\infty} \frac{L^{\zeta,x}(L^{\nu,1}(T))}{L^{\nu,1}(T)} \lim_{T\to\infty} \frac{L^{\nu,1}(T)}{T} = ap_x + (1-a)q_x.
\end{aligned}$$

For the case $a = 1$, the above argument does not work because

$$\lim_{T\to\infty} \frac{L^{\nu,1}(T)}{T} = 1 - a = 0, \tag{1.14}$$

and so $\lim_{T\to\infty} L^{\nu,1}(T)$ may be a finite number. However, because

$$\lim_{T\to\infty} \frac{L^{\nu,0}(T)}{T} = a = 1,$$

we still have

$$\lim_{T\to\infty} \frac{L^{\xi,x}(L^{\nu,0}(T))}{L^{\nu,0}(T)} = p_x. \tag{1.15}$$

Now, (1.14) and $L^{\zeta,x}(L^{\nu,1}(T)) \leq L^{\nu,1}(T)$ imply that

$$\lim_{T\to\infty} \frac{L^{\zeta,x}(L^{\nu,1}(T))}{T} = 0.$$

Using (1.13), we have that

$$\lim_{T\to\infty} \frac{L^{\xi\oslash_\nu\zeta,x}(T)}{T} = \lim_{T\to\infty} (\frac{L^{\xi,x}(L^{\nu,0}(T))}{L^{\nu,0}(T)} \frac{L^{\nu,0}(T)}{T}) + \lim_{T\to\infty} \frac{L^{\zeta,x}(L^{\nu,1}(T))}{T} = ap_x = p_x.$$

The case for $a = 0$ is completely symmetric. □

To prove part (a) of Theorem 1.3.1, we use the mathematical induction on the number of outcomes that appear in the collective. In the arguments, we decompose a collective into two subsequences according to a subset of outcomes A and use Axiom A3 to obtain the result. Here, we introduce a decomposition operator that will simplify those arguments. Let $A \subset X$ be a subset and let $\xi \in X^{\mathbb{N}}$. Define $\nu^{\xi,A}$ as $\nu_t^{\xi,A} = 1$ if $\xi_t \in A$ and $\nu_t^{\xi,A} = 0$ otherwise, which indicates whether a particular element in ξ belongs to A or not. To decompose ξ according to A via $\nu^{\xi,A}$, we introduce another function θ^ν. For any $\nu \in \{0,1\}^{\mathbb{N}}$, define the function θ^ν as follows:

$$\theta^\nu(0) \text{ is the least } t' \text{ such that } \nu_{t'} = 1; \tag{1.16}$$

$$\theta^\nu(t+1) \text{ is the least } t' \text{ such that } t' > \theta^\nu(t) \text{ and } \nu_{t'} = 1. \tag{1.17}$$

θ^ν records the places where ν has value 1. Notice that $\theta^\nu(t)$ may not be well-defined for all $t \in \mathbb{N}$. It is easy to check that it is well-defined for all t if and only if ν has infinitely many 1's.

Now, we can form the subsequence ξ^A of ξ obtained by eliminating elements in ξ that are not in A, using the function $\theta^{\nu^{\xi,A}}$: The sequence ξ^A is defined as ($x^0 \in X$ is a fixed outcome)

(1) if $\lim_{T\to\infty} |\{t : 0 \leq t \leq T-1,\ \nu_t^{\xi,A} = 1\}| = \infty$, then let $\xi_t^A = \xi_{\theta^{\nu^{\xi,A}}(t)}$ for all $t \in \mathbb{N}$;

(2) if $\lim_{T\to\infty} |\{t : 0 \le t \le T-1,\ \nu_t^{\xi,A} = 1\}| = K < \infty$, then let $\xi_t^A = \xi_{\theta^{\nu^{\xi,A}}(t)}$ for all $t = 0, ..., K-1$, and let $\xi_t^A = x^0$ for all $t \ge K$.

If infinitely many elements in ξ are in A, then ξ^A is a subsequence of ξ; otherwise, $\xi_t^A = x^0$ for all t large enough. It is straightforward to check that $\xi^{\overline{A}} \oslash_{\nu^{\xi,A}} \xi^A = \xi$ ($\overline{A}$ denotes the complement of A). That is, ξ can be decomposed into two disjoint subsequences ξ^A and $\xi^{\overline{A}}$ (provided that it has infinite elements belonging to A and $\overline{A}$) such that ξ^A is a sequence over A and $\xi^{\overline{A}}$ is a sequence over $\overline{A}$. Moreover, if ξ is a collective, so are ξ^A and $\nu^{\xi,A}$. Their relative frequencies can be found using conditional probability. The following lemma summarizes these properties.

Lemma 1.6.1. *Let $p \in \Delta(X)$. Suppose that ξ is a p-sequence. We have*

(1) $\nu^{\xi,A}$ is an $(1 - p_A, p_A)$-sequence, where $p_A = \sum_{x\in A} p_x$.

(2) if $p_A > 0$, then ξ^A is a p^A-sequence, where $p_x^A = \frac{p_x}{p_A}$ if $x \in A$ and $p_x^A = 0$ otherwise.

Proof. (1) It is straightforward to check that $L^{\nu^{\xi,A},1}(T) = \sum_{x\in A} L^{\xi,x}(T)$ for all $T > 0$. Thus,

$$\lim_{T\to\infty} \frac{L^{\nu^{\xi,A},1}(T)}{T} = \sum_{x\in A} \lim_{T\to\infty} \frac{L^{\xi,x}(T)}{T} = \sum_{x\in A} p_x = p_A.$$

(2) We first deal with outcomes not in A. For each $x \notin A$, by definition (here part (1) of the definition applies because $p_A > 0$) of ξ^A, $\xi_t^A \neq x$ for all $t \in \mathbb{N}$. Thus,

$$\lim_{T\to\infty} \frac{L^{\xi^A,x}(T)}{T} = 0 = p_x^A.$$

Now we consider outcomes in A. Since $p_A > 0$, $\lim_{T\to\infty} L^{\nu^{\xi,A},1}(T) = \infty$ by part (1) of this lemma. As a result, $\theta^{\nu^{\xi,A}}(t)$ is well-defined for all $t \in \mathbb{N}$. Because $\theta^{\nu^{\xi,A}}$ is strictly increasing by equation (1.17), we have that $\lim_{T\to\infty} \theta^{\nu^{\xi,A}}(T) \to \infty$.

We claim that

$$\lim_{T\to\infty} \frac{T}{\theta^{\nu^{\xi,A}}(T-1)+1} = p_A. \tag{1.18}$$

By construction, $\theta^{\nu^{\xi,A}}(T-1)$ is the place in $\nu^{\xi,A}$ where the T-th occurrence of 1 takes place. Thus, there are exactly T occurrences of 1's in the first $\theta^{\nu^{\xi,A}}(T-1)+1$ elements in $\nu^{\xi,A}$, and so

$$L^{\nu^{\xi,A},1}(\theta^{\nu^{\xi,A}}(T-1)+1) = T \text{ for all } T \geq 1. \tag{1.19}$$

Hence, by (1.19), the sequence $\{\frac{T}{\theta^{\nu^{\xi,A}}(T-1)+1}\}_{T=1}^{\infty}$ is a subsequence of $\{\frac{L^{\nu^{\xi,A},1}(T)}{T}\}_{T=1}^{\infty}$. By part (1), the latter sequence has limit p_A, and so the former sequence has the same limit. This validates (1.18).

For any $x \in A$,

$$\begin{aligned}
&\lim_{T\to\infty} \frac{L^{\xi^A,x}(T)}{T} = \lim_{T\to\infty} \frac{L^{\xi,x}(\theta^{\nu^{\xi,A}}(T-1)+1)}{\theta^{\nu^{\xi,A}}(T-1)+1} \frac{\theta^{\nu^{\xi,A}}(T-1)+1}{T} \\
&= \lim_{T\to\infty} \frac{L^{\xi,x}(\theta^{\nu^{\xi,A}}(T-1)+1)}{\theta^{\nu^{\xi,A}}(T-1)+1} \lim_{T\to\infty} \frac{\theta^{\nu^{\xi,A}}(T-1)+1}{T} \\
&= p_x \frac{1}{\lim_{T\to\infty} \frac{T}{\theta^{\nu^{\xi,A}}(T-1)+1}} = \frac{p_x}{p_A} = p_x^A.
\end{aligned} \tag{1.20}$$

It is straightforward to check that $L^{\xi^A,x}(T) = L^{\xi,x}(\theta^{\nu^{\xi,A}}(T-1)+1)$ and this gives the first equality in (1.20). The last equality in (1.20) comes from (1.18) and the fact that the limit of the inverses of a sequence is equal to the inverse of its limit. $\square$

We end this subsection with a lemma that links expected values to long-run averages. This lemma is the key step to obtain the long-run average criterion (1.10).

Lemma 1.6.2. *Let $h : X \to \mathbb{R}$ be any function. Suppose that $\xi \in X^{\mathbb{N}}$ is a p-sequence for some*

$p \in \Delta(X)$. *Then*

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} = \sum_{x\in X} p_x h(x).$$

Proof. For any $x' \in X$, $h(x') = \sum_{x\in X} c_{x'}(x)h(x)$, where $c_{x'}(x) = 1$ if $x = x'$ and $c_{x'}(x) = 0$ otherwise. We first show that for all $x^0 \in X$,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_{x^0}(\xi_t)}{T} = \sum_{x\in X} p_x c_{x^0}(x) = p_{x^0}. \tag{1.21}$$

For each $T \in \mathbb{N}$,

$$\sum_{t=0}^{T-1} c_{x^0}(\xi_t) = |\{t : 0 \le t \le T-1,\ \xi_t = x^0\}|.$$

Since ξ is a p-sequence, it follows from (1.4) that

$$\lim_{T\to\infty} \frac{|\{t : 0 \le t \le T-1,\ \xi_t = x^0\}|}{T} = p_{x^0}.$$

Thus, (1.21) holds. Therefore, we have

$$\lim_{T\to\infty} \textstyle\sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} = \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{\sum_{x\in X} c_x(\xi_t)h(x)}{T}$$

$$= \textstyle\sum_{x\in X} \lim_{T\to\infty} h(x) \sum_{t=0}^{T-1} \frac{c_x(\xi_t)}{T} = \sum_{x\in X} h(x)p_x = \sum_{x\in X} p_x h(x). \qquad \square$$

1.6.2 Proofs of Main Theorems

Proof of Theorem 1.3.1: (a) We claim that if $\precsim$ satisfies A1 and A3, then for any $p \in \Delta(X)$ and any p-sequences ξ and ζ, $\xi \sim \zeta$. Given this claim, it is straightforward to construct a preference relation $\precsim^P$ that satisfies (1.9).

First we give some notations. For any $x \in X$, we use $\mathbf{x}$ to denote the sequence such that for all $t \in \mathbb{N}$, $\mathbf{x}_t = x$. Also, for any $p \in \Delta(X)$, we define $S(p) = \{x \in X : p_x > 0\}$, that is, the set of outcomes in the support of p.

We prove the claim by induction on the number of outcomes in $S(p)$.

The basis: Suppose that $|S(p)| = 1$. Let $S(p) = \{x\}$. The difficulty here arises when a p-sequence still contains other outcomes than x.

Let ξ be a p-sequence. Consider $\xi^{\overline{\{x\}}}$ and $\nu^{\xi,\overline{\{x\}}}$. By construction, we have that $\xi = \mathbf{x} \oslash_{\nu^{\xi,\overline{\{x\}}}} \xi^{\overline{\{x\}}}$, and, by Lemma 1.6.1, $\nu^{\xi,\overline{\{x\}}}$ is an $(1,0)$-sequence. By A3,

$$\xi = \mathbf{x} \oslash_{\nu^{\xi,\overline{\{x\}}}} \xi^{\overline{\{x\}}} \precsim \mathbf{x} \oslash_{\mathbf{0}} \xi^{\overline{\{x\}}} = \mathbf{x} \text{ if and only if } \mathbf{x} \precsim \mathbf{x},$$

and since $\mathbf{x} \sim \mathbf{x}$ by A1, $\xi \sim \mathbf{x}$ (here, $\mathbf{0} = (0, 0, 0, ...)$ is the sequence consisting of 0's). Thus, for any an p-sequence sequence ξ, $\xi \sim \mathbf{x}$, and so by A1, for any two p-sequences ξ and ζ, we have $\xi \sim \zeta$.

The inductive step: Our induction hypothesis is that for any $p \in \Delta(X)$ with $|S(p)| \leq k$, $k \geq 1$, any two p-sequences are indifferent.

Consider any $q \in \Delta(X)$ with $|S(q)| = k + 1$ and any q-sequences ξ and ζ. We will show that $\xi \sim \zeta$.

Let $y \in S(q)$, and let $A = S(q) - \{y\}$. First we show that there are q-sequences ξ' and ζ' that consists of only outcomes in $S(q)$ such that $\xi \sim \xi'$ and $\zeta \sim \zeta'$. Moreover, we show that $\xi' \sim \zeta'$, and this will give us the desired result.

Notice that $\xi = \xi^{\overline{A}} \oslash_{\nu^{\xi,A}} \xi^{A}$ ($q_{\overline{A}} > 0$ and $q_A > 0$), and, by Lemma 1.6.1, $\nu^{\xi,A}$ is an $(1 - q_A, q_A)$-sequence. Because $\xi^{\overline{A}}$ is a $p^{\{y\}}$-sequence ($p^{\{y\}}_x = 1$ if $x = y$ and $p^{\{y\}}_x = 0$ otherwise), by the induction hypothesis, $\xi^{\overline{A}} \sim \mathbf{y}$.

Define $\xi' = \mathbf{y} \oslash_{\nu^{\xi,A}} \xi^{A}$. Then, by A3 and $\xi^{\overline{A}} \sim \mathbf{y}$, we have $\xi' \sim \xi$. Similarly, we define ζ' as

$\mathbf{y} \oslash_{\nu^{\zeta,A}} \zeta^A$, and, using the same arguments, we have $\zeta' \sim \zeta$.

By Lemma 1.6.1, both ξ^A and ζ^A are q^A-sequences. By the induction hypothesis, $\xi^A \sim \zeta^A$. Notice that ξ' can be written as $\xi^A \oslash_{\nu^{\xi,\overline{A}}} \mathbf{y}$ and ζ' can be written as $\zeta^A \oslash_{\nu^{\zeta,\overline{A}}} \mathbf{y}$. Since both $\nu^{\xi,\overline{A}}$ and $\nu^{\zeta,\overline{A}}$ are $(q_A, 1-q_A)$-sequences, it follows from A3 that $\xi' \sim \zeta'$

Because $\xi' \sim \xi$, $\zeta' \sim \zeta$, and $\xi' \sim \zeta'$, it follows from A1 that $\xi \sim \zeta$. This completes our inductive step. By mathematical induction, we have proved our claim.

Now, we define $\precsim^P$ to be such that $p \precsim^P q$ if and only if $\xi \precsim \zeta$ for some p-sequence ξ and some q-sequence ζ. The above claim shows that $\precsim^P$ is well-defined and satisfies condition (1.9).

(b) Suppose that there is a preference relation $\precsim^P$ over $\Delta(X)$ that satisfies (1.9).

(b.1) ($\Rightarrow$) Suppose that $\precsim$ satisfies A1, i.e, it is complete and transitive. First we show that $\precsim^P$ is transitive. Let $p,\ q,\ r \in \Delta(X)$ be such that $p \precsim^P q$ and $q \precsim^P r$. There are $\xi,\ \zeta,\ \eta \in \Omega_X$ such that ξ is a p-sequence, ζ is a q-sequence, and η is a r-sequence. Then, by (1.9), $\xi \precsim \zeta$ and $\zeta \precsim \eta$. Since $\precsim$ is transitive, it follows that $\xi \precsim \zeta$. Hence, by (1.9), $p \precsim^P r$. The proof for completeness is similar.

($\Leftarrow$) Suppose that $\precsim^P$ satisfies EU1. Let $\xi,\ \zeta,\ \eta \in \Omega_X$ be such that $\xi \precsim \zeta$ and $\zeta \precsim \eta$. Let $p,\ q,\ r \in \Delta(X)$ be such that ξ is a p-sequence, ζ is a q-sequence, and η is a r-sequence. It then follows, from (1.9), that $p \precsim^P q$ and $q \precsim^P r$. Since $\precsim^P$ is transitive, it follows that $p \precsim^P r$. Hence, by (1.9), $\xi \precsim \eta$. The proof for completeness is similar.

(b.2) ($\Rightarrow$) Suppose that $\precsim$ satisfies A2. We show that $\precsim^P$ satisfies EU2 by considering lotteries $p, q, r \in \Delta(X)$ that satisfy $p \prec^P q \prec^P r$. Let ξ be a p-sequence, ζ be a q-sequence, and η be a r-sequence. Then, by (1.9), $\xi \prec \zeta \prec \eta$. By A2, there exists a number $a \in K[0,1]$ and an $(a, 1-a)$-sequence $\nu \in \{0,1\}^{\mathbb{N}}$ such that $\zeta \sim \xi \oslash_\nu \eta$. By Lemma 1.3.1, $\xi \oslash_\nu \eta$ is an $(ap + (1-a)r)$-sequence. Thus, by (1.9) and $\zeta \sim \xi \oslash_\nu \eta$, $q \sim^P ap + (1-a)r$.

($\Leftarrow$) Suppose that $\precsim^P$ satisfies EU2. Suppose that ξ is a p-sequence, ζ is a q-sequence, and η is a r-sequence, and suppose that $\xi \prec \zeta \prec \eta$. Then, by (1.9), $p \prec^P q \prec^P r$. By EU2, there is a number $a \in K[0,1]$ such that $q \sim^P ap + (1-a)r$. Pick any $(a, 1-a)$-sequence ν, then $\xi \oslash_\nu \eta$ is an $(ap + (1-a)r)$-sequence. Hence, by (1.9) and $q \sim^P ap + (1-a)r$, $\eta \sim \xi \oslash_\nu \eta$.

(b.3) ($\Rightarrow$) Suppose that $\precsim$ satisfies A3. Recall that EU3 has two directions.

First we show that 'if' directions. Let $p, q, r \in \Delta(X)$, and let $a \in K[0,1]$ with $a \neq 0$. Suppose that $p \precsim^P q$. Let ξ be a p-sequence ξ and let ζ be a q-sequence. Then $\xi \precsim \zeta$. Pick any r-sequence η and pick any $(a, 1-a)$-sequence $\nu \in \{0,1\}^{\mathbb{N}}$. By Lemma 1.3.1, $\xi \oslash_\nu \eta$ is an $(ap+(1-a)r)$-sequence and $\zeta \oslash_\nu \eta$ is an $(aq + (1-a)r)$-sequence. By A3, $\xi \precsim \zeta$ implies that

$$\xi \oslash_\nu \eta \precsim \zeta \oslash_\nu \eta.$$

Thus, $ap + (1-a)r \precsim^P aq + (1-a)r$.

For the 'only if' direction, suppose that $ap + (1-a)r \precsim^P aq + (1-a)r$. Let ξ be a p-sequence, let ζ be a q-sequence, let η be a r-sequence, and let ν be an $(a, 1-a)$-sequence. By Lemma 1.3.1, $\xi \oslash_\nu \eta$ is an $(ap + (1-a)r)$-sequence and $\zeta \oslash_\nu \eta$ is an $(aq + (1-a)r)$-sequence. Hence, by (1.9), $\xi \oslash_\nu \eta \precsim \zeta \oslash_\nu \eta$. By A3, $\xi \precsim \zeta$, and so $p \precsim^P q$.

($\Leftarrow$) Suppose that $\precsim^P$ satisfies EU3. Again, A3 has two directions.

For the 'if' direction, suppose ξ is a p-sequence, ζ is a q-sequence, and η is a r-sequence. Also, suppose that $\xi \precsim \zeta$. Then, $p \precsim^P q$. For any $a \in K[0,1]$ with $a \neq 0$, using EU3, we have that $ap + (1-a)r \precsim^P aq + (1-a)r$. By Lemma 1.3.1, for any $(a, 1-a)$-sequences ν^1 and ν^2, $\xi \oslash_{\nu^1} \eta$ is an $(ap + (1-a)r)$-sequence and $\zeta \oslash_{\nu^2} \eta$ is an $(aq + (1-a)r)$-sequence. Therefore, we have $\xi \oslash_{\nu^1} \eta \precsim \zeta \oslash_{\nu^2} \eta$.

Consider the 'only if' direction. Suppose that ξ is a p-sequence, ζ is a q-sequence, and η is a

r-sequence, and suppose that ν^1 and ν^2 are $(a, 1-a)$-sequences with $a \in K[0,1]$ and $a \neq 0$. If $\xi \oslash_{\nu^1} \eta \precsim \zeta \oslash_{\nu^2} \eta$, it follows from Lemma 1.3.1 and (1.9) that $ap + (1-a)r \precsim^P aq + (1-a)r$. Then, by EU3, $p \precsim r$. Hence, we have $\xi \precsim \zeta$. $\square$

Proof of Theorem 1.3.2: There are two directions in this theorem. In part (a) of the proof, we show that 'if' direction, and in part (b) of the proof, we show the other direction.

(a) We show that existence of a representing utility function h of $\precsim$ implies that $\precsim$ satisfies A1-A3. Let $p, q \in \Delta(X)$. By Lemma 1.6.2, if ξ is a p-sequence, then $\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} = \sum_{x\in X} p_x h(x)$. Similarly, if ζ is a q-sequence, $\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\zeta_t)}{T} = \sum_{x\in X} q_x h(x)$. Therefore, by (1.10),

$$\xi \precsim \zeta \text{ if and only if } \sum_{x\in X} p_x h(x) \leq \sum_{x\in X} q_x h(x). \tag{1.22}$$

Now, define $\precsim^P$ over $\Delta(X)$ as $p \precsim^P q$ if and only if $\sum_{x\in X} p_x h(x) \leq \sum_{x\in X} q_x h(x)$. It is straightforward, using (1.22), to check that (1.9) holds for $\precsim$ and $\precsim^P$. By Theorem 1.2.1, it follows that $\precsim^P$ satisfies EU1-EU3. Then, by Theorem 1.3.1, we have that $\precsim$ satisfies A1-A3.

(b) Conversely, suppose that $\precsim$ satisfies A1-A3. By Theorem 1.3.1, there exists a preference relation $\precsim^P$ over $\Delta(X)$ that satisfies EU1-EU3 and satisfies condition (1.9). By Theorem 1.2.1, axioms EU1-EU3 hold if and only if there is a function $h : X \to K$ such that for all $p,\ q \in \Delta(X)$,

$$p \precsim^P q \text{ if and only if } \sum_{x\in X} p_x h(x) \leq \sum_{x\in X} q_x h(x).$$

By Lemma 1.6.2, if ξ is a p-sequence, then

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} = \sum_{x\in X} p_x h(x). \tag{1.23}$$

Therefore, for any p-sequence ξ and any q-sequence ζ,

$$\xi \precsim \zeta \text{ if and only if } p \precsim^P q \text{ if and only if } \sum_{x \in X} p_x h(x) \leq \sum_{x \in X} q_x h(x),$$

which, by (1.23), is equivalent to

$$\lim_{T \to \infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} \leq \lim_{T \to \infty} \sum_{t=0}^{T-1} \frac{h(\zeta_t)}{T}.$$

Thus, h is a representing utility function of $\precsim$. $\square$

Chapter 2

Does imperfect recognizability explain coexistence of money and higher-return assets?

2.1 Introduction

Over seventy years ago, Hicks [1935] asked: Why is money, which bears no interest, held when higher-return assets exist? One suggested answer is imperfect recognizability of higher-return assets (see, in particular, Freeman [1985], *et. al.* Lester et al. [2009], and Li and Rocheteau.). The idea that recognizability is a desirable property of a medium-of-exchange has a long history (see, for example, Jevons [1875][1]). Indeed, Alchain [1977] argues that the object with the smallest recognizability problem will emerge as the medium-of-exchange in equilibrium. Banerjee and Maskin [1996] formalize that idea in a model with competitive markets, and Williamson and Wright [1994] do so in a random matching model. Freeman [1985] addresses the Hicks problem directly.

[1] Jevons used the term identifiability to discuss recognizability.

He assumes that money is perfectly recognizable and that there is an extreme counterfeiting threat to higher-return assets (henceforth *bonds*). Lester et al. [2009] and Li and Rocheteau make similar assumptions in a random matching model. We reconsider their proposed solution to the Hicks problem.

Our reconsideration is concerned with two issues. First, in most of the above literature on imperfect recognizability, the analysis focuses on a certain type of pooling equilibrium: When the seller receives payment, he expects that it contains both genuine bonds and counterfeit bonds in the proportions in which they exist. However, starting with Cho and Kreps [1987], refinements have been proposed and most of them serve to eliminate some pooling equilibria. Second, in his solution to the Hicks problem, Freeman [1985] makes the extreme assumption that there is no cost to counterfeiting bonds. Our reconsideration studies robustness to the zero cost of counterfeiting, while at the same time invoking the refinement of Cho and Kreps, the so-called intuitive criterion.

We study these issues in a random matching model with two assets, money and bonds; money cannot be counterfeited, while bonds can be counterfeited. Aside from counterfeiting, the background setting resembles that in Zhu and Wallace [2007]. Our equilibrium concept is similar to that in Nosal and Wallace [2007], who study the counterfeiting of money. As they do, we model the pairwise meeting as a signaling game: The buyer makes a take-it-or-leave-it offer to the seller, with the seller not knowing the buyer's asset holding and, in particular, not knowing whether the buyer is offering genuine or counterfeit bonds. Our main finding is that with a sufficiently small but positive cost of counterfeiting, there is no monetary equilibrium. This result resembles what Nosal and Wallace [2007] find: in a model in which money can be counterfeited, the counterfeiting of money can be a threat to the monetary system.[2] Here, the counterfeiting of substitute media-of-exchange can be a threat to the monetary system.

[2]Nosal and Wallace [2007] seems to be the first paper on counterfeiting that applies the intuitive criterion. Other papers with private information about the quality of assets held use pooling equilibria without invoking it. See, for example, Green and Weber [1996], Velde et al. [1999], Williamson [2002], and Williamson and Wright [1994].

The reasoning is straightforward. First, under the assumptions made—and, quite plausibly—counterfeit bonds have value only if there is a pooling equilibrium in which prior to maturity they masquerade as genuine bonds in transactions. But, as shown here, there is no pooling equilibrium (that satisfies the intuitive criterion). Hence, there is no equilibrium with (costly) counterfeiting. Is there an equilibrium without counterfeiting? If there is, then, as might be expected, agents trade only with bonds because they bear interest. But, if the cost of counterfeiting bonds is sufficiently small, even for a small fraction of the population, then that cannot be an equilibrium. (There can, of course, be a non-monetary equilibrium.)

Those results suggest that a solution to the Hicks problem based on imperfect recognizability encounters difficulties. They point, however, to another solution: legal restrictions. In many countries, there are legal restrictions against the issue of assets that are intended to circulate as substitute media-of-exchange.[3] Our findings provide a new rationale for such restrictions. Given the threat of counterfeiting substitute of media-of-exchange, an effective prohibition against their use can restore monetary equilibrium.

2.2 The model

We study a random matching model with two divisible assets—fiat money and bonds. Instead of considering the general equilibrium of this model, we will focus on the partial equilibrium within one representative period, taking the continuation values of money as given. This approach suffices for our main result, a nonexistence result.

There are two stages in the period. Stage 1 has portfolios decisions, while stage 2 has pairwise

[3]For example, in Canada, such legal restriction is stated as follows: "Every bank or other person who issues, reissues, makes, draws or endorses any bill, bond, note, cheque or other instrument, intended to circulate as money or to be used as a substitute for money, is guilty of an offense against this Act." cited in Wallace [1988], p. 29.

meetings with production and consumption. There is a continuum of agents, each of whom begins the period with some amount of money, and the money holdings is distributed according to a distribution function $\eta \in \Delta(M)$ with $\eta(\{0\}) < 1$, where $M = [0, \overline{m}]$ is the set of possible money holdings. At the beginning of this period, half of the agents are chosen randomly to be buyers and half to be sellers and, in addition, each buyer has a technology that allows him to counterfeit bonds at cost $c \in C = [0, \infty)$ (with zero marginal cost). This cost is randomly distributed (independent of the wealth distribution) according to a distribution function $\mu \in \Delta(C)$.

At stage 1, the agents make portfolio choices: they can buy (genuine) discount bonds at price $p \in (0, 1]$ in terms of money. One unit of bond is a transferable claim to one unit of money at the end of the date. After observing the counterfeiting cost, buyers can choose to counterfeit. At the end of stage 1, each buyer is characterized by (m, b, t), where $m \in M$ is the money holding, $b \in B = [0, \overline{b}]$ is the (genuine) bond holding (measured at maturity value), and $t \in \{h, l\}$, where h (high) means that the person has not produced counterfeit bonds and l (low) means that the person has produced counterfeit bonds and has incurred cost c; each seller is characterized by his nominal wealth, denoted $z \in Z = [0, \overline{m} + \overline{b}]$. The stage-1 buyer decisions give rise to a distribution φ over $M \times B \times \{h, l\}$, the set of buyer types. Similarly, the stage-1 seller decisions give rise to a distribution π over Z, the set of seller types.

At stage 2, there are random pairwise meetings, with each meeting consisting of a buyer and a seller. We assume that the wealth of the seller is known, but that the buyer type is private information. In a meeting between a buyer of type (m, b, t) and a seller with wealth z, the buyer makes a take-it-or-leave-it offer, (x_m, x_b, y), where x_m and x_b denote the proposed money and bond transfers to the seller, respectively, and $y \in Y = [0, \infty)$ denotes the amount of good to be transferred to the buyer.

