I. Foundations: The Theory of Rational Beliefs

2. On the structure and diversity of rational beliefs (Edited)*

Mordecai Kurz

Department of Economics, Serra Street at Galvez, Stanford University, Stanford, CA 94305-6072, USA

Summary. The paper proposes that the theory of expectations be reformulated under the assumption that agents do not know the structural relations (such as equilibrium prices) of the economy. Instead, we postulate that they can observe past data of the economy and form probability beliefs based on the data generated by the economy. Using past data agents can compute relative frequencies and the basic assumption of the theory is that the system which generates the data is stable in the sense that the empirically computed relative frequencies converge. It is then shown that the limit of these relative frequencies induce a probability on the space of infinite sequences of the observables in the economy. This probability is stationary. A *belief* of an agent is a probability on the space of infinite sequences of the observable variables in the economy. Such a probability represents the "theory" or "hypothesis" of the agent about the mechanism which generates the data. A belief is said to be *compatible with the data* if under the proposed probability belief the economy would generate the same limit of the relative frequencies as computed from the real data. A theory which is "compatible with the data" is a theory which cannot be rejected by the data. A belief is said to be a *Rational Belief* if it is (i) compatible with the data and (ii) satisfies a certain technical condition. The Main Theorem provides a characterization of all Rational Beliefs.

JEL Classification Numbers: D81, D84, E37.

1 Introduction

The formation of expectations and probability beliefs has played a central role in the formulation of dynamic equilibria and rational expectations has been at the foundations of most expectations models in recent years. Yet, the theory of rational expectations in economics and game theory is based on the premise that agents know a great deal about the basic structure of their environment.

In economics agents are assumed to have knowledge about demand and supply functions, of how to extract present and future general equilibrium prices, and about the stochastic law of motion of the economy over time. In game situations it is assumed that players know the structure of the game and the equilibrium (random) strategies of the opponents. For the sake of terminological clarity we shall say that these agents possess "*structural knowledge.*" We suggest this term in order to distinguish it from the often used term of "information" which is employed to indicate the degree of observability of the state.

It is usually hard to conceive of how agents come to possess structural knowledge. The recent response to this problem has been to formulate dynamic processes of learning which aim to show how agents learn what they know when formulating their beliefs. The problem is that this research has not solved the initial problem. Without engaging in a full scale survey of the results of the recent effort, we think it is accurate to say that there are examples worked out where complete learning does take place. However, in general, the learning approach has not been able to provide a satisfactory mechanism for agents to acquire full structural knowledge and hold rational expectations. This conclusion has a counterpart in the statistical literature where a spirited debate has been taking place about the "Bayes consistency" problem (see Diaconis and Freedman [1986] for an excellent recent survey). We note that "Bayes consistency" may fail even when the statistician is able to conduct independent, repeated controlled experiments[1]. The problem is then compounded by the fact that in almost all instances, a learning economic agent cannot obtain independent observations and must be content with the actual data generated by the system.

A central characteristic of rational expectations equilibria is the fact that in such equilibria all agents hold the same probability belief and make the same forecasts i.e. those implied by the stochastic law of motion of the economy in the given equilibrium. This is also true of Bayesian equilibria. Although not mandated by the axioms of subjective probability, applications of the Bayesian approach in the social sciences almost always make the "common prior" assumption which requires all agents to have the same prior if they have the same information. In fact, Bayesians insist that this assumption is justified both on general principles, as well as being a prerequisite for the consistency of any equilibrium concept (see, for example, Aumann [1976] and [1987] page 12 and Harsanyi [1967–1968]).

We suggest that it is an empirical fact that intelligent economic agents may exhibit drastic differences in beliefs even when they have the same information. Consequently, the study of the diversity of rational beliefs is important and the characterization of the conditions which permit this diversity to arise is one of the objectives of this paper.

---

[1] Feldman [1991] uses the results surveyed by Diaconis and Freedman [1986] to demonstrate the problematics of Bayes learning in economics even in the i.i.d case.