Regarding payoffs, the buyers have utility function $u(y)$, where y is the amount of good they consume, and the sellers have cost y, where y is the amount of good they produce. We use $V(w)$ to denote the expected continuation value to an agent at the end of that period if his wealth level is w. We assume that u is strictly increasing and concave, and that $u'(0) = \infty$. For the continuation value, we assume only that V is continuously differentiable with $V' > 0$ on $[0, \overline{m} + \bar{b}]$.

Given the continuation value and the buyer offer (x_m, x_b, y), the payoffs depend on the seller's response and the buyer's type. If the seller says *yes* and the buyer's type is h, then the seller's payoff is $-y + V(z + x_m + x_b)$ and the buyer's is $u(y) + V(m + b - x_m - x_b)$; if the seller says *yes* and the buyer's type is l, then the seller's payoff is $-y + V(z + x_m)$ (because counterfeits are worthless at maturity) and the buyer's is $u(y) + V(m + b - x_m) - c$; if the seller says no, then the seller's payoff is $V(z)$ and that of an h-type buyer is $u(0) + V(m + b)$, while that of an l-type buyer is $u(0) + V(m + b) - c$.[4]

We analyze the model by backward induction beginning with stage-2. Then, given conclusions about stage-2 equilibrium, we draw conclusions about equilibrium for both stages.

2.3 Equilibrium in the pairwise meetings

Each pairwise meeting can be modeled as an incomplete information game with the distributions of types given by φ and π. Because the wealth level of the seller is common knowledge in a meeting, each level z defines a proper subgame which has the form of a signaling game (buyers as senders with offers as messages and sellers as receivers). For a given z, a buyer's strategy is a function $s_z^{B2} = (s_{z,m}^{B2}, s_{z,b}^{B2}, s_{z,y}^{B2}) : M \times B \times \{h, l\} \rightarrow M \times B \times Y$, which assigns to each buyer type an offer (the 1st component is offer of money, the 2nd is offer of bonds, and the 3rd is output) and is subject

[4] Although a buyer's type has three components, only the third—whether the buyer holds counterfeit bonds or not—affects the seller's payoff. For this reason, the label type-t buyer is useful and means a buyer of type (m, b, t) for some portfolio (m, b).

to the feasibility constraints that limit asset transfers ($s_{z,m}^{B2}(m,b,t) \leq m$ and $s_{z,b}^{B2}(m,b,h) \leq b$). A seller *strategy* and *belief* is a function $(s_z^{S2}, \psi_z) : M \times B \times Y \rightarrow \{yes, no\} \times [0,1]$, which assigns to each buyer offer observed by the seller a response to the offer and a probability that the buyer's type is l.

The definition we use is in the spirit of perfect Bayesian equilibrium. It differs from the standard perfect Bayesian equilibrium because the set of strategies is infinite. The challenge in such cases is to specify a reasonable restriction on seller beliefs in equilibrium. Our approach is to restrict the support of the belief without imposing quantitative restrictions when Bayes rule is not applicable.

Definition 2.3.1. Given the distributions φ and π and the continuation value V, a *stage-2 equilibrium* is a triple (s^{B2}, s^{S2}, ψ), where $(s^{B2}, s^{S2}, \psi) = \{(s_z^{B2}, s_z^{S2}, \psi_z)\}_{z \in Z}$, and s_z^{B2} is a buyer strategy given seller wealth z and (s_z^{S2}, ψ_z) is a seller belief and strategy given seller wealth z, that satisfies the following conditions:

(i) for each $z \in \text{supp}(\pi)$ and each $(m,b,t) \in \text{supp}(\varphi)$, $s_z^{B2}(m,b,t)$ is a best response to s_z^{S2};

(ii) for each $z \in \text{supp}(\pi)$ and each buyer offer (x_m, x_b, y), $s_z^{S2}(x_m, x_b, y)$ is a best response to (x_m, x_b, y) given $\psi_z(x_m, x_b, y)$;

(iii) for each $z \in \text{supp}(\pi)$, if there is some $(m,b,t) \in \text{supp}(\varphi)$ such that $s_z^{B2}(m,b,t) = (x_m, x_b, y)$, then $\psi_z(x_m, x_b, y)$ is calculated via Bayse' rule; if Bayes' rule is not applicable, then

$$|\delta_t^h - \psi_z(x_m, x_b, y)| > 0 \text{ iff } \{(m,b,t) : s_z^{B2}(m,b,t) = (x_m, x_b, y)\} \cap \text{supp}(\varphi) \neq \emptyset,$$

where $\delta_h^h = 1$ and $\delta_l^h = 0$.

(iv) for each $z \in \text{supp}(\pi)$ and for each (x_m, x_b, y) such that $x_m + x_b \leq \sup\{m + b : (m,b,h) \in$

$\text{supp}(\varphi)\}$,

$$\psi_z(x_m, x_b, y) = 0 \text{ if } \varphi(M \times B \times \{l\}) = 0 \text{ and } \psi_z(x_m, x_b, y) = 1 \text{ if } \varphi(M \times B \times \{l\}) = 1.$$

Conditions (i) and (ii) are standard conditions of sequential optimality. Conditions (iii) and (iv) are mild restrictions on $\psi_z(x_m, x_b, y)$ for games with continuous strategy sets. In the discrete case, Bayes' rule requires that the belief assigns positive probability to the event that the buyer's type is t if the event that the observed offer comes from a buyer of type t has positive probability. In the continuous case, condition (iii) requires that the former event have positive probability if the later event has nonempty intersection with the support of the prior distribution, which is analogous to having positive probability in the discrete case. This condition also appears in Ramey [1996], who considers signalling games with continuous strategy sets. Condition (iv) requires that after observing any offer consistent with the wealth distribution, the belief assigns probability 1 (0) to the event that the buyer's type is t, if the prior probability of the buyer's being type t is 1 (0), given that such an offer is consistent with the wealth distribution. This requirement would be implied by sequential equilibrium if the game had finite strategy sets.

In addition to the restrictions in Definition 2.3.1, there is a refinement literature that imposes restrictions on beliefs for off-equilibrium actions in general signaling games. We follow most of the literature and adopt the intuitive criterion proposed by Cho and Kreps [1987]. Roughly speaking, an equilibrium of a signaling game fails this criterion if there is an off-equilibrium message and a sender type such that that type will receive a higher payoff by sending that off-equilibrium message, given that the receiver's belief is restricted to assign probability zero over sender types that will get a strictly lower payoff by sending that message. In the context of our game, a message is an offer. We say that an offer (x_m, x_b, y) is an equilibrium offer given seller's wealth level z if there is a buyer type $(m, b, t) \in \text{supp}(\varphi)$ such that $s_z^{B2}(m, b, t) = (x_m, x_b, y)$. An offer is an off-equilibrium offer given z if it is not an equilibrium offer given z.

We will show that this criterion serves to eliminate pooling equilibria of the following sort:

Definition 2.3.2. A *pooling equilibrium* (in bonds) is a stage-2 equilibrium (s^{B2}, s^{S2}, ψ) such that there is some wealth level $z \in \text{supp}(\pi)$ and some equilibrium offer (x_m, x_b, y) given z satisfying (i) $s_z^{S2}(x_m, x_b, y) = yes$; (ii) $0 < \psi_z(x_m, x_b, y) < 1$; and (iii) $x_b > 0$.

That is, in a pooling equilibrium, the seller accepts an equilibrium offer that includes bonds and assigns positive probability to both buyer types.

To show that a candidate equilibrium fails the intuitive criterion, it is sufficient to find a buyer type (m, b, h) and an off-equilibrium offer (x_m, x_b, y), given some seller wealth z, satisfying the following conditions:

(IC1) for any type-l buyer, offering (x_m, x_b, y) gives that buyer a strictly lower payoff than his equilibrium payoff regardless of the seller's response;

(IC2) for the buyer of type (m, b, h), offering (x_m, x_b, y) gives a strictly higher payoff than his equilibrium payoff, given that the seller's best response is conditioned on the belief that with probability 1 the buyer's type is h.

Given a candidate pooling equilibrium, we will show that there exists a deviating offer for a buyer type (m, b, h) satisfying the above conditions. The buyer type is the type that makes the pooling equilibrium offer mentioned in Definition 2.3.2. The offer has smaller proposed trades of both the good and bonds. Therefore, it can only come from a type-h buyer and, as a consequence, is accepted by the seller. The new part of the argument excludes the possibility that the deviating offer is offered in equilibrium by some other buyer. This possibility does not arise in models with only two types, as is the case in Nosal and Wallace [2007].

Theorem 2.3.1. *There is no pooling equilibrium that satisfies the intuitive criterion.*

Proof. Suppose that there exists such a stage-2 equilibrium (s^{B2}, s^{S2}, ψ), with $s_z^{B2}(m, b, h) = (x_m, x_b, y)$ satisfying conditions (i)-(iii) in Definition 2.3.2 with $z \in \text{supp}(\pi)$. Let $\psi_z(x_m, x_b, y) = \alpha \in (0, 1)$. Consider the buyer type (m, b, h) and the wealth level z. Because (x_m, x_b, y) is an equilibrium offer and $\alpha \in (0, 1)$, we can assume that $(m, b, h) \in \text{supp}(\varphi)$ by condition (iii) in Definition 2.3.1.

We first construct a deviating offer $(x_m, x_b - \epsilon, y - \epsilon')$ for the buyer type (m, b, h). Then we show that it satisfies (IC1) and (IC2), if it is an off-equilibrium offer. Finally, we verify that this offer is an off-equilibrium offer.

There exists $\epsilon > 0$ and $\epsilon' > 0$ such that

$$u(y - \epsilon') + V(m + b - x_m - x_b + \epsilon) > u(y) + V(m + b - x_m - x_b) \tag{2.1}$$

and

$$-y + \epsilon' + V(z + x_m + x_b - \epsilon) > -y + (1 - \alpha)V(z + x_m + x_b) + \alpha V(z + x_m) \tag{2.2}$$

hold.

Inequality (2.1) says that the offer $(x_m, x_b - \epsilon, y - \epsilon')$ gives higher payoff to the buyer of type (m, b, h), given that the seller accepts it. Inequality (2.2) implies that if the seller with wealth level z believes with probability 1 that the buyer's type is h after observing offer $(x_m, x_b - \epsilon, y - \epsilon')$, then his best response is to accept it. This shows that (IC2) is satisfied.

To find such ϵ and ϵ', first notice that inequality (2.2) is equivalent to $\epsilon' > -f(\epsilon)$, where

$$\begin{aligned} f(\rho) = {} & \alpha[V(z + x_m + x_b - \rho) - V(z + x_m)] \\ & -(1 - \alpha)[V(z + x_m + x_b) - V(z + x_m + x_b - \rho)]. \end{aligned}$$

Because

$$f(0) = \alpha[V(z + x_m + x_b) - V(z + x_m)] > 0,$$

and V is continuous and strictly increasing, we can choose $\epsilon > 0$ so that $f(\epsilon) > 0$. If so, then inequality (2.2) holds for any $\epsilon' \geq 0$. Then, for any such ϵ, choose $\epsilon' > 0$ so that

$$u(y) - u(y - \epsilon') < V(m + b - x_m - x_b + \epsilon) - V(m + b - x_m - x_b). \tag{2.3}$$

Because V is strictly increasing and u is continuous, such ϵ and ϵ' exists.

Now we verify (IC1) by showing that for any type-l buyer, offering $(x_m, x_b - \epsilon, y - \epsilon')$ leads to a strictly worse payoff than his equilibrium payoff. It is, of course, enough to deal with type-l buyers whose money holding is at least x_m. So consider an arbitrary type-l buyer; namely type (m', b', l), with $m' \geq x_m$ and $s_z^{B2}(m', b', l) = (x'_m, x'_b, y')$. We consider two cases.

Case (a): $m' > 0$. Because V is continuously differentiable with $V' > 0$ and $u'(0) = \infty$, there exists a feasible offer $(x''_m, 0, y'')$ satisfying

$$u(y'') + V(m' + b' - x''_m) > u(0) + V(m' + b') \tag{2.4}$$

and

$$-y'' + V(z + x''_m) > V(z). \tag{2.5}$$

That is, the seller gets a strictly higher payoff by accepting this offer than not doing so, regardless of his belief, and the buyer gets a strictly higher payoff by offering $(x''_m, 0, y'')$ than no trade. It follows that $s_z^{B2}(m', b', l) = (x'_m, x'_b, y')$ is accepted. Therefore, the equilibrium payoff is

$$\begin{aligned} u(y') + V_b(m' + b' - x'_m) &\geq \max\{u(y) + V_b(m' + b' - x_m), u(y'') + V_b(m' + b' - x''_m)\} \qquad (2.6) \\ &> \max\{u(y - \epsilon') + V_b(m' + b' - x_m), u(0) + V_b(m' + b')\}. \end{aligned}$$

Inequality (2.6) implies that such a type-l buyer receives a strictly lower payoff by offering $(x_m, x_b - \epsilon, y - \epsilon')$ (whether the seller accepts it or not) than the equilibrium payoff. (The first inequality in (2.6) follows from the fact that s is an equilibrium strategy and that the offers (x_m, x_b, y) and $(x''_m, 0, y'')$ are feasible for the buyer of type (m', b', l), while the second inequality in (2.6) follows from strict monotonicity of u and inequality (2.4).)

Case (b): $m' = 0$. Then, $x_m = 0$. In this case, using only strict monotonicity of u, $(0, x_b, y)$ is a strictly better offer (which is accepted by the seller) than $(0, x_b - \epsilon, y - \epsilon')$ (whether it is accepted or not).

Now we show that $(x_m, x_b - \epsilon, y - \epsilon')$ is an off-equilibrium offer. Because this offer gives any type-l buyer a strictly lower payoff than his equilibrium payoff, this cannot be an equilibrium offer from such a buyer. Thus, it suffices to show that $(x_m, x_b - \epsilon, y - \epsilon')$ is not an equilibrium offer made by any type-h buyer. Suppose it is. Because (x_m, x_b, y) is accepted by the seller, inequality (2.1) implies that buyer of type (m, b, h) obtains a higher payoff by making the offer $(x_m, x_b - \epsilon, y - \epsilon')$ provided that the seller accepts it. We have shown that this offer is not from a type-l buyer. This and condition (iii) in Definition 1 imply that $\psi_z(x_m, x_b - \epsilon, y - \epsilon') = 0$. Notice that condition (iii) only applies to equilibrium offers. Then (2.2) implies that the seller accepts the offer $(x_m, x_b - \epsilon, y - \epsilon')$. Hence, the buyer of type (m, b, h) gets a strictly higher payoff by offering $(x_m, x_b - \epsilon, y - \epsilon')$ than his equilibrium payoff. This contradicts the assumption that (s^{B2}, s^{S2}, ψ), in which (x_m, x_b, y) is an equilibrium offer by a type-h buyer, is an equilibrium. Therefore, $(x_m, x_b - \epsilon, y - \epsilon')$ is an off-equilibrium offer, and this completes the proof. $\square$

Theorem 2.3.1 shows that no pooling equilibrium exists at stage 2. In the next section, we show that this result implies no counterfeiting in equilibrium unless there are people with zero counterfeiting cost.

2.4 Equilibrium

An equilibrium of both stages consists of actions in both stages. We have discussed agents' strategies at stage 2 in the last section. Now we turn to strategies at stage 1. Recall that each agent begins the period with some money according to the distribution function $\eta \in \Delta(M)$. Each agent is assigned to be a buyer or a seller. The sellers then decide their bond holdings. The buyers, after observing their counterfeiting costs, decide their bond holding and whether to counterfeit or not. A buyer's strategy is a function $s^{B1} = (s_b^{B1}, s_c^{B1}) : M \times C \to B \times \{h, l\}$, where the 1st component is the bond holding and the 2nd is the counterfeiting decision with h indicating not to counterfeit and l indicating to counterfeit, with the constraints that $s_b^{B1}(m, c) \leq \frac{m}{p}$. A seller's strategy is a function $s^{S1} : M \to B$ such that $s^{S1}(m) \leq \frac{m}{p}$.

We impose the following two requirements for an equilibrium $(s^{B1}, s^{S1}, s^{B2}, s^{S2})$:

(E1) At stage 2, the buyer's and seller's actions (s^{B2}, s^{S2}) consist of a stage-2 equilibrium for some seller belief ψ, with the distributions φ and π calculated from η, μ and s^{B1}, s^{S1};

(E2) At stage 1, s^{B1} and s^{S1} are optimal given the equilibrium actions (s^{B2}, s^{S2}) at stage 2.

We show that if the cost of counterfeiting is positive for everyone (that is, $\mu(c > 0) = 1$), then there is no counterfeiting in any equilibrium in which the stage-2 equilibrium satisfies the intuitive criterion. Based on Theorem 2.3.1, it is sufficient to show that counterfeiting implies pooling (see Theorem 2.4.1). The proof of Theorem 2.4.1, which appears in the appendix, would be easy if the type space were finite, as would be the case if money and bonds came in indivisible units.[5] Obviously, the existence of counterfeiting implies that there are type-l buyers making the same offers as some type-h buyers, but the challenge with an infinite type-space is the possibility that the seller may take this as a negligible event and assign probability zero to it. This difficulty is

[5]If money and bonds were indivisible, then it would be necessary to introduce lotteries, as is done in Nosal and Wallace [2007].

resolved in the proof via a measure-theoretic argument.

Theorem 2.4.1. *Suppose that $\mu(c > 0) = 1$. There is counterfeiting in an equilibrium only if the stage-2 equilibrium is a pooling equilibrium.*

Theorem 2.3.1 and Theorem 2.4.1 imply that if $\mu(c > 0) = 1$, then there is no counterfeiting in equilibrium. The remaining question is whether there is any equilibrium. We structure the discussion in terms of three cases for μ and p.

Case 1: $p < 1$ and $\mu(c > 0) = 1$.

Because there is no counterfeiting, in any equilibrium bonds offered at stage 2 are regarded as genuine—i.e., $\varphi(M \times B \times \{h\}) = 1$. It follows that no one leaves stage 1 with money and that all stage-2 trade is conducted with bonds. However, the existence of equilibrium depends on the distribution of costs, μ. The following theorem gives a sufficient condition for nonexistence.

Theorem 2.4.2. *If $\mu(c > 0) = 1$ and if for any $\varepsilon > 0$, $\mu(c \in [0, \varepsilon]) > 0$, then there is no equilibrium.*

Proof. Suppose that, by contradiction, there is an equilibrium $(s^{B1}, s^{S1}, s^{B2}, s^{S2})$ with seller belief ψ in the stage-2 equilibrium. Because there is no counterfeiting, i.e., $\varphi(M \times B \times \{h\}) = 1$, by condition (iv) in Definition 2.3.1, for any (x_m, x_b, y) such that $x_m + x_b \leq \sup\{m + b : (m, b, h) \in \operatorname{supp}(\varphi)\}$, $\psi_z(x_m, x_b, y) = 0$. This implies that holding a positive amount of money in the portfolio choice at stage 1 is a strictly dominated strategy for any agent because $p < 1$ and V is strictly increasing.

Thus, each buyer is characterized by his bond holding b. Because V is strictly increasing and continuously differentiable, and $u'(0) = \infty$, each buyer with positive bond holding will make an offer $(0, x_b, y)$ with $x_b > 0$ and $y > 0$ that is accepted by the seller, no matter what wealth level the seller has. Now, because $\eta(\{0\}) < 1$ and so $\varphi(\{0\} \times \{0\} \times \{h\}) < 1$, there is a positive measure of buyers with positive bond holdings.

For each possible bond holding $b > 0$, define

$$G(b) =_{def} \int_{z \in Z} [V(b) - V(b - x_{b,z})] d\pi(z),$$

where $s_z^{B2}(0, b, h) = (0, x_{b,z}, y_z)$. $G(b)$ is the gross gain from counterfeiting for a buyer with bond holding b. As we have seen, for any $b > 0$, $x_{b,z} > 0$ and so $G(b) > 0$. It then follows that

$$\varphi(\{(0, b, h) : G(b) > 0\}) = \varphi(\{(0, b, h) : b > 0\}) > 0. \tag{2.7}$$

If for all $n \in \mathbb{N}$, $\varphi(\{(0, b, h) : G(b) \leq \frac{1}{n}\}) = 1$, it follows that

$$\varphi(\{(0, b, h) : G(b) = 0\}) = \varphi(\bigcap_{n=1}^{\infty} \{(0, b, h) : G(b) \leq \frac{1}{n}\}) = \lim_{n \to \infty} \varphi(\{(0, b, h) : G(b) \leq \frac{1}{n}\}) = 1,$$

which contradicts (2.7). Hence, there is some $n \in \mathbb{N}$ such that

$$\varphi(\{(0, b, h) : G(b) > \frac{1}{n}\}) > 0.$$

Take $\varepsilon = \frac{1}{2n}$. For any buyer with bond holdings b such that $G(b) > \frac{1}{n}$ and with counterfeiting cost $c \leq \varepsilon$, counterfeiting is a strictly better choice than not counterfeiting. Now, the measure of such buyers is

$$\mu(c \in [0, \varepsilon])\varphi(\{(0, b, h) : G(b) > \frac{1}{n}\}) > 0.$$

Hence, these buyers counterfeit, which contradicts the fact that there is no counterfeiting. $\square$

Although Theorem 2.4.2 is similar to the finding of Nosal and Wallace [2007], our proof is quite different from theirs. In their setup, money holdings are not divisible and can only be either 0 or 1. Moreover, only stationary equilibria are considered there. Hence, trades are determined in any monetary equilibrium. In contrast, however, we assume only that trades happen in pairwise meetings in an equilibrium.

It follows from Theorem 2.4.2 that there is no equilibrium unless c is sufficiently large across all agents. Although we have focused on partial equilibrium, the partial equilibrium conditions, (E1) and (E2), are necessary for any general equilibrium of this model. Hence, we have shown that there is no monetary equilibrium (in the general equilibrium model) with continuation values satisfying our regularity conditions. In that sense, the counterfeiting of bonds threatens the monetary system.

Case 2: $p = 1$.

In this case, for any distribution of counterfeiting costs, there is an equilibrium in which no one counterfeits and bonds are ignored (none are purchased at stage 1).[6] This case does not deal with coexistence, because the bonds do not bear interest, but it may still be relevant. One interpretation is that both assets are currencies: one not subject to a counterfeiting threat and the other subject to such a threat.

Case 3: $\mu(c = 0) > 0$.

In this case, Theorem 2.3.1 holds, but Theorem 2.4.1 does not. Hence, there may be an equilibrium in which those buyers with zero cost of counterfeiting counterfeit, the counterfeits are not traded, and only money is used at stage 2. However, $\mu(c = 0) > 0$ is a special and unappealing case.

[6] See Zhu [2005] for a general equilibrium existence result where bonds are ignored. Although, there, both the buyer's and the seller's money holdings are observable, any equilibrium there corresponds to an equilibrium here because the seller cares only the trade and the money holding of the buyer has no effect on the seller's payoff.

2.5 Concluding remarks

So far, nothing has been said about who offers bonds or the determination of p. Governments rarely offer securities which can easily substitute for their monies. Therefore, let's consider the possibility that the transferable bonds are offered by intermediaries. In particular, suppose that the government at stage 1 makes available one-period discount bonds that are *not* suitable to be traded in pairwise meetings—perhaps, because they are in large denominations or are book-entry securities. The transferable bonds in the model could arise from the following intermediation activity: An intermediary holds as assets bonds issued by the government and offers one-period transferable bonds that are designed to be traded in pairwise meetings. If $p' < 1$ is the price at which the government offers its bonds, then the intermediary can offer its transferable bonds at $p \in (p', 1)$ and earn revenue proportional to $p - p'$. But if such intermediary bonds are subject to being counterfeited, then all the results above apply. In particular, unless the cost of counterfeiting the intermediary bonds is sufficiently high, there is no equilibrium with such intermediation.

This does not, of course, imply that there is an equilibrium in which no one chooses to enter the intermediation business. It does, however, imply that a legal restriction that prohibits such intermediation does not eliminate equilibria and is consistent with a monetary equilibrium. In that sense, the above analysis provides a new rationale for such a legal restriction.

We make many assumptions to obtain our results. Some assumptions are mainly made for expositional purposes, while others are more important. The first group includes almost everything about the environment. For example, the timing of the assignment of people to buyer and seller status is not essential. Nor is the assumption that every meeting is between a buyer and seller. The assumption that there is only one round of trade prior to maturity can be relaxed to allow for any finite number of rounds of such trade and, perhaps, for an infinite number. And the nonexistence of a pooling equilibrium depends only on the assumption that ending trade with a counterfeit bond is worth less than ending with a genuine bond. Nor do the results depend on the assumption that the marginal cost of producing counterfeits is zero. Finally, the assumption that the seller types are common knowledge can also be dropped. If seller wealth is private information, then a buyer offer consists of a menu, each item of which is intended for a different seller. All the analysis above can be modified to permit this and we suspect that all the main results still hold.

The more significant assumptions are about the game played and the equilibrium concept—in particular, the formulation of the stage-2 game as a signaling game and the use of the intuitive criterion. The timing of counterfeiting decisions seems crucial as well: We use the assumption that counterfeiting happens before the pairwise meetings to obtain our results. See Lester et al. [2009] and Li and Rocheteau for an alternative timing. Obviously, the nonexistence result holds for refinements of the intuitive criterion, but may not hold for alternatives to it.[7] One route to generalization beyond the signaling game formulation is to use a notion of the (pairwise) core under asymmetric information. However, as yet, there is no widely accepted notion of that core.

2.6 Appendix

Proof of Theorem 2.4.1: Suppose that there is counterfeiting in equilibrium; i.e., $q_l = \varphi(M \times B \times \{l\}) > 0$. Notice that $q_l < 1$ in equilibrium. Let φ_t denote the conditional distribution of φ over $M \times B \times \{t\}$, $t = h, l$. For any buyer type (m, b, l) that is in the support of φ, define

$$A_{m,b} = \{z \in \text{supp}(\pi) : s_{z,b}^{B2}(m, b, l) > 0,\ \psi_z(s_z^{B2}(m, b, l)) < 1, s_z^{S2}(s_z^{B2}(m, b, l)) = yes\}.$$

This is the set of seller types such that the buyer of type (m, b, l) offers a positive amount of bonds to them, they assign positive probability that the offer comes from a type-h buyer, and

[7] One alternative is Mailath et al. [1993]. For an epistemic justification for the intuitive criterion in terms of rationality, see Battigalli and Siniscalchi [2002].

they respond with *yes*.

If there is some $(m,b) \in \text{supp}(\varphi_l)$, and some $z \in A_{m,b}$ with $\psi_z(s_z^{B2}(m,b,l)) > 0$, then (s^{B2}, s^{S2}, ψ) is a pooling equilibrium. So assume that for all $(m,b) \in \text{supp}(\varphi_l)$, and for all $z \in A_{m,b}$, $\psi_z(s_z^{B2}(m,b,l)) = 0$. Now we show that this assumption leads to a contradiction.

The above assumption implies, using condition (iii) in Definition 2.3.1, that for all $(m,b) \in \text{supp}(\varphi_l)$ and for all $z \in A_{m,b}$,

$$\varphi_h(\{(m',b') : s_z^{B2}(m',b',h) = s_z^{B2}(m,b,l)\}) > 0, \tag{2.8}$$

and

$$\varphi_l(\{(m',b') : s_z^{B2}(m',b',l) = s_z^{B2}(m,b,l)\}) = 0. \tag{2.9}$$

To see this, clearly if (2.9) does not hold, $\psi_z(s_z^{B2}(m,b,l)) > 0$; if (2.9) holds but not (2.8), then the Bayes' rule is not applicable and so condition (iii) in Definition 2.3.1 implies that $\psi_z(s_z^{B2}(m,b,l)) > 0$.

Let

$$T = \{(m,b,z) : (m,b) \in \text{supp}(\varphi_l), z \in A_{m,b}\}.$$

This consists of portfolio types of type-l buyers and wealth levels of sellers such that counterfeits are offered and accepted.

There are two valid ways to calculate the measure of the set T. We will show that it is zero by one method and positive by the other. That is a contradiction. Let $\varphi_l \otimes \pi$ denote the product measure of φ_l and π over $(M \times B) \times Z$.

First, it can be calculated via the sets X_z, which are defined as

$$X_z = \{(x_m, x_b, y) : (x_m, x_b, y) = s_z^{B2}(m, b, l) \text{ for some } (m, b) \text{ such that } z \in A_{m,b}\},$$

where $z \in A = \bigcup_{(m,b)\in\text{supp}(\varphi_l)} A_{m,b}$. The set X_z consists of equilibrium buyer offers to sellers with wealth z that include counterfeit bonds.

It is easy to see that

$$T \subseteq \{(m, b, z) : s_z^{B2}(m, b, l) \in X_z, z \in A\} = U.$$

Now, the measure of U can be computed as follows: First compute the measure of the set $\{(m, b) : s_z^{B2}(m, b, l) \in X_z\}$ for each $z \in A$, then integrate these measures over z. Inequality (2.8) implies that for each offer in X_z, the set of portfolio types of type-h buyers who offer it has positive measure—i.e., for each $z \in A$ and for each $(x_m, x_b, y) \in X_z$, $\varphi_h(B^z_{x_m,x_b,y}) > 0$, where

$$B^z_{x_m,x_b,y} = \{(m, b) : s_z^{B2}(m, b, h) = (x_m, x_b, y)\}.$$

The set $B^z_{x_m,x_b,y}$ consists of type-h buyers who make the offer (x_m, x_b, y) to sellers with wealth z. For different offers (x_m, x_b, y) and (x'_m, x'_b, y') in X_z, the sets $B^z_{x_m,x_b,y}$ and $B^z_{x'_m,x'_b,y'}$ are disjoint. Because

$$\varphi_h\Big(\bigcup_{(x_m,x_b,y)\in X_z} B^z_{x_m,x_b,y}\Big) \leq 1 \text{ and } \varphi_h(B^z_{x_m,x_b,y}) > 0 \text{ for all } (x_m, x_b, y) \in X_z,$$

the set X_z is at most countably infinite.[8]

[8]Formally, this follows from Billingsley [1995], Theorem 10.2 (iv).

Now, (2.9) implies that for each $z \in A$ and for each $(x_m, x_b, y) \in X_z$,

$$\varphi_l(\{(m,b) : s_z^{B2}(m,b,l) = (x_m, x_b, y)\}) = 0.$$

Hence,

$$\begin{aligned}(\varphi_l \otimes \pi)(T) \leq (\varphi_l \otimes \pi)(U) &= \int_{z\in A} \varphi_l(\{(m,b) : s_z^{B2}(m,b,l) \in X_z\}) d\pi(z) \\ &= \int_{z\in A} \sum_{(x_m,x_b,y)\in X_z} \varphi_l(\{(m,b) : s_z^{B2}(m,b,l) = (x_m,x_b,y)\}) d\pi(z) = 0.\end{aligned} \tag{2.10}$$

Second, it can be calculated directly. We first claim that $\pi(A_{m,b}) > 0$ for all type-(m,b,l) buyer with positive counterfeiting cost and $(m,b) \in \mathrm{supp}(\varphi_l)$. Because $\mu(c > 0) = 1$, this implies that $\pi(A_{m,b}) > 0$ for almost all $(m,b) \in \mathrm{supp}(\varphi_l)$. Given this claim, we have

$$(\varphi_l \otimes \pi)(T) = \int_{(m,b)\in \mathrm{supp}(\varphi_l)} \pi(A_{m,b}) d\varphi_l(m,b) > 0. \tag{2.11}$$

Now, (2.10) and (2.11) contradict each other.

To show that $\pi(A_{m,b}) > 0$ for all type-(m,b,l) buyer with positive counterfeiting cost and $(m,b) \in \mathrm{supp}(\varphi_l)$, first notice that for each $z \notin A_{m,b}$, there are two possibilities:

(1) $z \notin \mathrm{supp}(\pi)$;

(2) $z \in \mathrm{supp}(\pi)$, divided into three subcases:

(2.a) the seller rejects the offer $s_z^{B2}(m,b,l)$;

(2.b) the seller accepts the offer and $s_{z,b}^{B2}(m,b,l) = 0$;

(2.c) $\psi_z(s_z^{B2}(m,b,l)) = 1$ and the seller accepts the offer.

If $\pi(A_{m,b}) = 0$, consider changing the offer of the buyer of type (m, b, l) to the following: For z's not in the support, or for z's in the support but belonging to cases (2.a) or (2.b), no change; for z's in the support and belonging to case (2.c), change the offer to $(s^{B2}_{z,m}(m, b, l), 0, s^{B2}_{z,y}(m, b, l))$. The new offer will be accepted if and only if the original one is accepted (recall that ψ_z is the probability that the buyer holds counterfeit bonds).

By not counterfeiting and giving this new offer, the buyer of type (m, b, l) gets exactly the same expected payoff in the pairwise meetings without occurring the cost of counterfeiting, which is positive. This profitable deviation cannot happen because (s^{B2}, s^{S2}, ψ) is a stage-2 equilibrium. Hence, $\pi(A_{m,b}) > 0$. $\square$

Chapter 3

Complexity and Mixed Strategy Equilibria

3.1 Introduction

Unpredictable behavior is central for optimal play in many strategic situations because a predictable pattern leaves a player vulnerable to exploitation—think of the direction of tennis serves or soccer penalty kicks or the pattern of bluffing in poker. In fact, Walker and Wooders [2001] for tennis and Palacios-Huerta [2003] for soccer find evidence that supports such unpredictable behavior. In this chapter, we propose a theory of unpredictable behavior in the context of (infinitely) repeated two-person zero-sum games in which the stage games have no pure strategy equilibrium. We focus on repeated play because equilibrium makes no predictions about a single play (von Neumann and Morgenstern [1944], p. 147, expresses a similar concern). Recognizing this fact, there is a literature which identifies mixed strategies with beliefs and which makes predictions about *beliefs* in one-shot games (see, for example, Harsanyi [1973] and Aumman and Brandenburger [1995]). In contrast, our theory, based on complexity considerations, has predictions about players' *actions*. For example, in a repeated matching pennies game, the sequence of plays that alternates between heads and tails is not an equilibrium play in our theory—even though it could have resulted from an i.i.d. random process (and, in fact, is no more likely or unlikely than any other sequence). Moreover, our theory has implications for empirical tests of unpredictable behavior. We find that the failure of some tests for randomness does not justify rejection of equilibrium play.

We endow each player a set of feasible functions that the player can use to implement strategies, and measure complexity using the computability relation over functions. Roughly speaking, a function is computable from another if the former can be obtained from the latter through a mechanical process. We assume that, first, if a function is feasible to a player, so is any function computable from it. Our framework then has three ingredients: a finite zero-sum game as the stage game, and a set of feasible functions for each player. For payoffs, we adopt the long-run average criterion, which we will argue is more appropriate in this context, and we show that there is no equilibrium with the discounting criterion. Our first result is a *necessary* condition for equilibrium unpredictable behavior to exist. If the stage game has no pure strategy equilibrium, then, to obtain equilibrium in the repeated game with the sets of feasible functions that has the same equilibrium payoffs as the stage game, it is necessary that each player has a feasible function that is not feasible to the other player. One corollary of this result is that if such equilibrium exists, the equilibrium strategy of one player is not computable by the other.

To obtain a *sufficient* condition for existence, we use Kolmogorov complexity (Kolmogorov [1965]) to consider another dimension of complexity. This complexity measures the minimal description length of a finite object, using a function as a language. It is used here to understand how uncomputable a function (which can be identified with an infinite sequence) is relative to a set of feasible functions. It is known that a *complex sequence*—i.e., a sequence such that the Kolmogorov complexities (relative to a set of feasible functions) of its initial segments are essentially the lengths of those segments—represents a highly uncomputable function, and it can be thought of as the limit of finite sequences that are hard to compute. Extending this idea in our two-player setting, we give a new concept called *mutual complexity*: It assumes that each player has a complex sequence relative to the other player's set of feasible functions. Adding another assumption that there exists a black box that generates all feasible functions within a finite process for each player, mutual complexity guarantees existence of an equilibrium with equilibrium payoffs equal to those of the stage game. Moreover, for any mixed strategy equilibrium of the stage game, there is a corresponding equilibrium in the repeated game.

Such an existence cannot be obtained with the complexity notion in the machine game literature (see Ben-Porath [1993] and Osborne and Rubinstein [1994]), which uses finite automata to implement strategies and measures complexity with the number of states in a player's automaton. The existence result in that literature still relies on the use of mixed strategies. We overcome this difficulty by considering uncomputable sequences (which are necessary for equilibrium) relative to the other player.

Although it has some implications on equilibrium strategies, the existence result is not very informative about the meaning of unpredictability in our framework. In particular, there is an empirical literature which rejects or accepts equilibrium hypothesis in repeated zero-sum games by employing statistical tests on observed behavior (see, for example, Brown and Rosenthal [1990], Walker and Wooders [2001], and Palacios-Huerta [2003]; however, O'Neill [1991] doubts the relevance of all such tests to reject the equilibrium hypothesis), and thus identifies the meaning of unpredictability with passing all such tests with respect to the i.i.d. distribution generated by equilibrium mixed strategies. But is that the meaning of unpredictability in our model? To answer this question, we follow Martin-Löf [1966] to define idealized statistical tests relative to a set of feasible functions, which are based on properties with zero probability that can be detected by those functions. The answer is then *no*: Under mutual complexity, there always exists an equilibrium strategy that fails the Law of the Iterated Logarithm; moreover, to have a strategy that passes all such tests for both players, mutual complexity must hold. We are not able to fully characterize the meaning of unpredictability, but have identified a criterion called *stochasticity* that is very close to it. This notion would help to understand the empirical implications of our model.

The rest of the chapter is organized as follows: In Section 3.2 we formulate two repeated games: horizontal game and vertical games, and present nonexistence results; Section 3.3 has two parts: first we formulate the notion of complex sequences using Kolmogorov complexity and give an existence result; then we discuss the meaning of unpredictability in our model; in section 3.4 we clarify our results for vertical games; we give some discussions of our results and further research in Section 3.5; the proofs of the main theorems are in Section 3.6.

3.2 Repeated games with feasible functions

We give two formulations of the repeated game, which are called the *horizontal game* (HG) and the *vertical game* (VG), respectively. Both games consist of infinite repetitions of a finite zero-sum game, but they differ in their information structures. In both games each player is endowed with a set of feasible functions to implementable his strategies. We use computability theory (c.f. Odifreddi [1989]) to formulate those feasible functions, and give a nonexistence result in the end of this section.

3.2.1 Feasible functions

In our approach, players do not have access to all strategies in the repeated game. Instead of restricting attention to an arbitrary subset of strategies, we endow a player a set of feasible functions for him to implement his strategies. Because we only consider finite stage games, it is sufficient to consider functions over natural numbers. However, we include also *partial functions*—i.e., functions that may be undefined for some inputs—over vectors of natural numbers. We include those functions because many complicated functions with single variable are derived from more basic partial functions with more than one variable (for example, the function $f(n) = \max\{k : 2^k|n\}$ is derived from division, which is a partial function over pairs of natural numbers).

Formally, the set of all partial functions is denoted by

$$\mathcal{F} = \bigcup_{k=1}^{\infty}\{f : \overline{\mathbb{N}}^k \to \overline{\mathbb{N}} \mid \text{ if for some } i,\ n_i = \perp,\ f(n_1, ..., n_k) = \perp\} = \bigcup_{k=1}^{\infty}\mathcal{F}^k,$$

where $\overline{\mathbb{N}} = \mathbb{N} \cup \{\perp\}$; $\mathbb{N}$ is the set of natural numbers and the symbol $\perp$ means that the function is not defined there. We give some related notations here: The *domain* of a partial function f, denoted by $dom(f)$, is the set of elements in $\mathbb{N}^k$ over which f is defined. A function $f \in \mathcal{F}^k$ is *total* if $dom(f) = \mathbb{N}^k$. For any set $\mathcal{P} \subset \mathcal{F}$, we use $\mathcal{P}_T$ to denote the set of total functions in $\mathcal{P}$. Finally, two partial functions f and g are *congruent*, denoted by $f \simeq g$, if $dom(f) = dom(g)$ and $f(n_1, ..., n_k) = g(n_1, ..., n_k)$ for all $(n_1, ..., n_k) \in dom(f)$.

Through out this chapter we assume that if a function is feasible to a player, so is any function less complicated than that function. This assumption is natural for any theory of rational behavior. We use the computability relation, as developed in the Recursion theory (cf. Odifreddi [1989]), to formulate this assumption formally. Our main result needs other assumptions, and we will discuss those after we give our basic model.

The computability relation indicates whether a function is computable from another. This relation measures complexity qualitatively—a function is more complicated than another if the latter is computable from the former. This relation is formulated based on the idea that any player is able to compute certain basic functions and perform certain basic operations, and so if a function can be transformed into another via these basic operations, possibly with the aid of basic functions, then the latter is computable from the former. The basic functions include (1) the zero function Z defined by $Z(n) = 0$; (2) the successor function S defined by $S(n) = n + 1$;

(3) the projection function U_{ki} defined by $U_{ki}(n_1, ..., n_k) = n_i$, for all $k > 0,\ i = 1, .., k$. The basic operations include

(1) composition: from functions $f(m_1, ..., m_l)$ and $g_1(n_1, ..., n_k), ..., g_l(n_1, ..., n_k)$ obtain

$$h \simeq f(g_1(n_1, ..., n_k), ..., g_l(n_1, ..., n_k));$$

(2) primitive recursion: from functions f and g obtain h defined as

$$\text{for } m = 0,\ h(0, n_1, ..., n_k) \simeq f(n_1, ..., n_k);$$
$$\text{for } m > 0,\ h(m, n_1, ..., n_k) \simeq g(m-1, h(m-1, n_1, ..., n_k), n_1, ..., n_k).$$

(3) minimization: from function f obtain g defined as

$$g(m, n_1, ..., n_k) = \min\{n : f(n, n_1, ..., n_k) = 0$$
$$\text{and } f(l, n_1, ..., n_k) \text{ is defined and is positive for all } l < n\} \text{ if such minimum exists;}$$
$$g(m, n_1, ..., n_k) \text{ is undefined otherwise.}$$

The computability relation is defined as follows.

Definition 3.2.1. *A function $f \in \mathcal{F}$ is computable from a function $g \in \mathcal{F}$* if there is a sequence of functions in $\mathcal{F}$ $f_1, ..., f_n$ such that $f_1 = g,\ f_n = f$, and for each $i = 2, ..., n,\ f_i$ is either a basic function or is obtained from functions in $f_1, ..., f_{i-1}$ through a basic operation.

The following lemma shows that the computability is a preorder, but is *not* complete. This fact will be used later on.

Lemma 3.2.1. *(a) The computability relation is reflexive and transitive.*

(b) There exist total functions $f,\ g \in \mathcal{F}^1$ such that f is not computable from g and g is not computable from f.

Proof. (a) It is clear that this relation is reflexive. For transitivity, suppose that f is computable from g and g is computable from h. Then, there are two sequences $f_1, ..., f_n$ and $g_1, ..., g_m$ such that $f_1 = g$, $f_n = f$, $g_1 = h$, $g_m = g$ and they satisfy the condition in Definition 1. Consider then the sequence $g_1, ..., g_m, f_2, ..., f_n$. Then it satisfies the condition in Definition 1 and shows that f is computable from h.

(b) See Odifreddi [1989]. □

Now we are ready to formulate our assumption on the set of feasible functions $\mathcal{P} \subset \mathcal{F}$ to a player:

Assumption A1 If $f \in \mathcal{P}$ and if g is computable from f, then $g \in \mathcal{P}$.

Before we formulate our model, several remarks are necessary. First, any finite set $X = \{x_1, ..., x_m\}$ can be identified with a subset of natural numbers $\{1, ..., m\}$ (of course, there are alternative identifications), and the set of finite sequences over X, denoted by $X^{<\mathbb{N}}$, can be identified with $\mathbb{N}$ using the mapping $\sigma \mapsto \sum_{t=0}^{|\sigma|-1}(\sigma_t)m^t$ (where x_i is identified with i, $i = 1, ..., m$). Hence, any function from $\mathbb{N}$ to X, or equivalently, any infinite sequence over X, can be identified with a function in $\mathcal{F}$; similarly, any function from $X^{<\mathbb{N}}$ to a finite set Y can be identified with a function in $\mathcal{F}$ as well. If a set $\mathcal{P}$ satisfies (A1), then whether such functions belong to $\mathcal{P}$ or nor does not depend on what identification we use.

3.2.2 Horizontal and vertical games

Here we formulate our model. The stage game is a finite zero-sum two-person game $g = \langle X, Y, h\rangle$, where $X = \{x_1, ..., x_m\}$ is the set of actions for player 1, $Y = \{y_1, ..., y_n\}$ is the set of actions for player 2, and $h : X \times Y \to \mathbb{Q}$ is the von Neumann-Morgenstern utility function for player 1. We use

$$\Delta(X) = \{p \in [0,1]^m : p \in \mathbb{Q},\ \sum_{x \in X} p_x = 1\}$$

and

$$\Delta(Y) = \{q \in [0,1]^n : q \in \mathbb{Q},\ \sum_{y \in Y} q_y = 1\}$$

to denote the set of mixed strategies (with rational probability values) for player 1 and 2, respectively. Because h is rational-valued, there is always a mixed strategy equilibrium of g with rational probability values. We use $X^{<\mathbb{N}}$ ($Y^{<\mathbb{N}}$) to denote the set of finite sequences over X (Y). For any sequence $\xi \in X^{\mathbb{N}}$, we use $\xi[T]$ to denote its initial segment with length T, i.e., $\xi[T] = (\xi_0, \xi_1, ..., \xi_{T-1})$. Given the stage game g, we formulate the vertical and the horizontal games as follows:

Definition 3.2.2. Let $g = \langle X, Y, h\rangle$ be a finite zero-sum game and let $\mathcal{P}^1$, $\mathcal{P}^2$ be two sets of functions satisfying (A1). The *vertical game* $VG(g, \mathcal{P}^1, \mathcal{P}^2)$ is a triple $\langle \mathcal{X}, \mathcal{Y}, u_h\rangle$:

(a) $\mathcal{X} = \{a : Y^{<\mathbb{N}} \to X : a \in \mathcal{P}^1_T\}$ is the set of strategies for player 1;

(b) $\mathcal{Y} = \{b : X^{<\mathbb{N}} \to Y : b \in \mathcal{P}^2_T\}$ is the set of strategies for player 2;

(c) $u_h : X^{\mathbb{N}} \times Y^{\mathbb{N}} \to \mathbb{R}$ is player 1's payoff function defined as

$$u_h(\xi, \zeta) = \liminf_{T \to \infty} \sum_{t=0}^{T-1} \frac{h(\xi_t, \zeta_t)}{T}; \tag{3.1}$$

player 2's payoff function is $-u_h$.

Definition 3.2.3. Let $g = \langle X, Y, h\rangle$ be a finite zero-sum game and let $\mathcal{P}^1$, $\mathcal{P}^2$ be two sets of functions satisfying (A1). The *horizontal game* $HG(g, \mathcal{P}^1, \mathcal{P}^2)$ is a triple $\langle \mathcal{X}, \mathcal{Y}, u_h\rangle$:

(a) $\mathcal{X} = \{\xi : \mathbb{N} \to X : \xi \in \mathcal{P}^1_T\}$ is the set of strategies for player 1;

(b) $\mathcal{Y} = \{\zeta : \mathbb{N} \to Y : \zeta \in \mathcal{P}^2_T\}$ is the set of strategies for player 2;

(c) $u_h : X^{\mathbb{N}} \times Y^{\mathbb{N}} \to \mathbb{R}$ is player 1's payoff function defined as (3.1); player 2's payoff function is $-u_h$.

Although the formulation of vertical games resembles the standard formulation of repeated games, the horizontal games seem more appropriate to capture the unpredictable behavior behind mixed strategy equilibrium in the standard framework. One example of a horizontal game is found in Luce and Raiffa [1957], where they discuss two aerial strategists deciding the actions of their pilots in a conflict consisting of many identical aircraft fights. They use this example to illustrate the meaning of a mixed strategy, which is interpreted as the distribution of different actions assigned to the pilots. Nonetheless, in many applications, the vertical games may serve a more faithful description. In any case, most of our results hold for both horizontal games and vertical games. For notational ease we will focus on horizontal games, and some results specific to vertical games will be given in Section 3.4.

We use the long-run average criterion (3.1) for payoffs. Chapter 1 gives an axiomatization of expected utility from the frequentist perspective that serves a foundation for this criterion. Because the sequence $\sum_{t=0}^{T-1} \frac{h(\xi_t,\zeta_t)}{T}$ may not have a limit, it is necessary to introduce limit inferior or superior or something similar. Our main results are robust to any such notion between the limit inferior and limit superior. However, the following proposition shows that the discounting criterion, defined as

$$v_h(\xi, \zeta) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t h(\xi_t, \zeta_t) \tag{3.2}$$

with $\delta \in (0, 1)$ being the discounting fact, is not consistent with equilibrium existence in our model. Its proof is given in Section 3.6.

Proposition 3.2.1. *Let $g = \langle X, Y, h \rangle$ be a two-person zero-sum game without any pure strategy equilibrium. If $\mathcal{P}^1$ and $\mathcal{P}^2$ satisfy (A1), then there is no equilibrium in*

$$HG'(g, \mathcal{P}^1, \mathcal{P}^2) = \langle \mathcal{X}, \mathcal{Y}, v_h \rangle,$$

where v_h is defined by (3.2).

This proposition implies that, in order to obtain a consistent theory of unpredictable behavior with the discounting criterion, it is necessary to assume that the players are not able to compute certain basic functions or they are not able to perform certain basic operations. We leave this difficulty for further research and move on with the long-run average criterion.

Under the long-run average criterion, equilibrium existence is not trivial either. Even though equilibrium may exist, it may not describe unpredictable behavior. Indeed, if one player fully exploit the other player in equilibrium, both players' actions would not be called unpredictable. The following proposition gives sufficient conditions on $\mathcal{P}^1$ and $\mathcal{P}^2$ for such nonexistence. It proof is given in Section 3.6.

Proposition 3.2.2. *Let $g = \langle X, Y, h \rangle$ be a two-person zero-sum game without any pure strategy equilibrium for either player. Suppose that $\mathcal{P}^1$ and $\mathcal{P}^2$ satisfy (A1).*

(a) If $\mathcal{P}^1 = \mathcal{P}^2$, then there is no equilibrium in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$.

(b) If $\mathcal{P}^2 \subset \mathcal{P}^1$, then the equilibrium payoff (if equilibrium exists) in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$ for player 1 is

$$\min_{y \in Y} \max_{x \in X} h(x, y).$$

Part (a) gives sufficient conditions for nonexistence; part (b) gives sufficient conditions either for nonexistence or for player 1 to fully exploit player 2. Of course, a similar result holds for the case $\mathcal{P}^1 \subset \mathcal{P}^2$. This result resembles the findings in Ben-Porath [1993][1], which obtains the same equilibrium payoffs when player 1 has a substantially stronger computational power. However, mixed strategies are allowed there and the existence problem is trivial.

There are examples of games for which there is no equilibrium in the associated horizontal games at all if $\mathcal{P}^2 \subset \mathcal{P}^1$, and we give one below. On the other hand, there is also a horizontal game that has an equilibrium with $\mathcal{P}^1 \subset \mathcal{P}^2$. This shows that the existence issue for the case that one player's feasible functions are a subset of other player's depends on the particular notion that extends the long-run average criterion to sequences without limit average payoffs (recall that we assume that player 1 maximizes limit inferior of the average payoffs and player 2 maximizes limit superior in our definitions).

Example 3.2.1. Consider the matching pennies game $g^{MP} = \langle \{H, T\}, \{H, T\}, h \rangle$ with

$$h(H, H) = 1 = h(T, T) \text{ and } h(H, T) = 0 = h(T, H).$$

Suppose that $\mathcal{P}^2 \subset \mathcal{P}^1$. Then, if the game $HG(g^{MP}, \mathcal{P}^1, \mathcal{P}^2)$ has an equilibrium, the equilibrium payoff to player 1 is 1. Let ξ^* be an equilibrium strategy of player 1 in that horizontal game. Then we have

$$\liminf_{T \to \infty} \frac{\sum_{t=0}^{T-1} h(\xi_t^*, H)}{T} = 1 = \liminf_{T \to \infty} \frac{\sum_{t=0}^{T-1} h(\xi_t^*, T)}{T},$$

[1] In that paper, players use finite automata to implement strategies.

and hence,

$$\liminf_{T\to\infty} \frac{|\{0 \le t \le T-1 : \xi_t^* = H\}|}{T} = 1 = \liminf_{T\to\infty} \frac{|\{0 \le t \le T-1 : \xi_t^* = T\}|}{T}.$$

But this implies that

$$1 = \liminf_{T\to\infty} \frac{|\{0 \le t \le T-1 : \xi_t^* = H\}| + |\{0 \le t \le T-1 : \xi_t^* = T\}|}{T}$$

$$\ge \liminf_{T\to\infty} \frac{|\{0 \le t \le T-1 : \xi_t^* = H\}|}{T} + \liminf_{T\to\infty} \frac{|\{0 \le t \le T-1 : \xi_t^* = T\}|}{T} = 2,$$

a contradiction. Thus, there is no equilibrium in $HG(g^{MP}, \mathcal{P}^1, \mathcal{P}^2)$.

Proposition 3.2.2 and Example 3.2.1 shows that, for equilibrium unpredictable behavior to exist, it is necessary that both $\mathcal{P}^1 - \mathcal{P}^2$ and $\mathcal{P}^2 - \mathcal{P}^1$ are nonempty—a condition we call *mutual uncomputability*. This condition says that each player has a feasible function that is uncomputable from the other player's perspective, and this result links unpredictability to uncomputability.

Mutual uncomputability, however, does not seem sufficient for equilibrium existence. Consider the horizontal game associated with g^{MP} as in Example 3.2.1. Suppose that player 1 plays H at odd places and uses an uncomputable sequence (from player 2's perspective) at even places; the resulting strategy is uncomputable from player 2's perspective, but it can be easily exploited by player 2 by playing T at odd places. The point is that simply being uncomputable may still be exploitable. In the next section, we consider another dimension of complexity that enables us to give a sufficient condition for equilibrium existence that avoids this problem.

3.3 Complexity and unpredictable behavior

Up to now we have considered complexity in terms of the computability relation. This notion can only tell us whether one function is computable from another, but, in the case of a negative answer, there is no information about how uncomputable a function is relative to another. Here, we consider a notion of complexity, the *Kolmogorov complexity*, that measures how uncomputable a strategy is relative to a set of feasible functions, and use this measure to formulate a condition on $\mathcal{P}^1$ and $\mathcal{P}^2$, called *mutual complexity*, that guarantees equilibrium existence.

3.3.1 Kolmogorov complexity and existence

Kolmogorov complexity, first introduced by Kolmogorov [1965] to study the foundation of probability theory, measures the minimum description lengths of a finite object, say, in a countable set Z. We use finite sequences over $\{0,1\}$ as possible descriptions, and a partial function $f : \{0,1\}^{<\mathbb{N}} \to Z$ is called a *language*. The Kolmogorov complexity (w.r.t. f) of an object z is defined as

$$K_f(z) = \min\{|\sigma| : \sigma \in \{0,1\}^{<\mathbb{N}},\ f(\sigma) = z\},$$

and $K_f(z) = \infty$ if there is no σ such that $f(\sigma) = z$. A string in $dom(f)$ is called a *code-word*, or simply a *word*. In the literature (c.f., van Lambalgen [1987]), it is argued that only *prefix-free languages*—i.e. languages f whose words are prefix-free in the sense that if $\sigma,\ \tau \in dom(f)$, then σ is not an initial segment of τ—give meaningful measures of the Kolmogorov complexity.

We use Kolmogorov complexity to measure how uncomputable a strategy is relative to a set of feasible functions $\mathcal{P}$, and we consider only prefix-free languages in $\mathcal{P}$, denoted by $D(\mathcal{P})$. However, a strategy ξ, as a total function (from $\mathbb{N}$ to $\{1, ..., m\}$), is an infinite object—it can be identified with its graph

$$\{(0, \xi_0), (1, \xi_1), ..., (n, \xi_n), ...\},$$

which can be further identified with an infinite sequence over $\{0, 1\}$.[2] To measure the complexity of such an infinite sequence, we follow the literature (c.f. Downey et al. [2006]) and consider the complexities of its initial segments. The set of objects we consider here is then $Z = \{0, 1\}^{<\mathbb{N}}$.

Before we formulate our sufficient condition for equilibrium existence, we give one additional assumption on the set of feasible functions $\mathcal{P}$, which requires the player have a most complicated function that generates all other feasible functions.

Assumption A2 There is a total function $f^* \in \mathcal{P}$ such that any function $f \in \mathcal{P}$ is computable from f^*.

Assumption (A1) gives a lower bound on what $\mathcal{P}^i$ has to contain; the second assumption, on the other hand, gives an upper bound. Given (A1), (A2) is equivalent to a formal statement[3] which basically says that the player has a black box that generates all the feasible functions within a finite process. The black box can be thought of as the player's mind, and the finiteness of the generation process is then almost inevitable. This boundary of rationality given by (A2), although it is indefinite, is parameterized by the most complicated function, and it makes our analysis tractable.

We strengthen mutual uncomputability by requiring each player have a sequence that is highly complicated from the other player's perspective in terms of Kolmogoorov complexity. This requirement is formulated as follows.

Definition 3.3.1. Let $\mathcal{P}$ be a set of functions. A sequence $\xi \in \{0, 1\}^{\mathbb{N}}$ is a *complex sequence*

[2]The set $\mathbb{N}^2$ can be effectively enumerated, and a graph can be identified with its characteristic function.
[3]See Pippenger [1997] for the formal statement.

relative to $\mathcal{P}$ if for all $f \in D(\mathcal{P})$, there is a constant b such that for all $T > 0$,

$$K_f(\xi[T]) \geq T - b,$$

where $\xi[T] = (\xi_0, ..., \xi_{T-1})$ is the initial segment of ξ with length T.

We say that two sets $\mathcal{P}^1$, $\mathcal{P}^2$ are *mutually complex* if there are there are $\xi^1 \in \mathcal{P}^1$, $\xi^2 \in \mathcal{P}^2$ such that for both $i = 1, 2$, ξ^i is a complex sequence relative to $\mathcal{P}^{-i}$. Mutual complexity, together with (A1) and (A2), guarantees equilibrium existence. The proof of the following theorem is given in Section 3.6.