The prototype problem with which this paper is concerned may be simply explained with the aid of an example. Let y_t, $t = 0, 1, 2, \ldots$ be a sequence of random profits or rewards of a household, a corporation or an investment project. Let $0 < \gamma < 1$ be the discount rate employed and let the present value, at date t, of future rewards be defined by

$$p_t^* = \sum_{k=0}^{\infty} \gamma^{k+1} y_{t+k}. \tag{1}$$

An economic agent who observes the data needs to evaluate, at date t, the risky prospect p_t^*. The problem is that the agent *does not know the true probability of the random sequence* $\{y_t, t = 0, 1, 2, \ldots\}$. He does have a finite but massive amount of past data since $t = 0$ occurred a long time ago and all past data were recorded. Given these data the agent sets up to learn all that he can and then form a conditional probability belief Q^t about the future sequence of random variables y_{t+k}, $k > 0$. We aim to establish criteria to determine if a probability belief of an agent is "rational." Moreover, given such criteria of rationality, we want to characterize the structure of all rational beliefs and consequently have a better understanding of the causes for diversity among agents. We formulate the economic environment as a stochastic dynamical system and propose that a belief should be taken to be rational if it is compatible with the data. The idea that rationality of beliefs should be defined relative to what is learnable from the data, rather than relative to some model of the economy is the central driving force of our theory.

2 Model formulation and the issue of stationarity

We suppose that there is a finite number, K, of observables in the economy so that at each date t for $t = 0, 1, 2, \ldots$ all agents observe $x_t = (x_{1t}, x_{2t}, \ldots, x_{Kt}) \in X$ where $X \subseteq \mathbb{R}^K$ is the state space. The x_{it} are such quantities as GNP, prices of commodities or assets, profits of firms or data about climate conditions. The economic environment is represented by a dynamical system $(\Omega, \mathscr{F}, \Pi, T)$ defined on the non-negative integers $t \geq 0$ where

$\Omega = X^{\infty} \subseteq (R^K)^{\infty}$,

$\mathscr{F}^t = \sigma(x_0, x_1, \ldots, x_t), 0 \leq t \leq \infty$, – the σ-field generated by $(x_0, x_1, \ldots, x_t)$,

$\mathscr{F} = \sigma \left(\bigcup_{t=0}^{\infty} \mathscr{F}^t \right)$.

Π is a probability on measurable sets of infinite sequences in X^{∞}. Although we shall think of x as a random point in X^{∞} it is important for us to associate with such a point the starting date of the sequence. We use the notation $x^t = (x_t, x_{t+1}, \ldots)$ to identify a random sequence *from the perspective of date t*. The realization of the stochastic process is represented by the measurable transformation T. We assume, as is standard in probability theory, that T is

a shift transformation. Thus,

$$x^{t+1} = Tx^t. \tag{2}$$

One defines $T^2x = T(Tx)$ and, in general, $T^nx = T(T^{n-1}x)$. From the measurability of T it follows that the iterated maps T^n are also measurable transformations. Since we assume that the process starts at a date called $t = 0$ with $x^0 = x$ we have that

$$x^t = T^t x \qquad t = 0, 1, 2, \ldots$$

T is *not assumed to be invertible*. The economic meaning of this assumption is that any particular *future* evolution, x^t, of the economy is not associated with a *unique past* $T^{-1}(x^t)$; a future x^t may arise from many possible pasts! Since T is not assumed invertible we reserve the notation $T^{-n}S$ for the preimage of S under T^n. That is

$$T^{-n}S = \{x : T^n x \in S\}. \tag{3}$$

For this reason we think of $T^{-n}S$ as the set $S \subset X^\infty$ located n periods into the future. Or, if B is the set of $x \in \Omega$ such that $T^n x \in S$, then B is the set of points in X^∞ from which one reaches S in n steps.