Theorem 3.3.1. *Let g be a finite zero-sum game. Suppose that $\mathcal{P}^1$, $\mathcal{P}^2$ satisfy (A1) and (A2) and are mutually complex. Then there exists an equilibrium in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$ with the same equilibrium payoffs as those in g.*

Theorem 3.3.1 would be meaningless if mutual complexity was an empty assumption. The following proposition, which is a simple implication of what's well known in the literature (c.f. Downey et al. [2006]), states that it's satisfied by many pairs of sets. Its proof is given in Section 3.6.

Proposition 3.3.1. *There are uncountably many different pairs of sets $\mathcal{P}^1$ and $\mathcal{P}^2$ that satisfy (A1), (A2), and mutual complexity.*

Theorem 3.3.1 also conveys information about equilibrium behavior in terms of equilibrium payoffs: under mutual complexity, neither player can fully exploit the other. If g has no pure strategy equilibrium, then

$$\min_{y \in Y} \max_{x \in X} h(x, y) > v > \max_{x \in X} \min_{y \in Y} h(x, y),$$

where v is the (mixed strategy) equilibrium payoff to player 1 in g, and, by Theorem 3.3.1, v is also the equilibrium payoff in the horizontal game associated with g. This shows that in equilibrium each player's strategy is unpredictable to the other.

However, Theorem 3.3.1 does not convey the meaning of unpredictability. For example, what is the frequency of different actions in an equilibrium strategy? What statistical properties does an equilibrium strategy exhibit? We answer these questions in the next two subsections.

3.3.2 Unpredictable behavior

Here we introduce a criterion, called *stochasticity*, and compares this criterion with the meaning of unpredictability in horizontal games. We show that, in horizontal games of which the associated stage games have unique equilibria, this criterion almost characterizes equilibrium strategies under mutual complexity. As a result, we are able to show that the frequencies of different actions in equilibrium strategies in the horizontal game correspond to equilibrium mixed strategies in the associated stage game.

To define this criterion, we need one more concept: selection functions. A *selection function* is a total function from $\mathbb{N}$ to $\{0,1\}$ such that $|\{t : r(t) = 1\}| = \infty$. Given a finite set X and a sequence $\xi \in X^{\mathbb{N}}$, a selection function r induces a subsequence ξ^r as follows: $\xi^r_t = \xi_{g(t)}$, where $g(0) = \min\{s : r(s) = 1\}$ and $g(t+1) = \min\{s : s > g(t),\ r(s) = 1\}$—that is, ξ^r is the subsequence of ξ from places where r has value 1. We define stochasticity as follows.

Definition 3.3.2. Let $p \in \Delta(X)$ be a probability distribution and let $\mathcal{P}$ be a set of functions. A sequence $\xi \in X^{\mathbb{N}}$ is *p-stochastic relative to $\mathcal{P}$* if for any selection function $r \in \mathcal{P}$ and for all $x \in X$,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi^r_t)}{T} = p_x,$$

where $c_x(y) = 1$ if $x = y$, and $c_x(y) = 0$ otherwise.

von Mises [1981] defines a sequence ξ to be p-stochastic if there is no *admissible* subsequence of ξ that has relative frequency different from p. Clearly, without the quantifier 'admissible,' there is no p-stochastic sequence unless p is degenerate. Here we define admissibility in terms of feasible functions in $\mathcal{P}$.[4]

We first show that if $\mathcal{P}^1$ and $\mathcal{P}^2$ satisfy mutual complexity, then both players $i = 1, 2$ can generate stochastic sequences relative to $\mathcal{P}^{-i}$.

Proposition 3.3.2. *Suppose that $\mathcal{P}^1$, $\mathcal{P}^2$ satisfy (A1), (A2), and mutual complexity. Then, for $p \in \Delta(X)$, there exists a strategy $\xi \in \mathcal{X}$ that is p-stochastic relative to $\mathcal{P}^2$*

Proposition 3.3.2 is well known in the literature, but for completeness, we give its proof in the Appendix. We then show that stochasticity is a sufficient condition for a strategy to be optimal in horizontal games under mutual complexity.

Theorem 3.3.2. *Suppose that $\mathcal{P}^1$, $\mathcal{P}^2$ satisfy (A1), (A2), and mutual complexity. Then, for any equilibrium strategy $p^* \in \Delta(X)$ in g, if $\xi \in \mathcal{X}$ is p^*-stochastic relative to $\mathcal{P}^2$, then it is an equilibrium strategy of player 1 in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$.*

Theorem 3.3.2 identifies a subset of equilibrium strategies. Together with Proposition 3.3.2, it implies that for any equilibrium mixed strategy p for player 1 in the stage game, there is an equilibrium strategy that is p-stochastic sequence relative to $\mathcal{P}^2$ for player 1 in the horizontal game. In particular, this implies that for any equilibrium mixed strategy p in the stage game, there is a corresponding equilibrium strategy in the horizontal game with the frequency of different actions equal to p. The symmetric statement for player 2, of course, holds as well.

[4] In the literature, stochasticity is defined with selection functions that are history-dependent—i.e., $r : X^{<\mathbb{N}} \to \{0, 1\}$. That concept is close to unpredictability in vertical games, where players observe past plays. We discuss this in Section 3.4.

The proof of Theorem 3.3.2 is given in Section 3.6, but we shall illustrate the intuition with the horizontal game associated with the matching pennies g^{MP}. Suppose that $\xi \in \{H,T\}^{\mathbb{N}}$ is an equilibrium strategy for player 1 in $HG(g^{MP}, \mathcal{P}^1, \mathcal{P}^2)$. Then, for any player 2's strategy $\zeta \in \{H,T\}^{\mathbb{N}}$, $u_h(\xi,\zeta) \geq 0$. Any player 2' strategy ζ can be identified with two selection functions: $r^{\zeta,H}$ defined by $r^{\zeta,H}(t) = 1$ if and only if $\zeta_t = H$ and $r^{\zeta,T}$ defined by $r^{\zeta,T}(t) = 1$ if and only if $\zeta_t = T$. Player 1's payoff by playing ξ against ζ is then determined by three frequencies: the frequency of H in $\xi^{r^{\zeta,H}}$, the frequency of H in $\xi^{r^{\zeta,T}}$, and frequency of H in ζ. Let's assume that these frequencies all exist and denote them by $p_1^{\xi,\zeta}$, $p_2^{\xi,\zeta}$, and p_H^{ζ}, respectively. Then player 1's payoff is

$$p_H^{\zeta}(2p_1^{\xi,\zeta} - 1) + (1 - p_H^{\zeta})(1 - 2p_2^{\xi,\zeta}).$$

Player 1 can guarantee 0 payoff if

$$p_1^{\xi,\zeta} = p_2^{\xi,\zeta} = \frac{1}{2} \text{ for all } \zeta \in \mathcal{Y}. \tag{3.3}$$

Because any ζ can be identified with a selection function, (3.3) is satisfied if ξ is a stochastic sequence relative to $\mathcal{P}^2$. This argument can be generalized to any horizontal games, but we deal with certain technical issues related to the problem that not every sequence of average payoffs has a limit.

After seeing a sufficient condition for unpredictability in horizontal games, we now give a necessary condition that is closely related to stochasticity. The proof of Theorem 3.3.3 is given in Section 3.6.

Theorem 3.3.3. *Suppose that $\mathcal{P}^1$, $\mathcal{P}^2$ satisfy (A1), (A2), and mutual complexity. Then, for any equilibrium strategy $\xi \in \mathcal{X}$ with limit relative frequency for each action $x \in X$ in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$ and any selection function $r \in \mathcal{P}^2$ such that $\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{r(t)}{T} > 0$, if for all $x \in X$, $p_x = \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t^r)}{T}$ exists, then $p = (p_x)_{x\in X}$ is an equilibrium mixed strategy in g.*

Although Theorem 3.3.3 gives a weaker condition than stochasticity, it helps identify the frequencies an equilibrium strategy in the horizontal game have: taking r as the selection function that has constant value 1, it implies that the frequency of an equilibrium strategy is an equilibrium mixed strategy of the associated stage game. Theorem 3.3.2 and Theorem 3.3.3 then give a formal correspondence between the relative frequency in equilibrium strategies in the horizontal game and equilibrium mixed strategies in the associated stage game.

The necessary condition in Theorem 3.3.3 is less informative than stochasticity in two aspects: First, it does not tell us whether the mentioned frequency existed or not; second, even if it exists, it only tells us that it is equal to one of the mixed strategy equilibrium of the associated stage game, and the frequencies in different subsequences may be be different. The first issue, we suspect, is related to how the long-run average criterion is extended to sequences of average payoffs that have no limits and so a complete solution seems not available; the second issue, however, does not appear if the associated stage game has a unique mixed strategy equilibrium, as in the matching pennies.

Corollary 3.3.1. *Suppose that $\mathcal{P}^1$, $\mathcal{P}^2$ satisfy (A1), (A2), and mutual complexity. Suppose that g has a unique equilibrium (p^1, p^2). Then, for any equilibrium strategy $\xi \in \mathcal{X}$ with limit relative frequency for each $x \in X$ in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$ and any selection function $r \in \mathcal{P}^2$ such that $\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{r(t)}{T} > 0$, if for all $x \in X$, $p_x = \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t^r)}{T}$ exists, then $p = p^1$.*

Corollary 3.3.1 shows that stochasticity is very close to the meaning of unpredictability in horizontal games of which the associated stage games have unique mixed strategy equilibria. This also renders some empirical implications and we shall return to this point in the concluding remarks.

Although we have compared the meaning of unpredictability in horizontal games with stochasticity, this criterion is not commonly used in the literature. Unpredictability is usually associated with statistical tests—in the context of zero-sum games, in particular, there is an empirical literature that tests equilibrium hypothesis using statistical tests. We consider this criterion in the next section and compare it with the meaning of unpredictability in horizontal games.

3.3.3 Statistical tests

In general, a statistical test corresponds to a probability law—i.e., a *describable* event that has probability 1. In the literature, these tests have been used to test equilibrium hypothesis in repeated zero-sum games. Some tests are very intuitive: For example, in a repeated matching pennies, a player is expected to play a strategy that satisfies the Law of Large Numbers according to the half-half distribution on H and T, and by observing a violation of this law one can reject the equilibrium hypothesis. Our result in the last section confirm that this expectation is justified in our framework. However, the question remains regarding what other probability laws the equilibrium strategies do and do not satisfy.

We address this question by invoking a criterion of unpredictability called *Martin-Löf randomness* given in Martin-Löf [1966] (hereafter M-L randomness). This criterion begins with a formal formulation of idealized statistical tests based on probability laws. We need a couple more concepts to define these tests and the criterion. Let X be a finite set. We endow the set of infinite sequences over X with the product topology, and we consider only Borel probability measures on it. Any open set can be written as a union of basic sets, where a basic set has the form $N_\sigma = \{\zeta \in X^{\mathbb{N}} : \sigma = \zeta[|\sigma|]\}$ for some $\sigma \in X^{<\mathbb{N}}$. We give a formal definition of randomness in the following.

Definition 3.3.3. Let X be a finite set and let $\mathcal{P}$ be a set of functions. Suppose that μ is a computable probability measure over $X^{\mathbb{N}}$,[5] i.e., the mapping $\sigma \mapsto \mu(N_\sigma)$ belongs to $\mathcal{T}$.[6] A sequence of open sets $\{V_t\}_{t=0}^{\infty}$ is *a μ-test relative to $\mathcal{P}$* if it satisfies the following conditions:

(1) There is a function $f : \mathbb{N} \to \mathbb{N} \times X^{<\mathbb{N}}$ in $\mathcal{P}_T$ such that for all $t \in \mathbb{N}$ and for all $\xi \in X^{\mathbb{N}}$,

$$\xi \in V_t \Leftrightarrow (\exists n)(f(n) = (t, \sigma) \wedge \sigma = \xi[|\sigma|]).$$

(2) For all $t \in \mathbb{N}$, $\mu(V_t) \leq 2^{-t}$.

A sequence $\xi \in X^{\mathbb{N}}$ is *μ-random relative to $\mathcal{P}$* if it passes all μ-tests relative to $\mathcal{P}$, i.e., for any μ-test $\{V_t\}_{t=0}^{\infty}$ relative to $\mathcal{P}$, $\xi \notin \bigcap_{t=0}^{\infty} V_t$.

The essence of this definition is to recognize that a probability law is not an arbitrary event that has probability 1, because there exists no sequence satisfies all such events with probability 1. The quantifier 'describable' is necessary but hard to define, and one attempt is given here by stating that an event with probability 1 is describable if there is a constructive proof for that fact. A test $\{V_t\}_{t=0}^{\infty}$ is used to establish a probability law that corresponds to the complement of $\bigcap_{t=0}^{\infty} V_t$, which corresponds to a probability zero property. Conditions (1) and (2) require that this test, as a sequence of open sets, can be generated by functions in $\mathcal{P}$, and hence it is constructive with respect to $\mathcal{P}$. A sequence is random if it passes all such tests, that is, it satisfies all probability laws thus defined. As a consequence, the set of random sequences, given a fixed measure, depends on the set of functions employed in a monotonic manner. If $\mathcal{P}^1 \subset \mathcal{P}^2$, then any test relative to $\mathcal{P}^1$ is also a test relative to $\mathcal{P}^2$. Therefore, the set of random sequences relative to $\mathcal{P}^2$ is a subset of random sequences relative to $\mathcal{P}^1$.

[5] Implicitly in the definition we assume that $\mu(N_\sigma)$ is always a rational number for μ to be computable. In the literature, computability of a measure is defined more generally, but this definition is sufficient for our purpose.

[6] The set $\mathcal{T}$ is the set of Turing-computable functions—i.e., the smallest set of functions that satisfy both (A1) and (A2).

We are mostly interested in Bernoulli measures—i.e., measures generated by a distribution over X in an i.i.d. manner. For any $p \in \Delta(X)$, we use μ_p to denote the Bernoulli measure over $X^{\mathbb{N}}$ generated by p: for all $\sigma \in X^{<\mathbb{N}}$, $\mu_p(N_\sigma) = \prod_{t=0}^{|\sigma|-1} p_{\sigma_t}$. In what follows, if M-L randomness is mentioned without reference to any probability measure, it means M-L randomness with respect to a i.i.d. measure μ_p. First we give an existence result that is known in the literature, and show that M-L randomness is a more informative criterion than stochasticity.

Proposition 3.3.3. *(a) Suppose that $\mathcal{P}^1$ and $\mathcal{P}^2$ satisfy (A1), (A2), and mutual complexity. Then for any $p \in \Delta(X)$, there exists a strategy $\xi \in \mathcal{X}$ that is μ_p-random relative to $\mathcal{P}^2$.*

(b) For any set $\mathcal{P}$ that satisfies (A1) and (A2), if $\xi \in X^{\mathbb{N}}$ is μ_p-random relative to $\mathcal{P}$, then it is p-stochastic relative to $\mathcal{P}$.

The proof of Proposition 3.3.3 is a slight generalization of well-known results in the literature, but it is given in the Appendix for completeness. The following theorem is a direct corollary of this proposition and Theorem 3.3.2.

Theorem 3.3.4. *Suppose that $\mathcal{P}^1$, $\mathcal{P}^2$ satisfy (A1), (A2), and mutual complexity. Then, for any equilibrium strategy $p^* \in \Delta(X)$ in g, there exists an equilibrium strategy $\xi \in \mathcal{X}$ for player 1 that is μ_{p^*}-random relative to $\mathcal{P}^2$ in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$.*

Theorem 3.3.4 shows that, under mutual complexity, there exist equilibrium strategies that satisfy all probability laws with respect to the i.i.d. distributions generated by the equilibrium mixed strategies of the associated stage game. It does not, however, imply that all equilibrium strategies satisfy all such probability laws; in fact, there always exists an equilibrium strategy that fails a particular probability law with respect to the i.i.d. distribution, the Law of the Iterated Logarithm.

Theorem 3.3.5. *Suppose that $\mathcal{P}^1$, $\mathcal{P}^2$ satisfy (A1), (A2), and mutual complexity. Suppose that g has no pure strategy equilibrium. Let p^* be an equilibrium mixed strategy for player 1 in g, and* *let $x \in X$ be an action such that $p^*_x > 0$. There exists an equilibrium strategy ξ in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$ that is p^*-stochastic and satisfies*

$$\lim_{T\to\infty} \frac{\sum_{t=0}^{T} c_x(\xi_t) - Tp^*_x}{\sqrt{2p^*_x(1-p^*_x)T\log\log T}} = \infty. \tag{3.4}$$

The proof of Theorem 3.3.5 is given in Section 3.6. The Law of the Iterated Logarithm states that, if ξ is a M-L random sequence with respect to the i.i.d. distribution μ_{p^*},

$$\limsup_{T\to\infty} \frac{|\sum_{t=0}^{T} c_x(\xi_t) - Tp^*_x|}{\sqrt{2p^*_x(1-p^*_x)T\log\log T}} = 1. \tag{3.5}$$

Equation (3.4) shows that there is always an equilibrium strategy that fails the Law of the Iterated Logarithm. Moreover, this shows that the subset of equilibrium strategies identified by Theorem 3.3.2 is a strict subset of equilibrium strategies identified by Theorem 3.3.4—i.e., there are equilibrium strategies that are stochastic but not M-L random.

Our results regarding the meaning of unpredictable behavior all rely on mutual complexity, which may be stronger than what is necessary for existence. However, the conclusion that if equilibrium unpredictable behavior exists, there always exists equilibrium strategies that are not M-L random is general.

Theorem 3.3.6. *Let $\mathcal{P}^1$, $\mathcal{P}^2$ be two sets of functions satisfying (A1) and (A2). Let $p \in \Delta(X)$ and $q \in \Delta(Y)$ be non-degenerate distributions.*[7] *Suppose that there is a μ_p-random sequence relative to $\mathcal{P}^2$ in $\mathcal{X}$ and there is a μ_q-random relative to $\mathcal{P}^1$ in $\mathcal{Y}$. Then $\mathcal{P}^1$, $\mathcal{P}^2$ are mutually complex.*

The proof of Theorem 3.3.6 is given in Section 3.6. This theorem implies that if mutual complexity does not hold, at least one player is not able to generate a M-L random sequence, and so that player's equilibrium strategies cannot be M-L random. We have seen that if mutual

[7] A distribution p is non-degenerate if $p_{x_1} > 0$ and $p_{x_2} > 0$ for two different elements x_1, x_2.

complexity holds, there exist equilibrium strategies that are not M-L random. Therefore, this theorem, together with Theorem 3.3.5, implies that the meaning of unpredictability in horizontal games is strictly weaker than M-L randomness: Failure of certain probability laws does not entail a rejection of the equilibrium hypothesis in our model. This also suggests that the common practice which rejects the equilibrium hypothesis by using statistical tests in the context of repeated zero-sum games without any restrictions needs a further reconsideration.

3.4 Results for vertical games

In this section we clarify our results for vertical games. Both Proposition 3.2.1 and Proposition 3.2.2 hold for vertical games without any change. Theorem 3.3.1 holds for vertical games as well.

However, the criterion stochasticity has to be strengthened for vertical games. In particular, the selection functions have to be history dependent for vertical games, while they are history independent for horizontal games. Thus, given a finite set X, we redefine a *selection function for* X as a total function $r : X^{<\mathbb{N}} \to \{0, 1\}$, which, as we have seen before, can be identified with a total function over natural numbers. Given a sequence $\xi \in X^{\mathbb{N}}$, we use r to choose a subsequence ξ^r from ξ as follows: $\xi^r_t = \xi_{g(t)}$, where $g(0) = \min\{t : r(\xi[t]) = 1\}$ and $g(t) = \min\{s : r(\xi[s]) = 1,\ s > g(t-1)\}$ for $t > 0$. Obviously, such a selection function may not produce an infinite subsequence, but only produce a finite initial segment. We call the criterion of unpredictability based these selection functions the *strong stochasticity*.

Definition 3.4.1. Let $p \in \Delta(X)$ be a probability distribution and let $\mathcal{P}$ be a set of functions. A sequence $\xi \in X^{\mathbb{N}}$ is *strongly p-stochastic relative to* $\mathcal{P}$ if for any selection function $r \in \mathcal{P}$ for X

such that ξ^r is an infinite sequence,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t^r)}{T} = p_x \text{ for all } x \in X,$$

where $c_x(y) = 1$ if $x = y$, and $c_x(y) = 0$ otherwise.

Replacing stochasticity by strong stochasticity, Proposition 3.3.2 and Theorem 3.3.2 continue to hold, with the understanding that a sequence $\xi \in X^{\mathbb{N}}$ is identified with the strategy $a : Y^{<\mathbb{N}} \to X$ such that for all $\sigma \in Y^{<\mathbb{N}}$, $a(\sigma) = \xi_{|\sigma|}$ in the vertical game.

Theorem 3.3.3 holds for vertical games as well, but only for equilibrium strategies that are history independent—i.e., strategies that can be identified with sequences in $X^{\mathbb{N}}$ as above; in fact, for history independent equilibrium strategies, we are able to obtain a slightly stronger necessary condition as follows.

Theorem 3.4.1. *Suppose that $\mathcal{P}^1$, $\mathcal{P}^2$ satisfy (A1), (A2), and mutual complexity. Then, for any equilibrium strategy $\xi \in \mathcal{X}$ with limit relative frequency for each $x \in X$ in $VG(g, \mathcal{P}^1, \mathcal{P}^2)$ and any selection function $r \in \mathcal{P}^2$ for X such that $\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{r(\xi[t])}{T} > 0$, if for all $x \in X$, $p_x = \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t^r)}{T}$ exists, then $p = (p_x)_{x\in X}$ is an equilibrium mixed strategy in g.*

The proof of Theorem 3.4.1 is given in Section 3.6. If r is history independent, the condition here is equivalent to that in Theorem 3.3.3; Theorem 3.4.1 is stronger because it applies to history dependent selection functions. However, this does not give a close characterization of unpredictability as in horizontal games: There are many history dependent strategies that are not characterized by any condition close to stochasticity. For example, the following strategy a is an equilibrium strategy in $VG(g^{MP}, \mathcal{P}^1, \mathcal{P}^2)$ with mutual complexity: $a(\sigma) = H$ if $\sigma_t = H$ for all $t = 0, ..., |\sigma| - 1$, and $a(\sigma) = \xi_{|\sigma|-1}$ otherwise, where ξ is a strongly $(\frac{1}{2}, \frac{1}{2})$-stochastic sequence relative to $\mathcal{P}^2$.

Nevertheless, Theorem 3.4.1 does imply that the meaning of unpredictability is different in vertical games and horizontal game: There are stochastic sequences that do not satisfy the condition in Theorem 3.4.1.

Finally, all the results in Section 3.3 apply to vertical games as well, even if we replace stochasticity with strong stochasticity. In both vertical and horizontal games, applying any arbitrary statistical test to reject the equilibrium hypothesis is not justified.

3.5 Discussions

In this section, we discuss implications of our results to the literature, and then point out possible interpretations of Nash equilibrium in our model, and, finally, consider two possible extensions.

3.5.1 Unpredictability in game theory

Unpredictable behavior is usually assumed to exist in the literature since von Neumann and Morgenstern [1944]. Our nonexistence results, however, suggest that some minimum complexity requirement, mutual uncomputability, on players' ability to generate strategies is necessary for unpredictable behavior. Although there is a huge literature that accumulates evidence of unpredictable behavior and its statistical properties, little has been said about the complexity of such behavior.

On the other hand, the same literature, by and large, reports that players do not generate 'random' strategies in terms of statistical tests in the context of repeated zero-sum games. However, our characterization results suggest that not all the tests are relevant to the equilibrium hypothesis. We are able to show that in both vertical and horizontal games, if equilibrium unpredictable behavior exists, there always exist equilibrium strategies that fail some probability laws. Moreover, we obtain a close characterization of unpredictable behavior in horizontal games in terms of stochasticity. The matter in vertical games seems more complicated, but stochasticity certainly plays an important role. These results may help identify relevant tests to the equilibrium hypothesis in our model.

3.5.2 Interpretation of Nash equilibrium

One conceptual difficulty encountered in our model is the interpretation of Nash equilibrium. It should be emphasized here that in zero-sum games, Nash equilibrium can be reduced to maximin criterion and there is no issue of equilibrium selection. One possible interpretation is that both players' are fully rational (and hence players are able to use the maximin criterion), and the limitation of strategy sets comes from limitation, not on rationality, but on implementation of strategies. This interpretation has been adopted by Chatterjee and Sabourian [2008] for games with finite automata.

On the other hand, a bounded rationality interpretation is also possible in our model: Namely, a player only knows his feasible functions. Our result shows that, if, in addition, the player knows that he has a feasible complex sequence relative to the other player's set of feasible functions and that the other player also has a complex sequence relative to his feasible functions, then the player can generate a maximin strategy from his feasible functions.

3.5.3 Extensions

We propose two possible extensions here. The first possibility is to introduce a learning process, and the second is to consider finite sequences instead of infinite ones. Unlike the extant learning literature, in our model the main challenge is to understand whether a player can learn the other player's strategy or not and how, and a modeling issue is whether we want a model in which the horizontal or vertical game is repeated infinitely many times.

Regardless of how the modeling issue is resolved, this project may help us understand a stability issue regarding mixed strategy equilibria. For example, consider the horizontal game associated with the following coordination game $g^C = \langle \{H, T\}, \{H, T\}, h_1, h_2 \rangle$ with

$$h_1(H, H) = 1 = h_1(T, T) \text{ and } h_1(H, T) = 0 = h_1(T, H);\; h_1 = h_2.$$

Although g^C is not a zero-sum game, all the results we have apply to the mixed strategy equilibrium $(\frac{1}{2}, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})$. However, intuitively, any unpredictable strategy in the associated horizontal game is not stable: Players will attempt to be predictable. On the other hand, one may suspect that the unpredictable equilibrium we have for zero-sum games are stable. But the question remains how they learn to be unpredictable and learn that the other is unpredictable in zero-sum games and how they learn to be predictable in coordination games. Nonetheless, our model gives a definition of unpredictability and so these questions become meaningful.

Extending the model to accommodate finite sequences seems very natural because we live in a finite world, and it may give us sharper empirical implications than what we have here. Moreover, the Kolmogorov complexity was originally introduced to study the complexity of finite sequences. However, a serious modeling issue here bout language dependence arises: Although Kolmogorov complexity can be defined for any language, it gives very different measures across different language. Which language should we use? There seems no widely accepted answer to this question.

3.6 Proofs of the main theorems

In this section we give the proofs of the main theorems. In the appendix, we review some extant results regarding Kolmogorov complexity and Martin-Löf randomness that are necessary in the proofs. We begin with the proofs of the nonexistence results, and then we present the proofs of the other theorems in the second subsection.

3.6.1 Proofs of the nonexistence results

Proof of Proposition 3.2.1: We give the proof for vertical games. The proof for horizontal games is very similar and is omitted. We show that there is no ε-equilibrium for ε sufficiently small. Suppose that (a^*, b^*) is an ε-equilibrium. Let

$$c = \max_{(x,y)\in X\times Y} |h(x,y)|.$$

For each $T \in \mathbb{N}$, consider the following strategy b' such that for all $\sigma \in X^{<\mathbb{N}}$, $b'(\sigma) = y_{\zeta_{|\sigma|}}$, where ζ is defined as follows:

$$\begin{aligned} &\text{for } t = 0,\ \zeta_0 = \min\{i : y_i \in \arg\min_{y\in Y} h(a^*(\epsilon), y)\}; \\ &\text{for } t = 1, ..., T,\ \zeta_t = \min\{i : y_i \in \arg\min_{y\in Y} h(a^*(y_\zeta[t]), y)\}; \\ &\text{for } t > T+1,\ \zeta_t = 1. \end{aligned} \tag{3.6}$$

Since ζ_t is constant for all $t > T$, $\zeta \in \mathcal{P}^2$ and so $b' \in \mathcal{Y}$. We have

$$v(a^*, b') \leq (1-\delta)\sum_{t=0}^{T} \delta^t v_1 + (1-\delta)\sum_{t=T+1}^{\infty} c = v_1 + \delta^{T+1}(c - v_1).$$

Since (a^*, b^*) is an ε-equilibrium, it follows that

$$v(a^*, b^*) \le v(a^*, b') + \varepsilon \le v_1 + \varepsilon + \delta^{T+1}(c - v_1).$$

Since this holds for all T, we have

$$v(a^*, b^*) \le \lim_{T\to\infty} v_1 + \varepsilon + \delta^{T+1}(c - v_1) = v_1 + \varepsilon.$$

Similarly, we can show that

$$v(a^*, b^*) \ge v_2 - \varepsilon.$$

It then follows that, for any $\varepsilon < \frac{v_2 - v_1}{2}$,

$$v_1 + \varepsilon < v_2 - \varepsilon \le v(a^*, b^*) \le v_1 + \varepsilon,$$

a contradiction. $\square$

Proof of Proposition 3.2.2: (a) We consider vertical games here, and the proof for horizontal games is almost identical. We first index the actions in Y as $Y = \{y_1, ..., y_n\}$. Since there is no pure strategy equilibrium in g, it follows that

$$v_1 = \max_{x\in X} \min_{y\in Y} h(x, y) < \min_{y\in Y} \max_{x\in X} h(x, y) = v_2.$$

Suppose that, to the contrary, (a^*, b^*) is an equilibrium strategy profile in $VG(g, \mathcal{P}, \mathcal{P})$. Consider the strategy b' such that for all $\sigma \in X^{<\mathbb{N}}$, $b'(\sigma) = y_{\zeta_{|\sigma|}}$, where ζ is defined as follows:

$$\begin{aligned} t = 0,\ \zeta_0 &= \min\{i : y_i \in \arg\min_{y\in Y} h(a^*(\epsilon), y)\}; \\ t > 0,\ \zeta_t &= \min\{i : y_i \in \arg\min_{y\in Y} h(a^*(y_\zeta[t]), y)\}, \end{aligned} \tag{3.7}$$

where $y_\zeta[t] = (y_{\zeta_0}, ..., y_{\zeta_{t-1}})$. $\zeta \in \mathcal{P}$ because $a^* \in \mathcal{P}$ and $\mathcal{P}$ is closed under composition and primitive recursion. Thus, $b' \in \mathcal{Y}$. By construction, we have that

$$h(a^*(y_\zeta[t]), \zeta_t) \leq \max_{x \in X} \min_{y \in Y} h(x, y) = v_1.$$

Then,

$$u(a^*, b') \leq \liminf_{T \to \infty} v_1 = v_1.$$

Since b^* is a best response to a^*, it follows that

$$u(a^*, b^*) \leq u(a^*, b') \leq v_1.$$

Similarly, we can show that

$$u(a^*, b^*) \geq v_2.$$

But then

$$v_2 > v_1 \geq u(a^*, b^*) \geq v_2,$$

a contradiction.

(b) We first index the actions in X as $X = \{x_1, ..., x_m\}$. For any $b \in \mathcal{Y}$, construct the strategy a' such that for all $\sigma \in Y^{<\mathbb{N}}$, $a'(\sigma) = x_{\zeta_{|\sigma|}}$, where ζ is defined as follows:

$$\begin{aligned} t = 0,\ \zeta_0 &= \min\{i : x_i \in \arg\max_{x \in X} h(x, b(\epsilon))\}; \\ t > 0,\ \zeta_t &= \min\{i : x_i \in \arg\max_{x \in X} h(x, b(x_\zeta[t]))\}, \end{aligned} \tag{3.8}$$

with $x_\zeta[t] = (x_{\zeta_0}, ..., x_{\zeta_{t-1}})$. $\zeta \in \mathcal{P}^2$ because $b \in \mathcal{P}^2$ and $\mathcal{P}^2$ is closed under composition and primitive recursion. Thus, $a' \in \mathcal{X}$ since $\mathcal{P}^2 \subset \mathcal{P}^1$. By construction, we have that for all $t \in \mathbb{N}$,

$$h(\zeta_t, b(x_\zeta[t])) \geq \min_{y \in Y} \max_{x \in X} h(x, y) = v_2.$$

Then, $u_h(a', b) \geq \liminf_{T \to \infty} v_2 = v_2$. It follows that $\sup_{a \in \mathcal{X}} u_h(a, b) \geq v_2$. Now, let

$$y^* \in \arg\min_{y \in Y} (\max_{x \in X} h(x, y)).$$

Let $b \in \mathcal{P}^2$ be such that $b(\tau) = y^*$ for all $\tau \in X^{<\mathbb{N}}$. Then $\sup_{a \in \mathcal{X}} u_h(a, b) = v_2$. Thus, we have that $\min_{b \in \mathcal{Y}} \inf_{a \in \mathcal{X}} u_h(a, b) = v_2$, and hence, the value of $VG(g, \mathcal{P}^1, \mathcal{P}^2)$ is v_2. $\square$

3.6.2 Proofs of the existence and the unpredictable behavior results

In this section, we will begin with two theorems (Theorems 3.6.1 and 3.6.2) which show that, under (A1) and (A2), if each player has a (strongly) stochastic sequence relative to the opponent's set of feasible functions which has an equilibrium frequency in the stage game, then these (strongly) stochastic sequences constitute an equilibrium in (VG) HG. This result, together with Proposition 3.3.2 and Proposition 3.3.3, then implies Theorem 3.3.1, Theorem 3.3.2, and Theorem 3.3.4. Proposition 3.3.2 is implied by Proposition 3.3.3, which is in turn a direct corollary of Theorem 3.7.2 and Theorem 3.7.6.

To prove Theorem 3.3.5, we consider a sequence of probability distributions $\mathbf{p} = \{p^t\}_{t=0}^{\infty}$ that converges to an equilibrium strategy $p \in \Delta(X)$ in g. Then, by Theorem 3.7.6, which can be found in the appendix, any sequence that is $\mu_{\mathbf{p}}$-random relative to $\mathcal{P}^2$ (for all $\sigma \in X^{<\mathbb{N}}$, $\mu_{\mathbf{p}}(N_\sigma) = \prod_{t=0}^{|\sigma|-1} p^t_{\sigma_t}$) is also p-stochastic relative to $\mathcal{P}^2$. By Theorem 3.6.1, this sequence is an equilibrium strategy. It then remains to show that there is a measure $\mu_{\mathbf{p}}$ such that any sequence ξ that is $\mu_{\mathbf{p}}$-random will fails the Law of the Iterated Logarithm for the i.i.d. distribution μ_p.

Theorem 3.7.7 gives a general law for $\mu_{\mathbf{p}}$, and, with appropriately chosen $\mathbf{p}$, we show that the sequence ξ will do the job.

To prove Theorem 3.3.6, first we show that any binary random sequence with respect to the uniform distribution is a complex sequence, and any random sequence that is generated by an non-degenerate i.i.d. measure can be used to compute a binary random sequence with respect to the uniform distribution. Of course, all these have to be relativized with respect to a computability constraint. This theorem is closely related to the principle that randomness is equivalent to extreme complexity, and interested readers may go to the survey paper Downey et al. [2006].

First we give a lemma concerning expected values. For (strongly) stochastic sequences, the expected values correspond to long-run averages.

Lemma 3.6.1. *Let X be a finite set. Let $\mathcal{P} \in \mathfrak{R}$ and let $p \in \Delta(X)$ be a distribution. Suppose that $h : X \to \mathbb{Q}$ is a function over X. If ξ is a strongly p-stochastic sequence relative to $\mathcal{P}$, then, for any selection function r in $\mathcal{P}$ such that $\xi^r \in X^{\mathbb{N}}$, we have*

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi^r_t)}{T} = \sum_{x\in X} p_x h(x).$$

Proof. Let r be a selection function in $\mathcal{P}$ such that $\xi^r \in X^{\mathbb{N}}$. Then, for any $x \in X$,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi^r_t)}{T} = p_x.$$

Therefore,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi^r_t)}{T} = \lim_{T\to\infty} \sum_{x\in X} \sum_{t=0}^{T-1} \frac{c_x(\xi^r_t)h(x)}{T} = \sum_{x\in X} p_x h(x).$$

□

A corresponding version of Lemma 3.6.1 holds for stochastic sequences, and its proof is almost identical to the above one.

Now, we shall give two theorems that guarantee the existence of equilibria in the vertical games and the horizontal games, respectively, that have the same equilibrium payoffs as the stage game.

Theorem 3.6.1. *Suppose that $\mathcal{P}^1$ and $\mathcal{P}^2$ satisfy (A1) and (A2). Let $(p^*, q^*) \in \Delta(X) \times \Delta(Y)$ be an equilibrium of g. Suppose that there are $\xi \in \mathcal{P}^1$, $\zeta \in \mathcal{P}^2$ such that ξ^* is p^*-stochastic relative to $\mathcal{P}^2$ and ζ^* is q^*-stochastic relative to $\mathcal{P}^1$. Then, any p-stochastic sequence relative $\mathcal{P}^2$ for some equilibrium mixed strategy $p \in \Delta(X)$ is an equilibrium strategy in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$.*

Proof. Let ξ^* be a p^*-stochastic sequence relative to $\mathcal{P}^2$ and let ζ^* be a q^*-stochastic sequence relative to $\mathcal{P}^1$.

First we show that

$$(\forall \xi \in \mathcal{X})(\limsup_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t, \zeta^*_t)}{T} \leq h(p^*, q^*)), \tag{3.9}$$

Because (p^*, q^*) is an equilibrium of the game g, $h(x, q^*) \leq h(p^*, q^*)$ for all $x \in X$.

Suppose that $\xi \in \mathcal{X}$, and so $\xi \in \mathcal{P}^1$. For each $x \in X$, let $r^x : \mathbb{N} \to \{0, 1\}$ be the selection function such that $r^x(t) = 1$ if $\xi_t = x$, and $r^x(t) = 0$ otherwise. Define

$$L_x(T) = |\{t \in \mathbb{N} : 0 \leq t \leq T-1,\ r^x(t) = 1\}| \text{ and } \zeta^x = (\zeta^*)^{r^x}.$$

It is easy to see that r^x is in $\mathcal{P}^1$ since ξ is. Let

$$\mathcal{E}^1 = \{x \in X : \lim_{T\to\infty} L_x(T) = \infty\} \text{ and } \mathcal{E}^2 = \{x \in X : \lim_{T\to\infty} L_x(T) < \infty\}.$$

For each $x \in \mathcal{E}^2$, let $B_x = \lim_{T\to\infty} L_x(T)$ and let $C_x = \sum_{t=0}^{B_x} h(x, \zeta^x_t)$. Then, for any $x \in \mathcal{E}^1$, by

Lemma 3.6.1,

$$\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{h(x,\zeta_t^x)}{T}=h(x,q^*)\le h(p^*,q^*).$$

We claim that for any $\varepsilon>0$, there is some T' such that $T>T'$ implies that

$$\sum_{t=0}^{T-1}\frac{h(\xi_t,\zeta_t^*)}{T}\le h(p^*,q^*)+\varepsilon. \tag{3.10}$$

Fix some $\varepsilon>0$. Let T_1 be so large that $T>T_1$ implies that, for all $x\in\mathcal{E}^1$,

$$\sum_{t=0}^{T-1}\frac{h(x,\zeta_t^x)}{T}\le h(p^*,q^*)+\frac{\varepsilon}{|X|}, \tag{3.11}$$

and, for all $x\in\mathcal{E}^2$,

$$\frac{C_x}{T}<\frac{\varepsilon}{|X|}. \tag{3.12}$$

Let T' be so large that, for all $x\in\mathcal{E}_1$, $L_x(T')>T_1$. If $T>T'$, then

$$\begin{aligned}\sum_{t=0}^{T-1}\frac{h(\xi_t,\zeta_t^*)}{T}&=\sum_{x\in\mathcal{E}_1}\frac{L_x(T)}{T}\sum_{t=0}^{L_x(T)-1}\frac{h(x,\zeta_t^x)}{L_x(T)}+\sum_{x\in\mathcal{E}_2}\sum_{t=0}^{L_x(T)-1}\frac{h(x,\zeta_t^x)}{T}\\&\le\sum_{i\in\mathcal{E}_1}\frac{L_x(T)}{T}(h(p^*,q^*)+\frac{\varepsilon}{|X|})+\sum_{x\in\mathcal{E}_2}\frac{\varepsilon}{|X|}\le h(p^*,q^*)+\varepsilon.\end{aligned} \tag{3.13}$$

Notice that L_x is weakly increasing, and $L_x(T)\le T$ for all T. Thus, $T>T'$ implies that $L_x(T)\ge L_x(T')>T_1$, and so $T>T_1$.

This proves the inequality (3.10), and it in turn implies that, for any $\varepsilon>0$, there is some T such that

$$\alpha_T=\sup_{T'>T}\sum_{t=0}^{T'-1}\frac{h(\xi_t,\zeta_t^*)}{T'}\le h(p^*,q^*)+\varepsilon.$$

Now, the sequence $\{\alpha_T\}_{T=0}^{\infty}$ is a decreasing sequence, and the above inequality shows that for any $\varepsilon > 0$, $\lim_{T\to\infty} \alpha_T \le h(p^*, q^*) + \varepsilon$. Thus, we have $\lim_{T\to\infty} \alpha_T \le h(p^*, q^*)$. This proves (3.9).

Now, (3.9) implies that

$$(\forall \xi \in \mathcal{X})(\liminf_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t, \zeta_t^*)}{T} \le h(p^*, q^*)). \tag{3.14}$$

Similarly, we can show that

$$(\forall \zeta \in \mathcal{Y})(\limsup_{T\to\infty} \sum_{t=0}^{T-1} \frac{-h(\xi_t^*, \zeta_t)}{T} \le -h(p^*, q^*)), \tag{3.15}$$

This implies that

$$(\forall \zeta \in \mathcal{Y})(\liminf_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t^*, \zeta_t)}{T} \ge h(p^*, q^*)). \tag{3.16}$$

By (3.14), we have for all $\xi \in \mathcal{X}$, $u_h(\xi, \zeta^*) \le h(p^*, q^*)$. By (3.16), we have for all $\zeta \in \mathcal{Y}$, $u_h(\xi^*, \zeta) \ge h(p^*, q^*)$. Therefore, we have

$$u_h(\xi^*, \zeta^*) \le h(p^*, q^*) \le u_h(\xi^*, \zeta^*).$$

This shows that (ξ^*, ζ^*) is an equilibrium.

On the other hand, if $\xi' \in \mathcal{X}$ is a p'-stochastic sequence for some equilibrium strategy p' in g, then $h(p', q^*) = h(p^*, q^*)$. Moreover, all the above argument holds for (ξ', ζ^*) and so ξ' is also an equilibrium strategy. □

Theorem 3.6.2. *Suppose that $\mathcal{P}^1$ and $\mathcal{P}^2$ satisfy (A1) and (A2). Let $(p^*, q^*) \in \Delta(X) \times \Delta(Y)$ be an equilibrium of g. Suppose that there are $\xi \in \mathcal{P}^1$, $\zeta \in \mathcal{P}^2$ such that ξ is p^*-stochastic relative* *to $\mathcal{P}^2$ and ζ is q^*-stochastic relative to $\mathcal{P}^1$. Then, any strongly p-stochastic relative $\mathcal{P}^2$ for some equilibrium mixed strategy $p \in \Delta(X)$ is an equilibrium strategy in $VG(g, \mathcal{P}^1, \mathcal{P}^2)$.*

Proof. For any strategy profile $(a, b) \in \mathcal{X} \times \mathcal{Y}$, let $\theta^{a,b}$ be the sequence of actions obtained by playing this profile. First we show that

$$(\forall a \in \mathcal{X})(\limsup_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\theta_t^{a,b^*})}{T} \leq h(p^*, q^*)), \tag{3.17}$$

where b^* is such that $b^*(\sigma) = \zeta^*_{|\sigma|}$ for all $\sigma \in X^{<\mathbb{N}}$. Because (p^*, q^*) is an equilibrium of the game g, $h(x, q^*) \leq h(p^*, q^*)$ for all $x \in X$.

Suppose that $a \in \mathcal{X}$, and so $a \in \mathcal{P}^1$. For each $x \in X$, let $r^x : Y^{<\mathbb{N}} \to \{0, 1\}$ be the selection function such that $r^x(\sigma) = 1$ if $a(\sigma) = x$, and $r^x(\sigma) = 0$ otherwise. Define

$$L_x(T) = |\{t \in \mathbb{N} : 0 \leq t \leq T-1,\ r^x(\zeta^*[t]) = 1\}| \text{ and } \zeta^x = (\zeta^*)^{r^x}.$$

It is easy to see that r^x is in $\mathcal{P}^1$ since a is. Let

$$\mathcal{E}^1 = \{x \in X : \lim_{T\to\infty} L_x(T) = \infty\} \text{ and } \mathcal{E}^2 = \{x \in X : \lim_{T\to\infty} L_x(T) < \infty\}.$$

For each $x \in \mathcal{E}^2$, let $B_x = \lim_{T\to\infty} L_x(T)$ and let $C_x = \sum_{t=0}^{B_x} h(x, \zeta_t^x)$. Then, for any $x \in \mathcal{E}^1$, by Lemma 3.6.1,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(x, \zeta_t^x)}{T} = h(x, q^*) \leq h(p^*, q^*).$$

We claim that for any $\varepsilon > 0$, there is some T' such that $T > T'$ implies that

$$\sum_{t=0}^{T-1} \frac{h(a(\zeta^*[t]), \zeta_t^*)}{T} \leq h(p^*, q^*) + \varepsilon. \tag{3.18}$$

Fix some $\varepsilon > 0$. Let T_1 be so large that $T > T_1$ implies that, for all $x \in \mathcal{E}^1$,

$$\sum_{t=0}^{T-1} \frac{h(x, \zeta_t^x)}{T} \leq h(p^*, q^*) + \frac{\varepsilon}{|X|}, \tag{3.19}$$

and, for all $x \in \mathcal{E}^2$,

$$\frac{C_x}{T} < \frac{\varepsilon}{|X|}. \tag{3.20}$$

Let T' be so large that, for all $x \in \mathcal{E}_1$, $L_x(T') > T_1$. If $T > T'$, then

$$\begin{aligned}\sum_{t=0}^{T-1} \frac{h(a(\zeta^*[t]), \zeta_t^*)}{T} &= \sum_{x \in \mathcal{E}_1} \frac{L_x(T)}{T} \sum_{t=0}^{L_x(T)-1} \frac{h(x, \zeta_t^x)}{L_x(T)} + \sum_{x \in \mathcal{E}_2} \sum_{t=0}^{L_x(T)-1} \frac{h(x, \zeta_t^x)}{T} \\ &\leq \sum_{i \in \mathcal{E}_1} \frac{L_x(T)}{T}(h(p^*, q^*) + \frac{\varepsilon}{|X|}) + \sum_{x \in \mathcal{E}_2} \frac{\varepsilon}{|X|} \leq h(p^*, q^*) + \varepsilon. \end{aligned} \tag{3.21}$$

Notice that L_x is weakly increasing, and $L_x(T) \leq T$ for all T. Thus, $T > T'$ implies that $L_x(T) \geq L_x(T') > T_1$, and so $T > T_1$.

This proves the inequality (3.18), and it in turn implies that, for any $\varepsilon > 0$, there is some T such that

$$\alpha_T = \sup_{T' > T} \sum_{t=0}^{T'-1} \frac{h(a(\zeta^*[t]), \zeta_t^*)}{T'} \leq h(p^*, q^*) + \varepsilon.$$

Now, the sequence $\{\alpha_T\}_{T=0}^{\infty}$ is a decreasing sequence, and the above inequality shows that for any $\varepsilon > 0$, $\lim_{T\to\infty} \alpha_T \leq h(p^*, q^*) + \varepsilon$. Thus, we have $\lim_{T\to\infty} \alpha_T \leq h(p^*, q^*)$. This proves (3.17).

Now, (3.17) implies that

$$(\forall a \in \mathcal{X})(\liminf_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\theta_t^{a,b^*})}{T} \leq h(p^*, q^*)). \tag{3.22}$$

Similarly, we can show that

$$(\forall b \in \mathcal{Y})(\limsup_{T\to\infty} \sum_{t=0}^{T-1} \frac{-h(\theta_t^{a^*,b})}{T} \leq -h(p^*,q^*)), \tag{3.23}$$

where a^* is such that $a^*(\sigma) = \xi_{|\sigma|}$ for all $\sigma \in Y^{<\mathbb{N}}$. This implies that

$$(\forall b \in \mathcal{Y})(\liminf_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\theta_t^{a^*,b})}{T} \geq h(p^*,q^*)). \tag{3.24}$$

By (3.22), we have for all $a \in \mathcal{X}$, $u_h(a,b^*) \leq h(p^*,q^*)$. By (3.24), we have for all $b \in \mathcal{Y}$, $u_h(a^*,b) \geq h(p^*,q^*)$. Therefore, we have

$$u_h(a^*,b^*) \leq h(p^*,q^*) \leq u_h(a^*,b^*).$$

Therefore, (a^*,b^*) is an equilibrium. $\square$

Theorems 3.6.1 and 3.6.2 are very similar, but neither theorem imply the other.

Now we proceed to the proof of Theorem 3.3.3.

Proof of Theorem 3.3.3: We consider two cases:

Case 1: $\lim_{T\to\infty} \sum_{t=0}^{T-1} r(t)/T = 1$. Suppose that p is not an equilibrium strategy in g. Then, there is an action $y_1 \in Y$ such that $h(p,y) < v$, where v is the equilibrium payoff for player 1 in g. Define ζ as $\zeta_t = y_1$ for all $t \in \mathbb{N}$. Let $1-r$ be the selection function such that $(1-r)(t) = 1-r(t)$ for all $t \in \mathbb{N}$, and let $S(T) = \sum_{t=0}^{T-1} r(t)$ for all $T \in \mathbb{N}$.

Then,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t,\zeta_t)}{T} = \lim_{T\to\infty} \sum_{t=0}^{S(T)-1} \frac{h(\xi_t^r,\zeta_t^r)}{T} + \lim_{T\to\infty} \sum_{t=0}^{T-S(T)-1} \frac{h(\xi_t^{1-r},\zeta_t^{1-r})}{T}.$$

Clearly,

$$\lim_{T\to\infty}\sum_{t=0}^{S(T)-1}\frac{h(\xi_t^r,\zeta_t^r)}{T}=\lim_{T\to\infty}\sum_{t=0}^{S(T)-1}\frac{h(\xi_t^r,y_1)}{S(T)}\lim_{T\to\infty}\frac{S(T)}{T}=h(p,y_1)<v.$$

Let $C=\max_{x\in X,y\in Y}|h(x,y)|$. Then,

$$\sum_{t=0}^{T-S(T)-1}\frac{|h(\xi_t^{1-r},\zeta_t^{1-r})|}{T}\leq\frac{(T-S(T))C}{T}\to 0,$$

and so

$$\lim_{T\to\infty}\sum_{t=0}^{T-S(T)-1}\frac{h(\xi_t^{1-r},\zeta_t^{1-r})}{T}=0.$$

Therefore, ξ is not an equilibrium strategy.

Case 2:: $\lim_{T\to\infty}\sum_{t=0}^{T-1}r(t)/T=\alpha<1$.

First we show that the sequence ξ^{1-r} has limit relative frequency for all $x\in X$. Indeed, for each $x\in X$,

$$\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{c_x(\xi_t^{1-r})}{T}=\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{c_x(\xi_t)(1-r(t))}{T-S(T)}$$

$$=\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{c_x(\xi_t)(1-r(t))}{T}\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{T}{T-S(T)}$$

$$=(\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{c_x(\xi_t)}{T}-\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{c_x(\xi_t)r(t)}{S(T)}\lim_{T\to\infty}\frac{S(T)}{T})\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{T}{T-S(T)}$$

$$=(q_x-\alpha p_x)\frac{1}{1-\alpha},$$

where $q_x=\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{c_x(\xi_t)}{T}$ and $S(T)=\sum_{t=0}^{T-1}r(t)$ for all $T\in\mathbb{N}$.

Suppose that p is not an equilibrium strategy in g. Then, there is an action $y_1\in Y$ such that $h(p,y)<v$, where v is the equilibrium payoff for player 1 in g. Moreover, there is an action $y_2\in Y$

such that $h(\frac{1}{1-\alpha}(q-\alpha p), y_2) \leq v$. Define ζ as follows: $\zeta_t = y_1$ if $r(t) = 1$ and $\zeta_t = y_2$ if $r(t) = 0$.

Then,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t, \zeta_t)}{T} = \lim_{T\to\infty} \sum_{t=0}^{S(T)-1} \frac{h(\xi_t^r, \zeta_t^r)}{T} + \lim_{T\to\infty} \sum_{t=0}^{T-S(T)-1} \frac{h(\xi_t^{1-r}, \zeta_t^{1-r})}{T}$$

$$= \lim_{T\to\infty} \sum_{t=0}^{S(T)-1} \frac{h(\xi_t^r, y_1)}{S(T)} \lim_{T\to\infty} \frac{S(T)}{T} + \lim_{T\to\infty} \sum_{t=0}^{T-S(T)-1} \frac{h(\xi_t^{1-r}, y_2)}{T-S(T)} \lim_{T\to\infty} \frac{T-S(T)}{T}$$

$$= \alpha h(p, y_1) + (1-\alpha) h(\frac{1}{1-\alpha}(q-\alpha p), y_2) < v.$$

Therefore, ξ is not an equilibrium strategy. $\square$

A similar proof works for Theorem 3.4.1.

Proof of Theorem 3.4.1: We consider two cases:

Case 1: $\lim_{T\to\infty} \sum_{t=0}^{T-1} r(\xi[t])/T = 1$. Suppose that p is not an equilibrium strategy in g. Then, there is an action $y_1 \in Y$ such that $h(p, y) < v$, where v is the equilibrium payoff for player 1 in g. Define b as $b(\sigma) = y_1$ for all $\sigma \in X^{<\mathbb{N}}$. Let $1-r$ be the selection function such that $(1-r)(\sigma) = 1 - r(\sigma)$ for all $\sigma \in X^{<\mathbb{N}}$, and let $S(T) = \sum_{t=0}^{T-1} r(\xi[t])$ for all $T \in \mathbb{N}$.

Then,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t, \zeta_t)}{T} = \lim_{T\to\infty} \sum_{t=0}^{S(T)-1} \frac{h(\xi_t^r, \zeta_t^r)}{T} + \lim_{T\to\infty} \sum_{t=0}^{T-S(T)-1} \frac{h(\xi_t^{1-r}, \zeta_t^{1-r})}{T}.$$

Clearly,

$$\lim_{T\to\infty} \sum_{t=0}^{S(T)-1} \frac{h(\xi_t^r, \zeta_t^r)}{T} = \lim_{T\to\infty} \sum_{t=0}^{S(T)-1} \frac{h(\xi_t^r, y_1)}{S(T)} \lim_{T\to\infty} \frac{S(T)}{T} = h(p, y_1) < v.$$

Let $C = \max_{x \in X, y \in Y} |h(x,y)|$. Then,

$$\sum_{t=0}^{T-S(T)-1} \frac{|h(\xi_t^{1-r}, \zeta_t^{1-r})|}{T} \leq \frac{(T-S(T))C}{T} \to 0,$$

and so

$$\lim_{T\to\infty} \sum_{t=0}^{T-S(T)-1} \frac{h(\xi_t^{1-r}, \zeta_t^{1-r})}{T} = 0.$$

Therefore, ξ is not an equilibrium strategy.

Case 2:: $\lim_{T\to\infty} \sum_{t=0}^{T-1} r(\xi[t])/T = \alpha < 1$.

First we show that the sequence ξ^{1-r} has limit relative frequency for all $x \in X$. Indeed, for each $x \in X$,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t^{1-r})}{T} = \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t)(1-r(\xi[t]))}{T-S(T)}$$

$$= \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t)(1-r(\xi[t]))}{T} \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{T}{T-S(T)}$$

$$= (\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t)}{T} - \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t) r(\xi[t])}{S(T)} \lim_{T\to\infty} \frac{S(T)}{T}) \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{T}{T-S(T)}$$

$$= (q_x - \alpha p_x) \frac{1}{1-\alpha},$$

where $q_x = \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_x(\xi_t)}{T}$ and $S(T) = \sum_{t=0}^{T-1} r(\xi[t])$ for all $T \in \mathbb{N}$.

Suppose that p is not an equilibrium strategy in g. Then, there is an action $y_1 \in Y$ such that $h(p,y) < v$, where v is the equilibrium payoff for player 1 in g. Moreover, there is an action $y_2 \in Y$ such that $h(\frac{1}{1-\alpha}(q - \alpha p), y_2) \leq v$. Define $b \in \mathcal{Y}$ as follows: $b(\sigma) = y_1$ if $r(\sigma) = 1$ and $b(\sigma) = y_2$ if $r(\sigma) = 0$.

Then,

$$\lim_{T\to\infty}\sum_{t=0}^{T-1}\frac{h(\theta_t^{\xi,b})}{T}=\lim_{T\to\infty}\sum_{t=0}^{S(T)-1}\frac{h(\xi_t^r,y_1)}{T}+\lim_{T\to\infty}\sum_{t=0}^{T-S(T)-1}\frac{h(\xi_t^{1-r},y_2)}{T}$$

$$=\lim_{T\to\infty}\sum_{t=0}^{S(T)-1}\frac{h(\xi_t^r,y_1)}{S(T)}\lim_{T\to\infty}\frac{S(T)}{T}+\lim_{T\to\infty}\sum_{t=0}^{T-S(T)-1}\frac{h(\xi_t^{1-r},y_2)}{T-S(T)}\lim_{T\to\infty}\frac{T-S(T)}{T}$$

$$=\alpha h(p,y_1)+(1-\alpha)h(\frac{1}{1-\alpha}(q-\alpha p),y_2)<v.$$

Therefore, ξ is not an equilibrium strategy. □

We are left with two more proofs for Theorem 3.3.5 and Theorem 3.3.6.

Proof of Theorem 3.3.5: p^* is an equilibrium mixed strategy, and x is an action such that $p^*_x > 0$. Let $x' \neq x$ be such that $p^*_{x'} > 0$. x' exists because g has no pure strategy equilibrium.

For any real number s, let $\llcorner s \lrcorner$ be the largest integer less than or equal to s. Construct the sequence $\mathbf{p} = (p^0, p^1, ...)$ as follows ($\bar{t}$ is the smallest t such that $\llcorner t^{0.4} \lrcorner > \frac{1}{p^*_x}$):

(a) $p^t_y = p^*_y$ if $y \neq x$ and $y \neq x'$;

(b) $p^t_x = p^*_x$ if $t \leq \bar{t}$ and $p^t_x = p^*_x - \frac{1}{\llcorner t^{0.4} \lrcorner}$ otherwise;

(c) $p^t_{x'} = p^*_{x'}$ if $t \leq \bar{t}$ and $p^t_{x'} = p^*_{x'} + \frac{1}{\llcorner t^{0.4} \lrcorner}$ otherwise.