Each of our agents does not have any knowledge of the causal structure of the economy. This leads to the assumption, that Π is not known by anyone. Since the objective of the observer is to discover, for each measurable set $S \in \mathscr{F}$, the true probability $\Pi(S)$ he has a natural way to proceed. Define

$$1_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S \end{cases}$$

and then compute

$$m^n(S)(x) = \frac{1}{n} \sum_{k=0}^{n-1} 1_S(T^k x). \tag{4}$$

$m^n(S)(x)$ is the relative frequency at which the dynamical system visits the set S given that it started at x. Our observing agent can conceivably learn something about the true Π only if $m^n(S)(x)$ converges so that with sufficient data the limit $\lim_{n \to \infty} m^n(S)(x)$ can be computed to any desired accuracy. This motivates the following:

Definition 1: A dynamical system $(\Omega, \mathscr{F}, \Pi, T)$ is said to be *stable* if for all finite-dimensional sets, or cylinders, $S \in \mathscr{F}$ the limit of $m^n(S)(x)$ exists Π a.e. and the limit is denoted by

$$\mathring{m}(S)(x) = \lim_{n \to \infty} m^n(S)(x) \qquad \Pi \text{ a.e.} \tag{5}$$

The system is said to be *strongly stable* if the limit of $m^n(S)(x)$ exists Π a.e. for all $S \in \mathscr{F}$.

The distinction between stability and *strong* stability is important. If S is a cylinder, one needs a *finite* number of observations in order to verify if $x \in S$ or not. However, for infinite dimensional sets a verification that $x \in S$ requires an *infinite* set of observations. We assume that $(\Omega, \mathscr{F}, \Pi, T)$ is only stable but our analysis adapts certain mathematical techniques which were developed for the study of strongly stable systems. Also, since $0 \le m^n(S)(x) \le 1$ the lack of convergence of $m^n(S)(x)$ means that for increasing lengths of time the means $m^n(S)(x)$ remain in different parts of the interval $[0, 1]$ without ever settling down. If this occurs there would be no *common* learning among the agents and little in the way of agreement. The requirement that $m^n(S)(x)$ converges a.e. for all finite cylinders S is a minimal condition needed to establish *common* learning of something meaningful about probabilities[2]. In the development below we shall assume that the limits in (5) are known to all agents.

An important case where a dynamical system has adequate repetition is the case of a stationary system. The dynamical system $(\Omega, \mathscr{F}, \Pi, T)$ is said to be *stationary* if the transformation T is measure preserving; that is, if for all $S \in \mathscr{F}$

$$\Pi(T^{-1}S) = \Pi(S).$$

When T preserves Π then Π is said to be *invariant* under T.

Almost all results in Ergodic Theory have been proved for the case of measure preserving transformations. When a dynamical system is stationary *and agents know that it is stationary* the questions raised in this paper have very clear answers. The main tool employed is Birkhoff's ergodic theorem (1931). To see the implications of this theorem to our problem introduce the following terms:

Definition 2: $S \in \mathscr{F}$ is said to be invariant with respect to T if $T^{-1}S = S$. A measurable function is said to be invariant with respect to T if for any $x \in \Omega$, $f(Tx) = f(x)$.

Definition 3: A dynamical system is said to be ergodic if $\Pi(S) = 0$ or $\Pi(S) = 1$ for all invariant sets S.

Now let the collection $\mathfrak{J}$ of invariant sets be defined by

$$\mathfrak{J} = \{S \in \mathscr{F} : T^{-1}S = S\}.$$

It is easily seen that $\mathfrak{J}$ is a sub σ-field of $\mathscr{F}$ and hence one can define the conditional probability of Π given $\mathfrak{J}$; we denote it by

$$\Pi(S | \mathfrak{J})(\omega) \quad \text{for all} \quad S \in \mathscr{F}, \quad \omega \in \Omega.$$

---

[2] The data may need to be "cleaned" for trend and deterministic cycles if "stability" is to be a useful tool. This is a standard practice in time series analysis. We note that deterministic cycles of amplitude up to M are identified by applying the definition of stability to the transformations T^n for $n = 2, 3, \ldots, M$ and determining the limits of such relative frequencies. For example T^2 calls for an examination of the data from the perspective of all even (or odd) dates: T^3 calls for shifting 3 dates at a time. More details on this issue are provided in Section 5.4 below.