By construction, $p^t_x = 0$ if and only if $p^*_x = 0$, and $\lim_{t\to\infty} p^t = p^*$. Clearly, $\mathbf{p}$ is computable. By Theorem 3.7.2, there is a $\mu_{\mathbf{p}}$-random sequence ξ relative to $\mathcal{P}^2$ in $\mathcal{P}^1$. Now, let $X_0 = \{y \in X : p_y > 0\}$, then the sequence ξ is $\mu_{\mathbf{p}}$-random can be regarded as a sequence in $X_0^{\mathbb{N}}$. Therefore, Theorem 3.7.6 is applicable and so ξ is p^*-stochastic. Thus, by Theorem 3.3.2, ξ is an equilibrium strategy for player 1 in $HG(g, \mathcal{P}^1, \mathcal{P}^2)$.

By Theorem 3.7.7, we have

$$\limsup_{T\to\infty} \frac{|\sum_{t=0}^{T-1}(c_x(\xi_t) - p_x^t)|}{\sqrt{2(\sum_{t=0}^{T-1} p_x^t(1-p_x^t))\log\log\sqrt{(\sum_{t=0}^{T-1} p_x^t(1-p_x^t))}}} = 1. \tag{3.25}$$

For any T large enough,

$$\frac{\sum_{t=0}^{T-1} c_x(\xi_t) - Tp_x^*}{\sqrt{2Tp_x^*(1-p_x^*)\log\log T}} = \frac{\sum_{t=0}^{T-1}(c_x(\xi_t) - p_x^t)}{\sqrt{2Tp_x^*(1-p_x^*)\log\log T}} + \frac{\sum_{t=\bar{t}+1}^{T-1} \frac{1}{\llcorner t^{0.4} \lrcorner}}{\sqrt{2Tp_x^*(1-p_x^*)\log\log T}}$$

We claim that

$$\lim_{T\to\infty} \frac{\sum_{t=\bar{t}+1}^{T-1} \frac{1}{\llcorner t^{0.4} \lrcorner}}{\sqrt{2Tp_x^*(1-p_x^*)\log\log T}} = \infty; \tag{3.26}$$

and there exists some $B > 0$ such that for all T large enough,

$$\frac{|\sum_{t=0}^{T-1}(c_x(\xi_t) - p_x^t)|}{\sqrt{2Tp_x^*(1-p_x^*)\log\log T}} < B. \tag{3.27}$$

The conclusion follows directly from (3.26) and (3.27).

For all t, we have $\llcorner t^{0.4} \lrcorner \le t^{0.4}$ and so $\frac{1}{t^{0.4}} \le \frac{1}{\llcorner t^{0.4} \lrcorner}$. It is easy to check that

$$\sum_{t=1}^{T-1} \frac{1}{\llcorner t^{0.4} \lrcorner} \ge \sum_{t=1}^{T-1} \frac{1}{t^{0.4}} \ge \int_{x=1}^{T-1} x^{-0.4} dx - 1 \ge (T-1)^{0.6} - 2.$$

Therefore, for T large enough,

$$\frac{\sum_{t=\bar{t}+1}^{T-1} \frac{1}{\llcorner t^{0.4} \lrcorner}}{\sqrt{2Tp_x^*(1-p_x^*)\log\log T}} \ge \frac{0.5T^{0.6}}{\sqrt{2Tp_x^*(1-p_x^*)\log\log T}} = A^0 \frac{T^{0.1}}{\sqrt{\log\log T}} \tag{3.28}$$

for some constant $A^0 > 0$. Because $\lim_{T\to\infty} \frac{T^{0.1}}{\sqrt{\log\log T}} = \infty$, (3.28) implies that (3.26).

Because of (3.25), to prove (3.27), it suffices to show that for T large enough,

$$\frac{\sqrt{2(\sum_{t=0}^{T-1} p_x^t(1-p_x^t))\log\log\sqrt{(\sum_{t=0}^{T-1} p_x^t(1-p_x^t))}}}{\sqrt{2Tp_x^*(1-p_x^*)\log\log T}} \tag{3.29}$$

is bounded. Now, for T large enough,

$$\sum_{t=0}^{T-1} p_x^t(1-p_x^t)) = Tp_x^*(1-p_x^*) + (2p_x^*-1)\sum_{t=\bar{t}+1}^{T-1}\frac{1}{\llcorner t^{0.4}\lrcorner} - \sum_{t=\bar{t}+1}^{T-1}(\frac{1}{\llcorner t^{0.4}\lrcorner})^2. \tag{3.30}$$

Because for t large enough, $\frac{1}{2}t^{0.4} < \llcorner t^{0.4}\lrcorner$, there is a constant $A > 0$ such that

$$\sum_{t=\bar{t}+1}^{T-1}\frac{1}{\llcorner t^{0.4}\lrcorner} < \sum_{t=\bar{t}+1}^{T-1}\frac{2}{t^{0.4}} + A < 2T^{0.6} + A.$$

Similarly, there is a constant $A' > 0$ such that

$$\sum_{t=\bar{t}+1}^{T-1}(\frac{1}{\llcorner t^{0.4}\lrcorner})^2 < \sum_{t=\bar{t}+1}^{T-1}\frac{4}{t^{0.8}} + A' < 4T^{0.2} + A'.$$

Hence,

$$\frac{\sum_{t=0}^{T-1} p_x^t(1-p_x^t)}{Tp_x^*(1-p_x^*)} < 1 + \frac{2|2p_x^*-1|}{T^{0.4}p_x^*(1-p_x^*)} + \frac{|2p_x^*-1|A + A'}{Tp_x^*(1-p_x^*)} + \frac{4}{T^{0.8}p_x^*(1-p_x^*)},$$

and so, for T large enough,

$$\frac{\sum_{t=0}^{T-1} p_x^t(1-p_x^t))}{Tp_x^*(1-p_x^*)} < 2.$$

Equation (3.30) also implies that, for T large enough,

$$\sum_{t=0}^{T-1} p_x^t(1-p_x^t) \leq (2|2p_x^*-1| + 5 + p_x^*(1-p_x^*))T = A''T,$$

and hence

$$\frac{\log\log\sqrt{(\sum_{t=0}^{T-1}p_x^t(1-p_x^t))}}{\log\log T} \leq \frac{\log(\log T+\log A'')}{\log\log T} \leq \frac{\log 2+\log\log T}{\log\log T}.$$

So for T large enough,

$$\frac{\log\log\sqrt{(\sum_{t=0}^{T-1}p_x^t(1-p_x^t))}}{\log\log T} \leq 2.$$

Thus, the expression in (3.29) is bounded by 2, and we have proved (3.27). □

Proof of Theorem 3.3.6: Let $p \in \Delta(X)$. Since p is not degenerate, there are $x_1 \neq x_2 \in X$ such that $p_{x_1} > 0$ and $p_{x_2} > 0$. Let ξ be a strategy for player 1 in HG that is μ_p-random relative to $\mathcal{P}^2$, and let ξ' be a strategy for player 2 in HG that is μ_q-random relative to $\mathcal{P}^1$. Construct a new sequence $\zeta \in (X \times X)^{\mathbb{N}}$ as follows: $\zeta_t = (\xi_{2t}, \xi_{2t+1})$ for all $t \in \mathbb{N}$. It is easy to check that ζ is $\mu_{p\otimes p}$-random. Define $\theta : \mathbb{N} \to \mathbb{N}$ as follows:

(a) $\theta(0) = \min\{t : \zeta_t = (x_1, x_2) \vee \zeta_t = (x_2, x_1)\}$;

(b) for $t > 0$, $\theta(t) = \min\{t' : t' > \theta(t-1) \wedge (\zeta_{t'} = (x_1, x_2) \vee \zeta_{t'} = (x_2, x_1))\}$.

θ is total since ζ is $\mu_{p\otimes p}$-random and $p_{x_1}p_{x_2} > 0$. Define ξ^1 as $\xi^1_t = 0$ if $\zeta_{\theta(t)} = (x_1, x_2)$ and $\xi^1_t = 1$ if $\zeta_{\theta(t)} = (x_2, x_1)$ for all $t \in \mathbb{N}$. By Theorem 3.7.4, it is easy to see that ξ^1 is $\mu_{(\frac{1}{2},\frac{1}{2})}$-random.

Clearly, $\xi^1 \in \mathcal{P}^1$, and so by Theorem 3.7.3, ξ' is μ_q-random relative to ξ^1, and, hence, by Theorem 3.7.3 again, ξ^1 is $\mu_{(\frac{1}{2},\frac{1}{2})}$-random relative to $\mathcal{P}^2$. By Theorem 3.7.1, ξ^1 is a complex sequence relative to $\mathcal{P}^2$. The existence of ξ^2 is completely symmetric. □

Finally, we give the proof of Proposition 3.3.1.

Proof of Proposition 3.3.1: By Theorem 3.7.1, we know that, for any $\xi, \zeta \in \{0,1\}^{\mathbb{N}}$, if ξ is $\mu_{(\frac{1}{2},\frac{1}{2})}$-random relative to $C(\zeta)$, then ξ is complex relative to $C(\zeta)$. Moreover, we know that, by Theorem 3.7.3, if $\xi \otimes \zeta$ is λ^4-random, then ξ is $\mu_{(\frac{1}{2},\frac{1}{2})}$-random relative to $C(\zeta)$ and vice versa.

Hence, $C(\xi)$ and $C(\zeta)$ are mutually complex. Now, by Proposition 3.7.1, the set

$$A = \{\xi \otimes \zeta : \xi \otimes \zeta \text{ is } \lambda^4\text{-random}\} \subset \{\xi \otimes \zeta : C(\xi) \text{ and } C(\zeta) \text{ are mutually complex}\}$$

has measure 1. Therefore, the set A is uncountable. Since for any ξ, the set of sequences ξ' such that $C(\xi) = C(\xi')$ is countable, we can then conclude that there are uncountably many different pairs of computability constraints that are mutually complex. $\square$

3.7 Appendix

3.7.1 Recursive functions

In this section we introduce $C(\xi)$, the set of functions computable in ξ, for any $\xi \in \{0,1\}^{\mathbb{N}}$.

Definition 3.7.1. Let $\xi \in \{0,1\}^{\mathbb{N}}$, which can be identified as a total function in $\mathcal{F}$. Define $C(\xi)$ to be the minimal class of functions in $\mathcal{F}$ that satisfies the following conditions:

(a) $C(\xi)$ contains functions ξ, Z, S, and P_{ki} for all $k, i > 0$.

(b) $C(\xi)$ is closed under composition, primitive recursion, and minimization.

If $f \in C(\xi)$, we say that f is *ξ-computable*. We also say that f is *computable* if $f \in \mathcal{T}$, where $\mathcal{T}$ is the smallest set that satisfies (a) and (b) without ξ in (a). It is well known that for any set $\mathcal{P}$ satisfying (A1) and (A2) there exists a sequence ξ such that $\mathcal{P} = C(\xi)$ (c.f. Pippenger [1997]).

For any finite set X, we can also consider the space $X^{\mathbb{N}}$, which is called the Cantor space. Any element in $X^{\mathbb{N}}$ is called a *Turing oracle*. In the same manner, we can also consider functions that are ξ-computable for any $\xi \in X^{\mathbb{N}}$. Clearly $C(\xi)$ satisfies (A1) and (A2). Conversely, for any set $\mathcal{P}$ satisfying (A1) and (A2), there is an oracle ξ such that $\mathcal{P}$ is the set of functions that can be computed by a Turing machine with oracle ξ. For a detailed discussion and proofs, please see Pippenger [1997], chapter 4. For any such a machine, we can give it a Gödel number, and we use $\varphi_e^{(k),\xi}$ to denote the function that is computed by the machine with Gödel number e and with oracle ξ, and the index k indicates that we put k numbers as input in the computation. We shall use this enumeration of functions to analyze the tests defined for randomness in section 3.3.

Consider a sequence $\{V_t\}_{t=0}^{\infty}$ of subsets of $X^{\mathbb{N}}$. Suppose there is a total function $f : \mathbb{N} \to \mathbb{N} \times X^{<\mathbb{N}}$ in $C(\xi)$ such that for all $t \in \mathbb{N}$ and for all $\zeta \in X^{\mathbb{N}}$,

$$\zeta \in V_t \Leftrightarrow (\exists n)(f(n) = (t, \sigma) \wedge \sigma = \zeta[|\sigma|]). \tag{3.31}$$

Then, we can find a total function $h \in \mathcal{P}^*$ such that[8]

$$\zeta \in V_t \Leftrightarrow \varphi_{h(t)}^{(1),\xi\oplus\zeta}(0) \downarrow, \tag{3.32}$$

where for any partial function f, $f(x_1, ..., x_k) \downarrow$ holds iff $f(x_1, ..., x_k) \neq \perp$.

In the following, we shall also consider functionals $\Psi : X^{\mathbb{N}} \times \mathbb{N}^2 \to \mathbb{N}$. Such a functional Ψ is ξ-*computable* if there is a number e such that for all ζ and for all x, y,

$$\Psi(\zeta, x, y) \simeq \varphi_e^{(2),\xi\oplus\zeta}(x, y).$$

It can also be shown that, for any sequence $\{V_t\}_{t=0}^{\infty}$, there is some $f \in C(\xi)$ such that (3.31) holds if and only if there is a computable functional such that

$$\zeta \in V_t \Leftrightarrow (\exists s)(\Psi(\xi, t, s) = 0). \tag{3.33}$$

[8]This follows directly from the Parametrization Theorem for relative computability. Please see Downey et al. [2006] for a more detailed discussion.

If there is such a ξ-computable that the above relation holds, then we say that $\{V_t\}$ is of $\Sigma_1^{0,\xi}$. In case that Ψ is computable, we say that $\{V_t\}$ is of Σ_1^0.

3.7.2 Effective randomness

In this section we review some results from effective randomness. We include this section mainly for self-containment. We will not give all the proofs, but will refer the readers to the survey paper Downey et al. [2006].

3.7.2.1 Transformations

There is a close connection between complex sequences and random sequences: any complex sequences is random with respect to the uniform distribution. The following theorem is well-known in this literature. The proof of this theorem with $\mathcal{P} = \mathcal{T}$ can be found in Downey et al. [2006], and all the arguments there can be relativized and the proof is general enough to cover the general case.

Theorem 3.7.1. *Let $\mathcal{P}$ satisfy (A1) and (A2). Then, $\xi \in \{0,1\}^{\mathbb{N}}$ is a complex sequence relative to $\mathcal{P}$ if and only if ξ is a λ-random sequence relative to $\mathcal{P}$, where $\lambda(N_\sigma) = 2^{-|\sigma|}$ for all $\sigma \in \{0,1\}^{<\mathbb{N}}$.*

Our first task is to show that if $\mathcal{P}^1$ includes a complex sequence relative to $\mathcal{P}^2$, then it also includes a μ_p-random sequence relative to $\mathcal{P}^2$. As the following theorem shows, in fact, a slightly stronger result holds. The proof for the case $|X| = 2$ can be found in Zvonkin and Levin [1970]. We follow a similar logic.

Theorem 3.7.2. *Let $\mathcal{P}$ satisfy (A1) and (A2). Let X be a finite set. Suppose that $\xi \in \{0,1\}^{\mathbb{N}}$ is a complex sequence relative to $\mathcal{P}$. Let $\mathbf{p} = \{p^t\}_{t=0}^{\infty}$ be a sequence over $\Delta(X)$ that is computable, i.e., is in $\mathcal{T}$. Then, there is a $\mu_{\mathbf{p}}$-sequence $\zeta \in X^{\mathbb{N}}$ relative to $\mathcal{P}$ belonging to $C(\xi)$.*

Proof. First we claim that there is a λ^X-random sequence $\xi' \in X^{\mathbb{N}}$ relative to $\mathcal{P}$ that is in $C(\xi)$, where $\lambda^X(N_\sigma) = |X|^{-|\sigma|}$ for all $\sigma \in X^{<\mathbb{N}}$. For a proof, see Calude [2002], Theorem 7.18.

There is a natural mapping Γ between $X^{\mathbb{N}}$ and $[0,1]$:

$$\Gamma(\zeta) = \sum_{t=0}^{\infty} \iota(\zeta_t)\frac{1}{n^{t+1}},$$

where $X = \{x_1, ..., x_n\}$ and $\iota(x) = i - 1$ if and only if $x = x_i$. Γ is onto but not one-to-one. However, the set $\{\zeta \in X^{\mathbb{N}} : \Gamma(\zeta) = \Gamma(\zeta') \text{ for some } \zeta' \neq \zeta\}$ is countable, since for any such ζ, $\Gamma(\zeta)$ is a rational number.

If $\Phi : X^{\mathbb{N}} \to X^{\mathbb{N}}$ is a (Borel) measurable function, then λ^X_Φ defined as $\lambda^X_\Phi(A) = \lambda^X(\Phi^{-1}(A))$ is also a measure over $X^{\mathbb{N}}$. We will construct a computable mapping Φ such that Φ maps a λ^X-random sequence relative to $\mathcal{P}$ to a $\mu_{\mathbf{p}}$-random sequence relative to $\mathcal{P}$.

We will define Φ via a monotone function $\phi : S \to T$, where S is a and T are trees over X (a tree is a subset of $X^{<\mathbb{N}}$ that is closed under initial segments). Such a function is monotone if $\sigma \subset \tau$ implies $\phi(\sigma) \subset \phi(\tau)$. If S is a tree, $[S] = \{\zeta \in X^{\mathbb{N}} : (\forall t)\zeta[t] \in S\}$. Given such a function, let

$$D(\phi) = \{\zeta \in [S] : \lim_{t\to\infty} \phi(\zeta[t]) = \infty\}.$$

Then, define $\Phi : D(\phi) \to X^{\mathbb{N}}$ by $\Phi(\zeta) = \bigcup_{t=0}^{\infty} \phi(\zeta[t])$. We say that Φ is *obtained from ϕ via a monotone function.*

We claim that there exists a computable monotone function ϕ such that $\mu_{\mathbf{p}} = \lambda^X_\Phi$ and $\lambda^X(D(\phi)) = 1$, where $\Phi : D(\phi) \to X^{\mathbb{N}}$ is obtained from ϕ via a monotone function.

Extend Γ to $X^{<\mathbb{N}}$ as $\Gamma(\sigma) = \sum_{t=0}^{|\sigma|-1} \frac{\iota(\sigma_t)}{n^{t+1}}$. Define $g : [0,1] \to [0,1]$ as

$$g(r) = \mu_{\mathbf{p}}(\{\zeta : \Gamma(\zeta) \leq r\}),$$

i.e., the distribution function of $\mu_{\mathbf{p}}$ over $[0,1]$. Define $h = g^{-1}$. Therefore, $r \leq g(s)$ if and only if $h(r) \leq s$. Then,

$$\mu_{\mathbf{p}}(\Gamma^{-1}([0,r])) = g(r) = \lambda^X(\Gamma^{-1}([0,g(r)])) = \lambda^X(\Gamma^{-1}(h^{-1}([0,r]))).$$

Define $g^0(\epsilon) = 0$ and $g^1(\epsilon) = 1$, and, for $\tau \in X^{<\mathbb{N}} - \{\epsilon\}$, define

$$g^0(\tau) = \sum\{\mu_{\mathbf{p}}(N_\sigma) : \Gamma(\sigma) \leq \Gamma(\tau) - \frac{1}{n^{|\tau|}}, |\sigma| = |\tau|\},$$

where $\sum \emptyset = 0$, and

$$g^1(\tau) = \sum\{\mu_{\mathbf{p}}(N_\sigma) : \Gamma(\sigma) \leq \Gamma(\tau), |\sigma| = |\tau|\}.$$

For any $\zeta \in X^{\mathbb{N}}$,

$$\Gamma(\zeta) \leq \Gamma(\tau) \Leftrightarrow \Gamma(\zeta[|\tau|]) \leq \Gamma(\tau) - \frac{1}{n^{|\tau|}} \vee \Gamma(\zeta) = \Gamma(\tau),$$

and

$$\Gamma(\zeta) \leq \Gamma(\tau) + \frac{1}{n^{|\tau|}} \Leftrightarrow \Gamma(\zeta[|\tau|]) \leq \Gamma(\tau) \vee \Gamma(\zeta) = \Gamma(\tau) + \frac{1}{n^{|\tau|}}.$$

Since $\mu_{\mathbf{p}}$ has no atoms, we have that

$$g^0(\tau) = g(\Gamma(\tau)) \text{ and } g^1(\tau) = g(\Gamma(\tau) + \frac{1}{n^{|\tau|}}).$$

Therefore, for each $t > 0$, the class of intervals

$$\{[g^0(\tau), g^1(\tau)] : \tau \in X^{<\mathbb{N}}, |\tau| = t\} \tag{3.34}$$

forms a partition of $[0,1]$.

Construct ϕ as follows: Given a string $\sigma \in X^{\mathbb{N}}$, let

$$a_\sigma = \Gamma(\sigma) \text{ and } b_\sigma = \Gamma(\sigma) + \frac{1}{n^{|\sigma|}}.$$

Let $\phi(\sigma)$ be the longest τ with $|\tau| \leq |\sigma|$ such that $[a_\sigma, b_\sigma] \subset [g^0(\tau), g^1(\tau)]$. Now, for any $\sigma \in X^{<\mathbb{N}}$, $\phi(\sigma)$ is well-defined, since the intervals in (3.34) forms a partition and $[g^0(\epsilon), g^1(\epsilon)] = [0, 1]$. To see that ϕ is monotone, suppose that $\sigma \subset \sigma'$ and $\tau = \phi(\sigma)$, $\tau' = \phi(\sigma')$. It is easy to check that $a_\sigma \leq a_{\sigma'}$ and $b_{\sigma'} \leq b_\sigma$. Now, if $\Gamma(\tau') \geq \Gamma(\tau) + \frac{1}{n^{|\tau|}}$, then

$$a_{\sigma'} \geq g^0(\tau') = g(\Gamma(\tau')) \geq g(\Gamma(\tau) + \frac{1}{n^{|\tau|}}) = g^1(\tau) \geq b_\sigma \geq b_{\sigma'},$$

a contradiction. On the other hand, we have $|\tau'| \geq |\tau|$. If $\Gamma(\tau') < \Gamma(\tau)$, then $\Gamma(\tau') \leq \Gamma(\tau) - \frac{1}{n^{|\tau|}}$, and hence,

$$b_\sigma \leq b_{\sigma'} \leq g^1(\tau') = g(\Gamma(\tau') + \frac{1}{n^{|\tau'|}}) \leq g(\Gamma(\tau)) = g^0(\tau) \leq a_\sigma,$$

a contradiction. Thus, we have $\tau \subset \tau'$.

Then we show that $X^{\mathbb{N}} - \{\zeta \in X^{\mathbb{N}} : h(\Gamma(\zeta)) = \frac{m}{n^t} \text{ for some } m, n, t \in \mathbb{N}\} \subseteq D(\phi)$. Suppose that $h(\Gamma(\zeta)) \neq \frac{m}{n^t}$ for any $m, n, t \in \mathbb{N}$. Let K be given. There exists some $l \in \mathbb{N}$ such that $h(\Gamma(\zeta)) \in (\frac{l}{n^K}, \frac{l+1}{n^K})$. Let

$$\varepsilon = \min\{h(\Gamma(\zeta)) - \frac{l}{n^K}, \frac{l+1}{n^K} - h(\Gamma(\zeta))\}.$$

Since h is continuous, there is some T such that $t \geq T$ implies that

$$\min\{|h(b_{\zeta[t]}) - h(\Gamma(\zeta))|, |h(\Gamma(\zeta) - h(a_{\zeta[t]})|\} \leq \frac{\varepsilon}{2} \text{ and so } [h(a_{\zeta[t]}), h(b_{\zeta[t]})] \subseteq (\frac{l}{n^K}, \frac{l+1}{n^K}).$$

Thus, if $t \geq \max\{T, K\}$, then

$$[a_{\zeta[t]}, b_{\zeta[t]}] \subset [g(\frac{l}{n^K}), g(\frac{l+1}{n^K})] = [g^0(\frac{l}{n^K}), g^1(\frac{l}{n^K})],$$

and so $|\phi(\zeta[t])| \geq K$. Therefore, $\lambda^X(D(\phi)) = 1$.

Now, we claim that if $\zeta \in D(\phi)$, then $\Gamma(\Phi(\zeta)) = h(\Gamma(\zeta))$. Let ε be given, and let K be so large that $\varepsilon < \frac{1}{n^{K-1}}$. Since $\zeta \in D(\phi)$, there exists T such that $t \geq T$ implies that $|\phi(\zeta[t])| \geq K$. Then, for all $t \geq T$,

$$h(\Gamma(\zeta)) \in [h(a_{\zeta[t]}), h(b_{\zeta[t]})] \subseteq [a_{\phi(\zeta[t])}, b_{\phi(\zeta[t])}],$$

and so

$$h(\Gamma(\zeta)) - \Gamma(\phi(\zeta[t])) \leq \frac{1}{n^K} \leq \varepsilon.$$

Thus,

$$\Gamma(\Phi(\zeta)) = \lim_{t\to\infty} \Gamma(\phi(\zeta[t])) = h(\Gamma(\zeta)).$$

Moreover, for any $r \in [0,1]$, there is a sequence $\zeta \in X^{\mathbb{N}}$ such that $\Gamma(\Phi(\zeta)) = r$, since h is strictly increasing and is continuous. Also, we have that

$$\Gamma(\Phi(\zeta)) \geq \Gamma(\Phi(\zeta')) \Leftrightarrow \Gamma(\zeta) \geq \Gamma(\zeta').$$

Now, we show that $\lambda^X_\Phi = \mu_{\mathbf{p}}$ by demonstrating that they share the same distribution function:

$$\lambda^X_\Phi(\{\zeta : \Gamma(\zeta) \leq \Gamma(\Phi(\zeta^*))\}) = \lambda^X(\{\zeta : \Gamma(\Phi(\zeta)) \leq \Gamma(\Phi(\zeta^*))\})$$

$$= \lambda^X(\{\zeta : \Gamma(\zeta) \leq \Gamma(\zeta^*)\}) = \Gamma(\zeta^*) = g(\Gamma(\Phi(\zeta^*))).$$

Notice that if $\Gamma(\zeta) = g(\frac{m}{n^t}) \in \mathbb{Q}$, then ζ is computable. Thus, $\xi' \in D(\phi)$. Let $\zeta' = \Phi(\xi')$. Now we show that ζ' is $\mu_{\mathbf{p}}$-random relative to $\mathcal{P}$. Suppose not. Then there is a $\mu_{\mathbf{p}}$-test $\{V_t\}_{t=0}^{\infty}$ relative to $\mathcal{P}$ such that $\zeta' \in \bigcap_{t=0}^{\infty} V_t$. Let $U_t = \Phi^{-1}(V_t)$. Then, for all t, $\lambda^X(U_t) = \mu_{\mathbf{p}}(V_t) \leq \frac{1}{2^t}$. Moreover, since ϕ is computable, $\{U_t\}_{t=0}^{\infty}$ is a λ^X-test relative to $\mathcal{P}$. But $\xi' \in \bigcap_{t=0}^{\infty} U_t$ since $\zeta' \in \bigcap_{t=0}^{\infty} V_t$, a contradiction. Since ϕ is computable, $\zeta' \in C(\xi') \subset C(\xi)$. □

We give an existence theorem for random sequences. Its proof for the case with $\mathcal{P} = \mathcal{T}$ can be found in Martin-Löf [1966] (with some minor modifications to accommodate general computable measures). See also Downey et al. [2006].

Proposition 3.7.1. *Suppose that X is a finite set and μ is a computable measure over $X^{\mathbb{N}}$. Let $\mathcal{P}$ satisfy (A1) and (A2). Then*

$$\mu(\{\xi \in X^{\mathbb{N}} : \xi \textit{ is } \mu\textit{-random relative to } \mathcal{P}\}) = 1.$$

3.7.2.2 Independence

Next, we shall consider independence. This concept is crucial to prove Theorem 3.3.6. First we give some notations. For any distribution $(p, q) \in \Delta(X) \times \Delta(Y)$, we define $p \otimes q$ to be the product measure of them over $X \times Y$. For any two sequences $\xi \in X^{\mathbb{N}}$ and $\zeta \in Y^{\mathbb{N}}$, we define $\xi \otimes \zeta$ as $(\xi \otimes \zeta)_t = (\xi_t, \zeta_t)$ for all $t \in \mathbb{N}$. In the axiomatic probability theory, independence of random variables is defined in terms of product measures: a random variable on X and a random variable on Y are independent in the standard theory if their joint distribution is a product distribution over $X \times Y$. Similar to Definition 3.3.3, we can define $\mu_{p \otimes q}$-randomness (relative to $\mathcal{P}^*$) in $(X \times Y)^{\mathbb{N}}$, where for all $x \in X$, $y \in Y$, $p \otimes q_{(x,y)} = p_x q_y$. The following theorem, essentially due to van Lambalgen van Lambalgen [1990], characterizes independence in terms of randomness with respect to product measures, which establishes a connection between our definition of independence and the measure theoretical definition. The proof for the case $|X| = |Y| = 2$ and $p = (\frac{1}{2}, \frac{1}{2}) = q$ can be found in Downey et al. [2006], Theorem 12.12, 12.13.

Theorem 3.7.3. *Consider two finite sets X and Y. Suppose $\xi \in X^{\mathbb{N}}$ and $\zeta \in Y^{\mathbb{N}}$, and suppose $p \in \Delta(X)$ and $q \in \Delta(Y)$.*

(a) If $\xi \otimes \zeta$ is $\mu_{p\otimes q}$-random, then ξ is μ_p-random relative to $C(\zeta)$.

(b) If ξ is μ_p-random relative to $C(\zeta)$ and ζ is μ_q-random, then $\xi \otimes \zeta$ is $\mu_{p\otimes q}$-random.

Proof. (a) Suppose that ξ is not p-random relative to $C(\zeta)$. Then $\xi \in \bigcap_{t=0}^{\infty} U_t^{\zeta}$ for some uniformly μ_p-test $\{U_t^{\zeta}\}_{t=0}^{\infty}$ relative to $C(\zeta)$ in $X^{\mathbb{N}}$ such that $\mu_p(U_t^{\zeta}) \leq \frac{1}{2^t}$. Since it is a test, by (3.32), let h be a total computable function such that $\xi' \in U_t^{\zeta}$ if and only if $\varphi_{h(t)}^{(1),\zeta\oplus\xi'}(0)\downarrow$. For any $\zeta' \in Y^{\mathbb{N}}$, define

$$U_{t,s}^{\zeta'} = \{\xi' \in X^{\mathbb{N}} : \varphi_{h(t)}^{(1),(\zeta'\oplus\xi')[2s]}(0)\downarrow\},\ U_t^{\zeta'}[\frac{1}{2^t}] = \bigcup_{\mu_p(U_{t,s}^{\rho})\leq\frac{1}{2^t}} U_{t,s}^{\zeta'}. \tag{3.35}$$

Let

$$V_t = \{\xi' \otimes \zeta' : \xi' \in X^{\mathbb{N}},\ \zeta' \in Y^{\mathbb{N}},\ \xi' \in U_t^{\zeta'}[\frac{1}{2^t}]\}. \tag{3.36}$$

We claim that $\{V_t\}_{t=0}^{\infty}$ is a $\mu_{p\otimes q}$-test.

Now, by (3.35) and (3.36), $\xi' \otimes \zeta' \in V_t$ if and only if $(\exists s)(\mu_p(U_{t,s}^{\zeta'}) \leq \frac{1}{2^t} \wedge \xi' \in U_{t,s}^{\zeta'})$ if and only if $(\exists s)(\mu_p(U_{t,s}^{\zeta'}) \leq \frac{1}{2^t} \wedge \varphi_{h(t)}^{(1),\zeta'\oplus\xi'[2s]}(0)\downarrow)$. We claim that the predicates $\varphi_{h(t)}^{(1),\zeta'\oplus\xi'[2s]}(0)\downarrow$ and $\mu_p(U_{t,s}^{\zeta'}) \leq \frac{1}{2^t}$ are computable.

(a.1) The functional $(\xi'\otimes\zeta', s) \mapsto \zeta'\oplus\xi'[2s]$ is computable in $(X\times Y)^{\mathbb{N}}\times\mathbb{N}$, and so is the predicate $\varphi_{h(t)}^{(1),\sigma}(0)\downarrow$ in $\mathbb{N}^2$ (in t and σ). Thus, by generalized composition, the predicate $\varphi_{h(t)}^{(1),\zeta'\oplus\xi'[2s]}(0)\downarrow$ is computable in $(X\times Y)^{\mathbb{N}}\times\mathbb{N}^2$ (in $\xi'\otimes\zeta'$, t, and s).

(a.2) It is easy to check that

$$U_{t,s}^{\zeta'} = \bigcup \{N_\sigma : \varphi_{h(t)}^{(1),(\zeta'[s]\oplus\sigma)}(0)\downarrow, |\sigma| = s\},$$

and

$$\mu_p(U_{t,s}^{\zeta'}) = \sum_\sigma \{\prod_{j=0,\ldots,s-1} p_{\sigma_j} : \varphi_{h(t)}^{(1),(\zeta'[s]\oplus\sigma)}(0)\downarrow \wedge |\sigma| = s\}.$$

The functional $(\xi' \otimes \zeta', s) \mapsto \zeta'[s]$ is computable in $(X \times Y)^{\mathbb{N}} \times \mathbb{N}$ and so is the predicate $\sum_\sigma\{\prod_{j=0,\ldots,s-1} p_{\sigma_j} : \varphi_{h(t)}^{(1),(\sigma'\oplus\sigma)}(0)\downarrow \wedge |\sigma| = s\}$ in $\mathbb{N}^3$ (in σ', t, and s). By generalized composition, the functional $(\xi' \otimes \zeta', s) \mapsto \mu_p(U_{t,s}^{\zeta'})$ is computable, and so the predicate $\mu_p(U_{t,s}^{\zeta'}) \leq \frac{1}{2^t}$ is computable.

Thus, there is a f in $\mathcal{P}^*$ such that (3.31) holds. Moreover,

$$\begin{aligned}
\mu_{p\otimes q}(V_t) &= \int_{(X\times Y)^{\mathbb{N}}} \chi_{V_t}(\xi' \otimes \zeta') d\mu_{p\otimes q}(\xi' \otimes \zeta') \\
&= \int_{Y^{\mathbb{N}}} \int_{X^{\mathbb{N}}} \chi_{U_t^{\zeta'}[\frac{1}{2^t}]}(\xi') d\mu_p(\xi') d\mu_q(\zeta') \\
&= \int_{Y^{\mathbb{N}}} \mu_p(U_t^{\zeta'}[\tfrac{1}{2^t}]) d\mu_q(\zeta') \leq \tfrac{1}{2^t}.
\end{aligned}$$

$\{V_t\}_{t=0}^\infty$ is a $\mu_{p\otimes q}$-test. But $\xi \otimes \zeta \in V_t$ for all $t \in \mathbb{N}$, and so $\xi \otimes \zeta$ is not $\mu_{p\otimes q}$-random.

(b) Suppose that $\xi \otimes \zeta \in (X \times Y)^{\mathbb{N}}$ is not $\mu_{p\otimes q}$-random. Then, $\xi \otimes \zeta \in \bigcap_{t=0}^\infty U_t$ for some $\mu_{p\otimes q}$-test $\{U_t\}$ in $(X \times Y)^{\mathbb{N}}$ such that $\mu_{p\otimes q}(U_t) \leq \frac{1}{4^t}$. Suppose that $\xi' \otimes \zeta' \in U_t$ if and only if $\varphi_{h(t)}^{(1);\xi'\otimes\zeta'}(0)\downarrow$, where h is a total and computable function. Let $\zeta' \in Y^{\mathbb{N}}$, and define

$$V_t^{\zeta'} = \{\xi' \in X^{\mathbb{N}} : \xi' \otimes \zeta' \in U_t\},\ W_t = \{\zeta' \in Y^{\mathbb{N}} : \mu_p(V_t^{\zeta'}) > \frac{1}{2^t}\}. \tag{3.37}$$

We claim that, for each $\zeta' \in Y^{\mathbb{N}}$, there is a ζ'-computable functional such that (3.33) holds for $\{V_t^{\zeta'}\}_{t=0}^\infty$.

(b.1) Now,

$$\xi' \in V_t^{\zeta'} \;\Leftrightarrow\; \varphi_{h(t)}^{(1),\xi'\otimes\zeta'}(0)\downarrow \Leftrightarrow (\exists s)\varphi_{h(t)}^{(1),(\xi'\otimes\zeta')[s]}(0)\downarrow .$$

Since the functional $(\xi', s) \mapsto (\xi' \otimes \zeta')[s]$ is ζ'-computable, and the predicate $\varphi_{h(t)}^{(1),\sigma}(0)\downarrow$ is computable in (t, σ), the claim is proved.

We claim that there is a computable functional that satisfies (3.33) for $\{W_t\}$ and $\mu_q(W_t) \leq \frac{1}{2^t}$.

(b.2) $\zeta' \in W_t$ if and only if $\mu_p(V_t^{\zeta'}) > \frac{1}{2^t}$. But if we define

$$V_{t,s}^{\zeta'} = \bigcup\{N_\sigma \subseteq X^{\mathbb{N}} : \sigma \in X^s \wedge \varphi_{h(t)}^{(1),\sigma\otimes(\zeta'[s])}(0)\downarrow\},$$

then $V_{t,s}^{\zeta'} \subseteq V_{t,s+1}^{\zeta'}$ for all $s \in \mathbb{N}$ and $V_t^{\zeta'} = \bigcup_{s=0}^{\infty} V_{t,s}^{\zeta'}$. Thus, $\mu_p(V_t^{\zeta'}) > \frac{1}{2^t}$ if and only if $(\exists s)(\mu_p(V_{t,s}^{\zeta'}) > \frac{1}{2^t})$. It is easy to check that

$$\mu_p(V_{t,s}^{\zeta'}) = \sum_{\sigma}\{\prod_{j=0,\ldots,s-1} p_{\sigma_j} : \varphi_{h(t)}^{(1),\sigma\otimes(\zeta'[s])}(0)\downarrow \wedge \sigma \in X^s\},$$

which is a computable functional of (ζ', s, t), and so $\{W_t\}_{t=0}^{\infty}$ is uniformly Σ_1^0. Now,

$$\begin{aligned}
\mu_{p\otimes q}(U_t) &= \textstyle\int_{(X\times Y)^{\mathbb{N}}} \chi_{U_t}(\xi'\otimes\zeta')d\mu_{p\otimes q}(\xi'\otimes\zeta') \\
&= \textstyle\int_{(X\times Y)^{\mathbb{N}}} \chi_{V_t^{\zeta'}}(\xi')d\mu_{p\otimes q}(\xi'\otimes\zeta') = \int_{Y^{\mathbb{N}}} \mu_p(V_t^{\zeta'})d\mu_q(\zeta') \\
&> \textstyle\int_{Y^{\mathbb{N}}} \frac{1}{2^t}\chi_{W_t}(\zeta')d\mu_q(\zeta') = \frac{1}{2^t}\mu_q(W_t).
\end{aligned}$$

Thus,

$$\mu_q(W_t) < 2^t \mu_{p\otimes q}(U_t) \leq \frac{1}{2^t}.$$

Since ζ is μ_q-random, by Solovay's Theorem (see Downey et al. [2006] Definition 3.2), there is some $L \in \mathbb{N}$ such that $\zeta \notin W_t$ for all $t \geq L$. Thus, by (3.37), for all $t \geq L$, $\mu_p(V_t^\zeta) \leq \frac{1}{2^t}$. By (b.1), $\{V_t^\zeta\}_{t=0}^\infty$ is a μ_p-test relative to $C(\zeta)$. But $\xi \in V_t^\zeta$ for all t, and so ξ is not p-random relative to $C(\zeta)$. □

3.7.2.3 Decomposition

Now we shall consider conditional probability and decomposition. We use the result here only for Theorem 3.3.6. Let ξ be a sequence in $X^{\mathbb{N}}$. For any $A \subset X$, we define $\nu^{\xi,A} \in \{0,1\}^{\mathbb{N}}$ to be the sequence such that $\nu_t^{\xi,A} = 1$ if $\xi_t \in A$ and $\nu_t^{\xi,A} = 0$ otherwise. $\nu^{\xi,A}$ records the occurrences of the event A.

For any $\nu \in \{0,1\}^{\mathbb{N}}$, we define $\theta^\nu : \mathbb{N} \to \mathbb{N}$ as follows:

$$\theta^\nu(0) \text{ is the least } t' \text{ such that } \nu_{t'} = 1; \tag{3.38}$$

$$\theta^\nu(t+1) \text{ is the least } t' \text{ such that } \nu_{t'} = 1 \text{ and } t' > \theta^\nu(t). \tag{3.39}$$

θ^ν is then a partial ν-computable function. We can extend this for strings τ in $\{0,1\}^{\mathbb{N}}$ as well:

$$\theta^\tau(0) \text{ is the least } t' \text{ such that } \tau_{t'} = 1;$$

$$\theta^\tau(t+1) \text{ is the least } t' \text{ such that } \tau_{t'} = 1 \text{ and } t' > \theta^\tau(t).$$

In this case, θ^τ is always a partial function.

Applying the construction to $\nu^{\xi,A}$, $\theta^{\nu^{\xi,A}}$ is then ξ-computable and is total if elements in A appear in ξ infinitely often. Define $\xi_t^A = \xi_{\theta^{\nu^{\xi,A}}(t)}$, $t \in \mathbb{N}$. $\xi^A \in A^{\mathbb{N}}$ if and only if $\theta^{\nu^{\xi,A}}$ is total. The sequence ξ^A records the happenings in ξ given the event A. Intuitively, ξ^A should be a random sequence as well, and it should follow the conditional distribution. If $A = X - \{x\}$ for some $x \in X$, then $\nu^{\xi,A}$ is also denoted as $\nu^{\xi,-x}$ and ξ^A is denoted as ξ^{-x}.

On the other hand, let $\xi,\ \zeta \in X^{\mathbb{N}}$ and let $\nu \in \{0,1\}^{\mathbb{N}}$. We shall define an inverse operator that is intended to model composite random processes. We shall now define the *shuffle* of ξ and ζ using ν, denoted by $\xi \oslash_{\nu} \zeta$, as follows: for all $t \in \mathbb{N}$,

$$(\xi \oslash_{\nu} \zeta)_t = (1-\nu_t)\xi_{\sum_{s=0}^{t-1}(1-\nu_s)} + \nu_t \zeta_{\sum_{s=0}^{t-1}\nu_s}. \tag{3.40}$$

The two sequences ξ and ζ can be thought of as two independent processes, and the shuffle of them using ν is the composite process taking ν as a random device. Intuitively, the shuffle is expected to be random and follow the distribution that is a convex combination of the two processes. Notice that for all $t \in \mathbb{N}$ such that $\theta^{\nu}(t)$ is defined, $(\xi \oslash_{\nu} \zeta)_{\theta^{\nu}(t)} = \zeta_t$. Similarly, for all $t \in \mathbb{N}$ such that $\theta^{\nu}(t)$ is defined, $(\xi \oslash_{\nu} \zeta)_{\theta^{1-\nu}(t)} = \xi_t$.

Likewise, for strings $\sigma,\ \sigma' \in X^{<\mathbb{N}}$ with and $\tau \in \{0,1\}^{\mathbb{N}}$ such that $|\sigma| = |\tau| = |\sigma'| = s$, $\sigma \oslash_{\tau} \sigma' \in X^s$ is defined as the follows: for all $t = 0, .., s-1$,

$$(\sigma \oslash_{\tau} \sigma')_t = (1-\tau_t)\sigma_{\sum_{u=0}^{t-1}(1-\tau_u)} + \tau_t \sigma'_{\sum_{u=0}^{t-1}\tau_u}.$$

For any $x \in X$, let $\mathbf{x}$ denote the sequence in $X^{\mathbb{N}}$ such that $\mathbf{x}_t = x$ for all $t \in \mathbb{N}$. Then, it is easy to check that $\mathbf{x} \oslash_{\nu^{\xi,-x}} \xi^{-x} = \xi$ for any $\xi \in X^{\mathbb{N}}$.

Now we shall show that all these intuitions are true. The way we prove this is to establish some measure theoretical lemmas, and then apply them to construct tests.

Lemma 3.7.1. *Consider any finite set X with $|X| \geq 2$.*

(a) Let $x \in X$. Consider the mapping $T : (X - \{x\})^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \to X^{\mathbb{N}}$ such that $T(\xi, \nu) = \mathbf{x} \oslash_{\nu} \xi$, and consider the probability distribution $p \in \Delta(X)$ such that $p_x < 1$. For any measurable set

$V \subseteq X^{\mathbb{N}}$, we have

$$\mu_p(V) = \mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(V)), \tag{3.41}$$

where $p^{-x} \in \Delta(X - \{x\})$ is defined as $p^{-x}_{x'} = \frac{p_{x'}}{1-p_x}$ for all $x' \neq x$.

(b) Let $p, q \in \Delta(X)$ and let $\alpha \in [0,1] \cap \mathbb{Q}$. Let $T : X^{\mathbb{N}} \times X^{\mathbb{N}} \times \{0,1\}^{\mathbb{N}} \to X^{\mathbb{N}}$ be such that $T(\xi, \zeta, \nu) = \xi \oslash_\nu \zeta$. For any measurable set $V \subseteq X^{\mathbb{N}}$,

$$\mu_{p\otimes q\otimes(\alpha,1-\alpha)}(T^{-1}(V)) = \mu_{\alpha p+(1-\alpha)q}(V). \tag{3.42}$$

Proof. (a) Since μ_p is regular, it suffices to show that (3.41) holds for all open sets V in $X^{\mathbb{N}}$. First we remark two facts. For any $\sigma,\ \sigma' \in X^{<\mathbb{N}}$, $T^{-1}(N_\sigma \cap N'_\sigma) = T^{-1}(N_\sigma) \cap T^{-1}(N'_\sigma)$. For any collection $G \subseteq X^{<\mathbb{N}}$, $\bigcup_{\sigma\in G} T^{-1}(N_\sigma) = T^{-1}(\bigcup_{\sigma\in G} N_\sigma)$. Now, any open set $V \subseteq X^{\mathbb{N}}$ can be written as $V = \bigcup_{\sigma\in G} N_\sigma$ for some prefix-free set $G \subseteq X^{<\mathbb{N}}$, and so if

$$\mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(N_\sigma)) = \mu_p(N_\sigma) \tag{3.43}$$

holds for all $\sigma \in X^{<\mathbb{N}}$, then

$$\mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(V)) = \sum_{\sigma\in G} \mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(N_\sigma)) = \sum_{\sigma\in G} \mu_p(N_\sigma) = \mu_p(V),$$

for any such open set V.

Now we show that (3.43) holds for any $\sigma \in X^{<\mathbb{N}}$. It is easy to check that $T^{-1}(N_\sigma) = N_{\sigma'} \times N_{\tau'}$, where $\tau'_t = 0$ if $\sigma_t = x$, $\tau'_t = 1$ otherwise, and $\sigma'_t = \sigma_{\theta^\tau(t)}$ for $t = 0, ..., |\sigma'| - 1$, where $|\sigma'| = |\{t : 0 \leq t \leq |\sigma| - 1 : \sigma_t \neq x\}|$.

Thus, $\mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(N_\sigma))$

$= \mu_{p^{-x}}(N_{\sigma'}) \times \mu_{(p_x,1-p_x)}(N_{\tau'})$

$= p_x^{|\tau'|-\sum_{t=0}^{|\tau'|-1}\tau'_t}(1-p_x)^{\sum_{t=0}^{|\tau'|-1}\tau'_t}(\frac{1}{1-p_x})^{|\sigma'|}\prod_{t=0}^{|\sigma'|-1} p_{\sigma'_t}$

$= p_x^{|\tau'|-\sum_{t=0}^{|\tau'|-1}\tau'_t}\prod_{t=0}^{|\sigma'|-1} p_{\sigma'_t}$ $(|\sigma'| = \sum_{t=0}^{|\tau'|-1}\tau'_t)$

$= \prod_{t=0}^{|\sigma|-1} p_{\sigma_t} = \mu_p(N_\sigma).$

(b) As in (a), it suffices to show that (3.42) holds for all basic open sets. Let $\sigma \in X^{<\mathbb{N}}$. $(\xi, \zeta, \nu) \in T^{-1}(N_\sigma)$ if and only if there are $\tau \in \{0,1\}^{<\mathbb{N}}$, $\sigma^1,\ \sigma^2 \in X^{<\mathbb{N}}$ such that

(b.1) $|\tau| = |\sigma|$;

(b.2) $|\sigma^1| = |\{s : 0 \le s \le |\tau|-1,\ \tau_s = 0\}|$ and $|\sigma^2| = |\{s : 0 \le s \le |\tau|-1,\ \tau_s = 1\}|$;

(b.3) for all $t = 0, ..., |\sigma^1|-1$, $\sigma_{\theta^{1-\tau}(t)} = \sigma^1_t$, and for all $t = 0, ..., |\sigma^2|-1$, $\sigma_{\theta^{\tau}(t)} = \sigma^2_t$;

(b.4) $\sigma^1 \subset \xi$, $\sigma^2 \subset \zeta$, and $\tau \subset \nu$.

Hence,

$$T^{-1}(N_\sigma) = \bigcup \{N_{\sigma^1} \times N_{\sigma^2} \times N_\tau : \tau,\ \sigma^1,\ \sigma^2 \text{satisfy (b.1-3) above}\}. \tag{3.44}$$

Notice that for each $\tau \in \{0,1\}^{\mathbb{N}}$, there is a unique pair (σ^1, σ^2) that satisfy (b.1-3) above.

$\mu_{p\otimes q\otimes(\alpha,1-\alpha)}(T^{-1}(N_\sigma))$

$= \sum_{|\tau|=|\sigma|}\prod_{s=0}^{|\tau|-1}(\alpha p_{\sigma_s})^{1-\tau_s}((1-\alpha)q_{\sigma_s})^{\tau_s}$

$= \prod_{s=0}^{|\sigma|-1}(\alpha p_{\sigma_s} + (1-\alpha)q_{\sigma_s}) = \mu_{\alpha p+(1-\alpha)q}(N_\sigma).$ $\square$

The following theorem states that Martin-Löf randomness is closed under conditional probability.

Theorem 3.7.4. *Let A be a subset of X and let $p \in \Delta(X)$ be such that $p_A =_{def} \sum_{x \in A} p_x > 0$. Suppose that $\xi \in X^{\mathbb{N}}$ is μ_p-random. Then,*

(a) $\nu^{\xi,A}$ is $\mu_{(1-p_A,p_A)}$-random relative to ξ^A;

(b) ξ^A is μ_{p^A}-random relative to $\nu^{\xi,A}$, where $p^A_x = \frac{p_x}{p_A}$ for all $x \in A$ and $p^A_x = 0$ otherwise.

Proof. We first show that the theorem holds for A of the form $X - \{x\}$.

(a) (for $A = X - \{x\}$) Suppose that $\nu^{\xi,-x} \in \bigcap_{t=0}^{\infty} U_t$, where $\{U_t\}_{t=0}^{\infty}$ is a $\mu_{(p_x,1-p_x)}$-test relative to ξ^{-x} with $\mu_{(p_x,1-p_x)}(U_t) \leq \frac{1}{2^t}$. Let $U_t = \{\nu \in \{0,1\}^{\mathbb{N}} : \varphi^{(1),\nu\oplus\xi^{-x}}_{h(t)}(0)\downarrow\}$, where h is a total computable function. For each $\zeta \in (X - \{x\})^{\mathbb{N}}$, define $U^{\zeta}_{t,s} = \{\nu \in \{0,1\}^{\mathbb{N}} : \varphi^{(1),(\nu\oplus\zeta)[2s]}_{h(t)}(0)\downarrow\}$. Let

$$U^{\zeta}_t[\frac{1}{2^t}] = \bigcup_{\mu_{(p_x,1-p_x)}(U^{\zeta}_{t,s}) \leq \frac{1}{2^t}} U^{\zeta}_{t,s}.$$

For each $\sigma \in (X - \{x\})^{<\mathbb{N}}$, we define

$$U^{\sigma}_t = \{\nu \in \{0,1\}^{\mathbb{N}} : \varphi^{(1),\nu[|\sigma|]\oplus\sigma}_{h(t)}(0)\downarrow\}.$$

By construction, we have that $\nu^{\xi,-x} \in \bigcap_{t=0}^{\infty} U^{\xi^{-x}}_t = \bigcap_{t=0}^{\infty} U^{\xi^{-x}}_t[\frac{1}{2^t}]$.

For each $t \in \mathbb{N}$, we define $V_t \subseteq X^{\mathbb{N}}$ as follows: $\xi' \in V_t$ if and only if there is some $s \in \mathbb{N}$ such that (notice that $(\xi')^{-x}[s]$ is defined if and only if $(\forall j < s)(\theta^{\nu^{\xi',-x}}(j)\downarrow)$)

(a1) $(\forall j < s)(\theta^{\nu^{\xi',-x}}(j)\downarrow)$;

(a2) $\varphi^{(1),\nu^{\xi',-x}[s]\oplus(\xi')^{-x}[s]}_{h(t)}(0)\downarrow$;

(a3) $\mu_{(p_x,1-p_x)}(U^{(\xi')^{-x}[s]}_t) \leq \frac{1}{2^t}$.

We claim that $\{V_t\}_{t=0}^{\infty}$ is is a μ_p-test with $\mu_p(V_t) \leq \frac{1}{2^t}$.

For the first half of the claim, it suffices to check that all predicates in (a1-3) are of Σ^0_1.

(a1) The functional $(\xi', j) \mapsto \theta^{\nu^{\xi', -x}}(j)$ is computable, and so the predicate in (a1) is of Σ^0_1 in $X^{\mathbb{N}} \times \mathbb{N}^2$.

(a2) The functional $(\xi', s) \mapsto (\nu^{\xi', -x}[s] \oplus (\xi')^{-x}[s])$ is computable (it is undefined if $\theta^{\nu^{\xi', -x}}(j)$ is undefined for some $j < s$). Thus, the predicate in (a2) is of Σ^0_1 in $X^{\mathbb{N}} \times \mathbb{N}^2$.

(a3) The functional $(\xi', s, t) \mapsto \mu_{(p_x, 1-p_x)}(U_t^{(\xi')^{-x}[s]})$ is computable since

$$U_t^\sigma = \bigcup\{N_\tau : \tau \in \{0,1\}^{\mathbb{N}}, |\tau| = |\sigma|, \varphi_{h(t)}^{(1),\tau\oplus\sigma}(0) \downarrow\}$$

and so

$$\mu_{(p_x,1-p_x)}(U_t^{(\xi')^{-x}[s]}) = \sum_{\tau\in\{0,1\}^{<\mathbb{N}}} \{ \prod_{j=0,\ldots,s-1} p_x^{1-\tau_j}(1-p_x)^{\tau_j} : \varphi_{h(t)}^{(1),(\tau\oplus(\xi')^{-x}[s])}(0) \downarrow \wedge |\tau| = s\}.$$

Thus, the predicate $\xi' \in V_t$ is Σ^0_1, and so $\{V_t\}_{t=0}^\infty$ is uniformly Σ^0_1.

Let $\Gamma = \{\xi' \in X^{\mathbb{N}} : (\forall t)(\theta^{\nu^{\xi', -x}}(t) \downarrow)\}$. Let $\Gamma_s = \{\xi' \in X^{\mathbb{N}} : \xi'_s \neq x\}$. Then $\mu_p(\Gamma_s) = 1 - p_x$, and $\{\Gamma_s\}_{s=0}^\infty$ is a sequence of independent events. Since $p_x < 1$,

$$\sum_{s=0}^\infty \mu_p(\Gamma_s) = \sum_{s=0}^\infty (1-p_x) = \infty$$

and so by the second Borel-Cantelli lemma,

$$\mu_p(\Gamma) = \mu_p(\bigcap_{t=0}^\infty \bigcup_{s=t}^\infty \Gamma_s) = 1.$$

For each $t \in \mathbb{N}$, let $V_t^1 = \Gamma \cap V_t$ and let $V_t^0 = (X^{\mathbb{N}} - \Gamma) \cap V_t$. So

$$\mu_p(V_t) = \mu_p(V_t^0) + \mu_p(V_t^1) = \mu_p(V_t^1).$$

Moreover, $\xi' \in V_t^1$ if and only if

$$\xi' \in \Gamma \wedge (\exists s)(\varphi_{h(s)}^{(1),\nu^{\xi',-x}[s]\oplus(\xi')^{-x}[s]}(0)\downarrow \wedge \mu_{(p_x,1-p_x)}(U_t^{(\xi')^{-x}[s]}) \leq \frac{1}{2^t})$$

if and only if $\theta^{\nu^{\xi',-x}}$ is total and $\nu^{\xi',-x} \in U_t^{(\xi')^{-x}}[\frac{1}{2^t}]$. Thus, if we define T to be such that $T(\zeta,\nu) = \mathbf{x} \oslash_\nu \zeta$, then, by Lemma 3.7.1 (a),

$\mu_p(V_t^1) = \mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(V_t^1))$

$\leq \int_{(X-\{x\})^{\mathbb{N}}} \int_{\{0,1\}^{\mathbb{N}}} \chi_{U_t^\zeta[\frac{1}{2^t}]}(\nu) d\mu_{(p_x,1-p_x)}(\nu) d\mu_{p^{-x}}(\zeta)$

$\leq \int_{(X-\{x\})^{\mathbb{N}}} \mu_{(p_x,1-p_x)}(U_t^\zeta[\frac{1}{2^t}]) d\mu_{p^{-x}}(\zeta) \leq \frac{1}{2^t}$.

Therefore, $\{V_t\}_{t=0}^\infty$ is a μ_p-test, and we have $\xi = \mathbf{x} \oslash_{\nu^{\xi,-x}} \xi^{-x} \in V_t^1 \subseteq V_t$ for all $t \in \mathbb{N}$. Hence, ξ is not μ_p-random in $X^{\mathbb{N}}$.

(b) (for $A = X - \{x\}$) Suppose that $\xi^{-x} \in \bigcap_{t=0}^\infty U_t^{\nu^{\xi,-x}}$, where $\{U_t^{\nu^{\xi,-x}}\}_{t=0}^\infty$ is a uniformly $\Sigma_1^{0,\nu^{\xi,-x}}$ sequence of sets in $(X-\{x\})^{\mathbb{N}}$ with $\mu_{p^{-x}}(U_t^{\nu^{\xi,-x}}) \leq \frac{1}{2^t}$. Let

$$U_t^{\nu^{\xi,-x}} = \{\zeta \in (X-\{x\})^{\mathbb{N}} : \varphi_{h(t)}^{(1),\zeta\oplus\nu^{\xi,-x}}(0)\downarrow\},$$

where h is a total computable function. For each $\nu \in \{0,1\}^{\mathbb{N}}$, define

$$U_{t,s}^\nu = \{\zeta \in (X-\{x\})^{\mathbb{N}} : \varphi_{h(t)}^{(1),(\zeta\oplus\nu)[2s]}(0)\downarrow\},\ U_t^\nu[\frac{1}{2^t}] = \bigcup_{\mu_{p^{-x}}(U_{t,s}^\nu)\leq\frac{1}{2^t}} U_{t,s}^\nu.$$

For each $\tau \in \{0,1\}^{<\mathbb{N}}$, we define

$$U_t^{\tau} = \{\zeta \in (X - \{x\})^{\mathbb{N}} : \varphi_{h(t)}^{(1),\zeta[|\tau|]\oplus\tau}(0)\downarrow\}.$$

For each $t \in \mathbb{N}$, we define $V_t \subseteq X^{\mathbb{N}}$ as follows: $\xi' \in V_t$ if and only if there is some $s \in \mathbb{N}$ such that (notice that $(\xi')^{-x}[s]$ is defined if and only if $(\forall j < s)(\theta^{\nu^{\xi',-x}}(j)\downarrow)$)

(b1) $(\forall j < s)(\theta^{\nu^{\xi',-x}}(j)\downarrow)$;

(b2) $\varphi_{h(t)}^{(1),(\xi')^{-x}[s]\oplus\nu^{\xi',-x}[s]}(0)\downarrow$;

(b3) $\mu_{p^{-x}}(U_t^{\nu^{\xi',-x}[s]}) \leq \frac{1}{2^t}$.

Using similar arguments as in (a), we can show that $\{V_t\}_{t=0}^{\infty}$ is uniformly Σ_1^0 and $\mu_p(V_t) \leq \frac{1}{2^t}$.

Define the set Γ as in (a). We have seen that $\mu_p(\Gamma) = 1$. For each $t \in \mathbb{N}$, let $V_t^1 = \Gamma \cap V_t$ and let $V_t^0 = (X^{\mathbb{N}} - \Gamma) \cap V_t$. So $\mu_p(V_t) = \mu_p(V_t^0) + \mu_p(V_t^1) = \mu_p(V_t^1)$. Moreover, $\xi' \in V_t^1$ if and only if

$$\xi' \in \Gamma \wedge (\exists s)(\varphi_{h(t)}^{(1),(\xi')^{-x}[s]\oplus\nu^{\xi',-x}[s]}(0)\downarrow \wedge \mu_{p^{-x}}(U_t^{\nu^{\xi',-x}[s]}) \leq \frac{1}{2^t})$$

if and only if $\theta^{\nu^{\xi',-x}}$ is total and $(\xi')^{-x} \in U_t^{\nu^{\xi',-x}}[\frac{1}{2^t}]$. Thus, if we define T to be such that $T(\zeta,\nu) = \mathbf{x} \oslash_{\nu} \zeta$, then, by Lemma 3.7.1 (a),

$\mu_p(V_t^1) = \mu_{p^{-x}\otimes(p_x,1-p_x)}(T^{-1}(V_t^1))$

$\leq \int_{\{0,1\}^{\mathbb{N}}} \int_{(X-\{x\})^{\mathbb{N}}} \chi_{U_t^{\nu}[\frac{1}{2^t}]}(\zeta) d\mu_{p^{-x}}(\zeta) d\mu_{(p_x,1-p_x)}(\nu)$

$\leq \int_{\{0,1\}^{\mathbb{N}}} \mu_{p^{-x}}(U_t^{\nu}[\frac{1}{2^t}]) d\mu_{(p_x,1-p_x)}(\nu) \leq \frac{1}{2^t}$.

Therefore, $\{V_t\}_{t=0}^{\infty}$ is an μ_p-test, and we have $\xi = \mathbf{x} \oslash_{\nu^{\xi,-x}} \xi^{-x} \in V_t$ for all $t \in \mathbb{N}$. Hence, ξ is not μ_p-random.

Now we prove the theorem for general A. Let $Y = A \cup \{y\}$, where $y \notin X$. Define $\Gamma : X^{\mathbb{N}} \to Y^{\mathbb{N}}$ as $\Gamma(\xi')_t = \xi'_t$ if $\xi'_t \in A$ and $\Gamma(\xi')_t = y$ otherwise. Γ is clearly computable. Let $q \in \Delta(Y)$ be such that $q_x = p_x$ for all $x \in A$ and $q_y = 1 - p_A$. We first show that $\Gamma(\xi) = \zeta$ is μ_q-random. Suppose not. Then $\zeta \in \bigcap_{t=0}^{\infty} U_t$ for some μ_q-test $\{U_t\}_{t=0}^{\infty}$. Define $V_t \subset X^{\mathbb{N}}$ as $V_t = \{\xi' \in X^{\mathbb{N}} : \Gamma(\xi') \in U_t\}$. We claim that for any open set $U \subset Y^{\mathbb{N}}$, $\mu_q(U) = \mu_p(\Gamma^{-1}(U))$. To see this, let $U = \bigcup_{\sigma \in G} N_\sigma$ for some prefix-free $G \subset Y^{<\mathbb{N}}$. Then $\Gamma^{-1}(U) = \bigcup_{\sigma \in G} \bigcup_{\tau \in H_\sigma} N_\tau$, where

$$H_\sigma = \{\tau \in X^{<\mathbb{N}} : |\sigma| = |\tau| \text{ and } \sigma_t = \tau_t \text{ if } \sigma_t \in A,\ \tau_t \in A \text{ otherwise}\}.$$

Notice that $H = \bigcap_{\sigma \in G} H_\sigma$ is also prefix-free. For each $\sigma \in G$,

$$\mu_q(N_\sigma) = \prod_{t=0}^{|\sigma|-1} q_{\sigma_t} = \prod_{\sigma_t \in A} p_{\sigma_t} \prod_{\sigma_t \notin A} (1 - p_A) = \sum_{\tau \in H_\sigma} \prod_{t=0}^{|\tau|-1} p_{\tau_t} = \sum_{\tau \in H_\sigma} \mu_p(N_\tau),$$

and thus

$$\mu_q(U) = \sum_{\sigma \in G} \mu_q(N_\sigma) = \sum_{\sigma \in G} \sum_{\tau \in H_\sigma} \mu_p(N_\tau) = \mu_p(\Gamma^{-1}(U)).$$

Since $V_t = \Gamma^{-1}(U_t)$, it follows that $\mu_p(V_t) = \mu_q(U_t) \leq \frac{1}{2^t}$ for all t. Hence, $\{V_t\}_{t=0}^{\infty}$ is a test. $\zeta \in \bigcap_{t=0}^{\infty} U_t$ implies that $\xi \in \bigcap_{t=0}^{\infty} V_t$ and so this is a contradiction to the fact that ξ is μ_p-random.

By construction, $\nu^{\zeta,-y} = \nu^{\xi,A}$ and $\xi^A = \zeta^{-y}$. As we have shown, $\nu^{\zeta,-y}$ is $\mu_{(q_y,1-q_y)} = \mu_{(1-p_A,p_A)}$-random relative to ζ^{-y}, and ζ^{-y} is $q^{-y} = \mu_{p^A}$-random. $\square$

3.7.2.4 Martingales and stochastic sequences

In this section we shall give a proof of Theorem 3.7.6. To this end, we need a new concept called martingales. A martingale is a formalization of a betting strategy. As we shall see, it is easy to show stochasticity using martingales instead of tests, and there is a characterization of Martin-Löf randomness using martingales. First we give the formal definition.

Definition 3.7.2. Let X be a finite set and let $\mathbf{p} = (p^0, p^1, ...)$ be a computable sequence of distributions over X such that $p^t_x > 0$ for all $x \in X$ and for all $t \in \mathbb{N}$. A function $M : X^{<\mathbb{N}} \to \mathbb{R}_+$ is a martingale with respect to $\mu_{\mathbf{p}}$ if for all $\sigma \in X^{<\mathbb{N}}$, $M(\sigma) = \sum_{x \in X} p^{|\sigma|}_x M(\sigma\langle x \rangle)$.

Let $\mathcal{P}$ satisfy (A1) and (A2). A martingale M is $\mathcal{P}$-effective if there is a sequence of martingales $\{M_t\}_{t=0}^{\infty}$ that satisfies the following properties:

(a) $M_t(\sigma) \in \mathbb{Q}_+$ for all $t \in \mathbb{N}$ and for all $\sigma \in X^{<\mathbb{N}}$;

(b) for each $t \in \mathbb{N}$, $M_t \in \mathcal{P}_T$;

(c) $\lim_{t \to \infty} M_t(\sigma) \uparrow M(\sigma)$ for all $\sigma \in X^{<\mathbb{N}}$.

In this case, we say that the sequence $\{M_t\}_{t=0}^{\infty}$ supports M.

The following theorem characterizes randomness in terms of martingales. The proof for the case $\mathcal{P} = \mathcal{T}$ can be found in Downey et al. [2006], and the proof there can be easily relativized to cover the general case.

Theorem 3.7.5. *Let $\mathcal{P}$ satisfy (A1) and (A2) and let $\mathbf{p}$ be a computable sequence such that $p^t_x > 0$ for all $x \in X$ and for all $t \in \mathbb{N}$. A sequence $\xi \in X^{\mathbb{N}}$ is $\mu_{\mathbf{p}}$-random relative to $\mathcal{P}$ if and only if for any $\mathcal{P}$-effective martingale M w.r.t. $\mu_{\mathbf{p}}$,*

$$\limsup_{T \to \infty} M(\xi[t]) < \infty.$$

We shall then discuss selection function and stochastic sequences. Let X be a finite set, and let $r : X^{<\mathbb{N}} \to \{0, 1\}$ be a selection function. Let $L^{\xi}_r(k) = |\{0 < t < k+1 : r(\xi[t-1]) = 1\}|$ to be the number of elements selected by r in $\xi[k]$. We have defined ξ^r before as the sequence obtained from ξ by applying r to it. With these notations, we now show that any $\mu_{\mathbf{p}}$-random sequence with $\lim_{t\to\infty} p^t = p$ is a p-stochastic sequence.

Theorem 3.7.6. *Let $\mathcal{P}$ satisfy (A1) and (A2). Let $u : X \to \mathbb{N}$ be a function. Suppose that ξ is $\mu_{\mathbf{p}}$-random relative to $\mathcal{P}$ with $p^t_x > 0$ for all $t \in \mathbb{N}$ and for all $x \in X$ and $\lim_{t\to\infty} p^t = p$, and suppose that r is a function in $\mathcal{P}_T$. If ξ^r is total, then*

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{u(\xi^r_t)}{T} = \sum_{x\in X} p_x u(x).$$

Consequently, ξ is a p-stochastic sequence relative to $\mathcal{P}$.

Proof. By Theorem 3.7.5, for any $\mathcal{P}$-effective martingale M,

$$\limsup_{T\to\infty} M(\xi[t]) < \infty.$$

It is sufficient to show that for each $x \in X$,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{\chi_x(\xi^r_t)}{T} = p_x,$$

where $\chi_x(y) = 1$ if $x = y$, $\chi_x(y) = 0$ otherwise.

Suppose that there exists some $\varepsilon > 0$ and a sequence $\{T_k\}_{k=0}^{\infty}$ such that for all $k \in \mathbb{N}$,

$$\sum_{t=0}^{T_k-1} \frac{\chi_x(\xi^r_t)}{T_k} \geq p_x + \varepsilon.$$

We shall define a martingale M as follows:

(a) $M(\epsilon) = 1$;

(b) $M(\sigma\langle x\rangle) = (1+\kappa(1-p_x^{|\sigma|}))M(\sigma)$ and $M(\sigma\langle y\rangle) = (1-\kappa p_x^{|\sigma|})M(\sigma)$ for all $y \neq x$ if $r(\sigma) = 1$;

(c) $M(\sigma\langle y\rangle) = M(\sigma)$ for all $y \in X$ if $r(\sigma) = 0$.

To check that M is a martingale, note that if $r(\sigma) = 1$, then

$$\sum_{y\in X} p_y^{|\sigma|} M(\sigma\langle y\rangle) = p_x^{|\sigma|}(1+\kappa(1-p_x^{|\sigma|}))M(\sigma) + \sum_{y\neq x} p_y^{|\sigma|}(1-\kappa p_x^{|\sigma|})M(\sigma)$$

$$= M(\sigma) + \kappa M(\sigma)(p_x^{|\sigma|}(1-p_x^{|\sigma|}) - (1-p_x^{|\sigma|})p_x^{|\sigma|}) = M(\sigma);$$

if $r(\sigma) = 0$, then $\sum_{y\in X} p_y^{|\sigma|} M(\sigma\langle y\rangle) = \sum_{y\in X} p_y^{|\sigma|} M(\sigma) = M(\sigma)$.

M is in $\mathcal{P}_T$ since r is. For $k \geq 1$, define

$$D_k = \{t \leq k-1 : r(\xi[t]) = 1, \xi_{t+1} = x\} \text{ and } E_k = \{t \leq k-1 : r(\xi[t]) = 1, \xi_{t+1} \neq x\}.$$

Then,

$$M(\xi[k]) = \prod_{t\in D_k}(1+\kappa(1-p_x^{t+1}))\prod_{t\in E_k}(1-\kappa p_x^{t+1}).$$

Let $l_k = (L_r^{\xi})^{-1}(T_k)$. Since ξ^r is total, l_k is well defined for all $k \in \mathbb{N}$.

Let $\delta = \min\{p_x, 1-p_x, \frac{\varepsilon}{2}\}$. Since $\lim_{t\to\infty} p^t = p$, let T be so large that $t \geq T$ implies that $|p_x^t - p_x| < \delta$. Let K be the first k such that $T_k > T$. Then, for all $k > K$,

$$M(\xi[l_k]) = \prod_{t\in D_{l_k}}(1+\kappa(1-p_x^{t+1}))\prod_{t\in E_{l_k}}(1-\kappa p_x^{t+1})$$

$$\geq \prod_{t\in D_{l_K}}(1+\kappa(1-p_x^{t+1})\prod_{t\in E_{l_K}}(1-\kappa p_x^{t+1}))(1+\kappa(1-p_x-\delta))^{|D_{l_k}-D_{l_K}|}(1-\kappa p_x-\kappa\delta)^{|E_{l_k}-E_{l_K}|}.$$

Let

$$A = \frac{\prod_{t\in D_{l_K}}(1+\kappa(1-p_x^{t+1})\prod_{t\in E_{l_K}}(1-\kappa p_x^{t+1}))}{(1+\kappa(1-p_x-\delta))^{L^1}(1-\kappa p_x-\kappa\delta)^{L^2}},$$

where

$$L^1 = |D_{l_K}| \text{ and } L^2 = |E_{l_K}|.$$

Since for each k, $|D_{l_k}| \geq T_k p_x + T_k \varepsilon$,

$$M(\xi[l_k]) \geq A((1+\kappa(1-p_x-\delta))^{p_x+\varepsilon}(1-\kappa p_x - \kappa\delta)^{1-p_x-\varepsilon})^{T_k}.$$

Define

$$F(\kappa) = (1+\kappa(1-p_x-\delta))^{p_x+\varepsilon}(1-\kappa p_x-\kappa\delta)^{1-p_x-\varepsilon}.$$

We have $\ln F(0) = 1$ and

$$(\ln F)'(0) = (p_x+\varepsilon)(1-p_x-\delta) - (1-p_x-\varepsilon)(p_x+\delta) = \varepsilon - \delta > 0.$$

Thus, for κ small enough, $F(\kappa) > 1$, and so

$$\limsup_{T\to\infty} M(\xi[T]) = \infty,$$

a contradiction. □

Theorem 3.7.6 is an effective version of the Law of Large Numbers. Here we report an effective version of the Law of the Iterative Logarithm.

Theorem 3.7.7. *Suppose that ξ is a $\mu_{\mathbf{p}}$-random sequence relative to $\mathcal{T}$ with $\mathbf{p} = (p^0, p^1, ..., p^t, ...)$. Then, for any $x \in X$,*

$$\limsup_{T\to\infty} \frac{|\sum_{t=0}^{T-1}(c_x(\xi_t) - p_x^t)|}{\sqrt{2(\sum_{t=0}^{T-1} p_x^t(1-p_x^t)) \log\log \sqrt{(\sum_{t=0}^{T-1} p_x^t(1-p_x^t))}}} = 1, \tag{3.45}$$

Proof. The positive part of equation (3.45) is equivalent to the following two conditions:

(a) for all rational $\varepsilon > 0$,

$$(\exists S)(\forall T \geq S)\sum_{t=0}^{T-1}(c_x(\xi_t) - p_x^t) \leq \sqrt{2(1+\varepsilon)(\sum_{t=0}^{T-1} p_x^t(1-p_x^t)) \log\log \sqrt{(\sum_{t=0}^{T-1} p_x^t(1-p_x^t))}}.$$

(b) for all rational $\varepsilon > 0$,

$$(\forall S)(\exists T \geq S)\sum_{t=0}^{T-1}(c_x(\xi_t) - p_x^t) \geq \sqrt{2(1-\varepsilon)(\sum_{t=0}^{T-1} p_x^t(1-p_x^t)) \log\log \sqrt{(\sum_{t=0}^{T-1} p_x^t(1-p_x^t))}}.$$

Let

$$E_T^\varepsilon = \{\zeta : \sum_{t=0}^{T-1}(c_x(\zeta_t) - p_x^t) > \sqrt{2(1+\varepsilon)(\sum_{t=0}^{T-1} p_x^t(1-p_x^t)) \log\log \sqrt{(\sum_{t=0}^{T-1} p_x^t(1-p_x^t))}}\},$$

and

$$F_T^\varepsilon = \{\zeta : \sum_{t=0}^{T-1}(c_x(\zeta_t) - p_x^t) < \sqrt{2(1-\varepsilon)(\sum_{t=0}^{T-1} p_x^t(1-p_x^t)) \log\log \sqrt{(\sum_{t=0}^{T-1} p_x^t(1-p_x^t))}}\}.$$

Clearly, condition (a) is equivalent to $\xi \notin \bigcap_{S=0}^{\infty}\bigcup_{T=S}^{\infty} E_T^\varepsilon$ and condition (b) is equivalent to $\xi \notin \bigcup_{S=0}^{\infty}\bigcap_{T=S}^{\infty} F_T^\varepsilon$. By Theorem 7.5.1 in Chung [1968],

$$\mu_{\mathbf{p}}(\bigcap_{S=0}^{\infty}\bigcup_{T=S}^{\infty} E_T^\varepsilon) = 0$$

and

$$\mu_{\mathbf{p}}(\bigcup_{S=0}^{\infty}\bigcap_{T=S}^{\infty}F_T^{\varepsilon})=0.$$

It then follows that $\mu_{\mathbf{p}}(\bigcap_{T=S}^{\infty}F_T^{\varepsilon})=0$ for any $S\in\mathbb{N}$. Because F_T^{ε} is computable (uniformly in T), $\{F_T^{\varepsilon}\}_{T=S}^{\infty}$ is a $\mu_{\mathbf{p}}$-test for any S (notice that $\mu_{\mathbf{p}}(F_T^{\varepsilon})$ is also computable). Therefore, $\xi\notin\bigcup_{S=0}^{\infty}\bigcap_{T=S}^{\infty}F_T^{\varepsilon}$.

On the other hand, the set E_T^{ε} is computable (uniformly in T) and so the set $\bigcup_{T=S}^{\infty}E_T^{\varepsilon}$ is of Σ_1^0 (uniformly in S). For $\{\bigcup_{T=S}^{\infty}E_T^{\varepsilon}\}_{S=0}^{\infty}$ to be a test, we need to show that $\mu_{\mathbf{p}}(\bigcup_{T=S}^{\infty}E_T^{\varepsilon})$ has a computable upper bound for all S. From the proof in Theorem 7.5.1 in Chung [1968], we know that there exists a constant $A>0$ and a number $\overline{k}>0$ such that for all $k\geq\overline{k}$ (with the provision that $c^2(1+\frac{\varepsilon}{2})<1+\varepsilon$), c.f. p. 216),

$$\mu_{\mathbf{p}}(\bigcup_{T=T_k}^{T_{k+1}-1}E_T^{\varepsilon})<\frac{A}{(k\log c)^{1+\frac{\varepsilon}{2}}},$$

where $T_k=\max\{T:\sqrt{\sum_{t=0}^{T}p_y^t(1-p_y^t)}\leq c^k\}$ and $c=1+\frac{\varepsilon}{10}$ (for ε small enough, $c^2(1+\frac{\varepsilon}{2})<1+\varepsilon$).

Let's define $G_0=\bigcup_{T=0}^{T_1-1}E_T^{\varepsilon}$ and $G_k=\bigcup_{T=T_k}^{T_{k+1}-1}E_T^{\varepsilon}$ for $k>0$. Clearly,

$$\bigcap_{S=0}^{\infty}\bigcup_{k=S}^{\infty}G_k=\bigcap_{S=0}^{\infty}\bigcup_{T=S}^{\infty}E_T^{\varepsilon}.$$

Now, because T_k is a computable function of k, $\{\bigcup_{k=S}^{\infty}G_k\}_{S=0}^{\infty}$ is also a sequence of uniformly Σ_1^0 sets. We now show that there is a computable mapping $i\mapsto S_i$ so that $\mu_{\mathbf{p}}(\bigcup_{k=S_i}^{\infty}G_k)\leq\frac{1}{2^i}$. We know that

$$\sum_{k=S}^{\infty}\frac{A}{(k\log c)^{1+\frac{\varepsilon}{2}}}\leq\int_{x=S-1}^{\infty}\frac{A}{(x\log c)^{1+\frac{\varepsilon}{2}}}=(S-1)^{-\frac{\varepsilon}{2}}(\log c)^{-1-\frac{\varepsilon}{2}}.$$

Let $B\in\mathbb{N}$ be such that $B>(A(\log c)^{-1-\frac{\varepsilon}{2}})^{\frac{2}{\varepsilon}}$ and let $N\in\mathbb{N}$ be such that $N>\frac{2}{\varepsilon}$. Take $S_i=B2^{Ni}+1$, and it follows that $\mu_{\mathbf{p}}(\bigcup_{k=S_i}^{\infty}G_k)\leq\frac{1}{2^i}$.

This shows that $\{\bigcup_{k=S}^{\infty} G_k\}_{S=0}^{\infty}$ is a $\mu_{\mathbf{p}}$-test, and so

$$\xi \notin \bigcap_{S=0}^{\infty} \bigcup_{k=S}^{\infty} G_k = \bigcap_{S=0}^{\infty} \bigcup_{T=S}^{\infty} E_T^{\varepsilon}.$$

$\square$

Bibliography

A. Alchain. Why money? *Journal of Money, Credit and Banking*, 9(1):133–40, 1977.

F. J. Anscombe and R. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.

R. J. Aumman and A. Brandenburger. Epistemic conditions for nash equilibrium. *Econometrica*, 63:1161–1180, 1995.

A. Banerjee and E. Maskin. A walrasian theory of barter and exchange. *Quarterly Journal of Economics*, 111(4):955–1005, 1996.

S. Barbera, P. J. Hammond, and C. Seidl. *Handbook of utility theory: Principles*, volume 1. Kluwer Academic Publishers, Dordrecht, the Netherlands, 1998.

P. Battigalli and M. Siniscalchi. Strong belief and forward induction reasoning. *Journal of Economic Theory*, 106:356 – 391, 2002.

E. Ben-Porath. Repeated games with finite automata. *Journal of Economic Theory*, 59:17–32, 1993.

P. Billingsley. *Probability and Measure*. John and Wiley and Sons, Inc., 1995.

J. Brown and R. Rosenthal. Testing the minimax hypothesis: A reexamination of oneills experiment. *Econometrica*, 58:1065–1081, 1990.

C. S. Calude. *Information and randomness: an algorithmic perspective*. Springer, 2nd edition, 2002.

K. Chatterjee and H. Sabourian. Complexity and game theory. 2008.

I. Cho and D. Kreps. Signaling games and stable equilibria. *Quarterly Journal of Economics*, 102 (2):179–222, 1987.

C.-L. Chung. *A course in probability theory*. Harcourt, Brace and World, Inc., 1968.

R. Downey, D. Hirschfeldt, A. Nies, and S. Terwijn. Calibrating randomness. *Bulletin of Symbolic Logic*, (71):1443–1489, 2006.

P. C. Fishburn. *Utility theory for decision making.* Wiley, New York, 1970.

S. Freeman. Transactions costs and the optimal quantity of money. *Journal of Political Economy*, 93(1):146–157, 1985.

D. Gillies. *Philosophical theories of probability.* Routlegde, London and New York, 2000.

E. Green and W. Weber. Will the new $100 decrease counterfeiting? *Federal Reserve Bank of Minneapolis Quarterly Review*, 20(3):3–10, 1996.

J. C. Harsanyi. Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2:1–23, 1973.

J. Hicks. A suggestion for simplifying the theory of money. *Economica New Series*, 2:1–19, 1935.

W. Jevons. *Money and the Mechanism of Exchange.* Appleton, London, 1875.

M. Kaneko and J. J. Kline. Inductive game theory: a basic scenario. *Journal of Mathematical Economics*, 44:1332–1363, 2008.

M. Kaneko and J. J. Kline. Transpersonal understanding through social roles, and emergence of cooperation. *Tsukuba University, Department of Social Systems and Management Discussion Paper Series, No.1228*, 2009.

A. N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission (Problemy Peredachi Informatsii)*, 1:1–7, 1965.

B. Lester, A. Postlewaite, and R. Wright. Information and liquidity. 2009.

Y.-T. Li and G. Rocheteau. On the threat of counterfeiting.

R. Duncan Luce and Howard Raiffa. *Games and decisions: Introduction and critical survey.* Dover, 1957.

G. Mailath, M. Okuno-Fujiwara, and A. Postlewaite. Belief-based refinements in signaling games. *Journal of Economic Theory*, 60:241–276, 1993.

P. Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

E. Nosal and N. Wallace. A model of (the threat of) counterfeiting. *Journal of Monetary Economics*, 54(34):994–1001, 2007.

P. G. Odifreddi. *Classical recursion theory*, volume I. Amsterdam-North Holland, 1989.

B. O'Neill. Comments on brown and rosenthal's reexamination. *Econometrica*, 59(2):503–507, 1991.

M. J. Osborne and A. Rubinstein. *A course in game theory.* MIT Press, 1994.

I. Palacios-Huerta. Professionals play minimax. *Review of Economic Studies*, 70(2):395–415, 2003.

N. Pippenger. *Theories of computability.* Cambridge University Press, Cambrigde, United Kingdom, 1997.

G. Ramey. D1 signaling equilibria with multiple signals and a continuum of types. *Journal of Economic Theory*, 69:508–531, 1996.

M. van Lambalgen. *Random sequences.* Ph. D. Diss. University of Amsterdam., 1987.

M. van Lambalgen. The axiomatization of randomness. *Journal of Symbolic Logic*, 55(3):1143–1167, 1990.

F. Velde, W. Weber, and R. Wright. A model of commodity money, with applications to gresham's law and the debasement puzzle. *Review of Economic Dynamics*, 2:291–323, 1999.

J. Ville. *Ĕtude critique de la concept du collectif.* Gauthier-Villars, 1939.

R. von Mises. *Probability, statistics, and truth.* Dover, 2nd edition, 1981.

J. von Neumann and O. Morgenstern. *Theory of games and economic behavior.* Princeton University, 1944.

A. Wald. Die widerspruchsfreiheit des kollektivsbegriffes. *American Economic Review*, 1938.

M. Walker and J. Wooders. Minimax play at wimbledon. *American Economic Review*, 91:1521–1538, 2001.

N. Wallace. A suggestion for oversimplifying the theory of money. *Economic Journal*, 98(390): 25–36, 1988.

R. Weatherford. *Philosophical foundations of probability theory.* Routledge and Kegan Paul, 1982.

S. Williamson. Private money and counterfeiting. *Federal Reserve Bank of Richmond Economic Quarterly*, 88:37–57, 2002.

S. Williamson and R. Wright. Barter and monetary exchange under private information. *American Economic Review*, 84:104–123, 1994.

T. Zhu. Existence of a monetary steady state in a matching model: divisible money. *Journal of Economic Theory*, 123:135–160, 2005.

T. Zhu and N. Wallace. Pairwise trade and coexistence of money and higher-return assets. *Journal of Economic Theory*, 133:524–535, 2007.

A. K. Zvonkin and L. A. Levin. The complexity of finite objects and the basing of the concepts of information and randomness on the theory of algorithms. *Uspehi Mat. Nauk.*, 25(6):85–127, 1970.

Vita

Tai-Wei Hu

Education

Ph.D. Economics, Penn State University, August 2009.

M.A. Economics, National Taiwan University, June 2005.

B.A. Accounting, National Taiwan University, June 2002.

Research Interests

Decision theory, Game theory, Monetary theory.

Publications

1. "On p-rationalizability and approximate common certainty of rationality," *Journal of Economic Theory.* Volume 136, Issue 1, September 2007, Pages 379-391.

2. "Coalition-proof trade and the Friedman rule in the Lagos-Wright model." with John Kennan and Neil Wallace. *Journal of Political Economy.* Volume 117, Issue 1, Pages 116-137.

3. "Expected utility theory from the frequentist perspective," forthcoming, *Economic Theory.*

Completed Manuscripts

1. "Does imperfect recognizability explain coexistence of money and higher-return assets?"

Link: www.taiweihu.weebly.com/research.html

2. "Complexity and mixed strategy equilibrium."

Link: www.taiweihu.weebly.com/research.html

Work in Progress

"Logic for games of perfect information and epistemic conditions for backward induction and for subgame perfectness."

Award

Award for Research Excellence, 2007, College of the Liberal Arts, Penn State.