Applying Birkhoff's ergodic theorem to our problem when $(\Omega, \mathscr{F}, \Pi, T)$ is stationary hence strongly stable, we do not assume the condition of strong stability but rather, *we prove it*. We then draw three implications:

(a) $\lim_{n \to \infty} m^n(S)(x) = m(S)(x)$ exists Π a.e. for all $S \in \mathscr{F}$,

(b) $m(S)(x) = \Pi(S|\mathfrak{J})(x)$ Π a.e. for all $S \in \mathscr{F}$,

(c) if $(\Omega, \mathscr{F}, \Pi, T)$ is ergodic then
$$m(S)(x) = m(S) = \Pi(S) \quad \Pi \text{ a.e.} \quad \text{for all} \quad S \in \mathscr{F}.$$

If the dynamical system is stationary *and the agents know that it is stationary* then they can calculate $m(\cdot)(x)$ and know that they have learned exactly the conditional probability $\Pi(\cdot|\mathfrak{J})(x)$. In the ergodic case the agents calculate the measure m and know that $m = \Pi$. The conclusion that in the non-ergodic case $m(S)(x) = \Pi(S|\mathfrak{J})(x)$ for all $S \in \mathscr{F}$, is sensible since in this case the sequence $(T^n x)$ will visit only the invariant sets which contain x and hence $\Pi(\cdot|\mathfrak{J})(x)$ is the only object which can be learned.

It is important to point out that when the dynamical system is stationary agents may not know that it is stationary. Moreover, there does not exist any statistical means by which agents can ascertain that a stationary system is, in fact, stationary. More important is the fact that the dynamical system may not be stationary. In this eventuality, even if we work with a stable system for which $m(\cdot)(x)$ exists, agents cannot use the ergodic theorem to determine what is it that they are learning.

In our view the determination if a dynamical system is stationary or not must originate with the foundation of the mechanism which gives rise to the system[3]. Thus stationarity is a logical implication of the underlying *theory* rather than an *empirical observation* which is deduced from the data. Apart from the fact that agents do not know the stochastic mechanism which generates the data it is relatively rare that economic theoretic reasoning enables us to make a *logical deduction* of what must be the nature of the probability laws under which the economic data is generated. More specifically, we suggest that although the assumption of stationarity is almost universally employed in applied economics, there is little theoretical justification for it. If anything, there are compelling reasons to question it. Economic growth has been associated with bursts of innovations, changes in technology and organizational structure; wars, depressions, major migrations, revolutions, etc. remain outliers in most empirical studies no matter what stationary model is employed.

---

[3] An example will illustrate the point. In certain applications in physics the description of stochastic dynamical systems arises from Hamiltonian structures. These Hamiltonians imply that the transformation of the dynamical systems is measure preserving and thus stochastically stationary. However, this stationarity can be traced to the fact that Hamiltonians are required to satisfy Liouville's theorem on the conservation of energy. Putting it differently, the stationarity of the dynamical system is proved as *a logical consequence of Liouville's Theorem* which, in turn, is proved from the underlying physical structure.

In addition to the above we have our own unique reason to allow for the possibility of non-stationarity. In Kurz [1994b] we incorporate our theory of Rational Beliefs into an equilibrium model. We then show that if agents believe that the economic system *may be* non-stationary, then equilibrium prices and quantities *become non-stationary even if the exogenous environment is, in fact, stationary*. The belief of agents in the possibility of non-stationarity becomes self justifying!

The validity of our approach does not depend on the existence of a conclusive proof for non-stationarity. It does, however, hinge on the fact that we do not have a conclusive theoretical reasoning to *compel* a rational agent to believe that his environment is stationary. We therefore only require that an economic agent not be declared irrational if he takes the view that the economic process at hand *may be* non-stationary.

The concept of "stability" is central to this paper. Since we insist that $(\Omega, \mathscr{F}, \Pi, T)$ may not be stationary[4] our method of analysis is reversed: We *assume* that $(\Omega, \mathscr{F}, \Pi, T)$ is stable and then regard the ergodic properties which emerge as the common learned knowledge of all agents. This calls for further clarification of the property of stability.

3 Stable systems

Definition 1 leaves open a subtle but important technical question. To see the problem denote by $\hat{\mathscr{F}}$ the field of all cylinders in $\mathscr{F}$; it satisfies $\hat{\mathscr{F}} = \bigcup_{t=0}^{\infty} \mathscr{F}^t$. The question is, then, which $x \in \Omega$ satisfy the requirement that $\mathring{m}(S)(x)$ is well defined *for all* $S \in \hat{\mathscr{F}}$. It would be desirable to have this property satisfied Π a.e. so as to enable an extension of $\mathring{m}(\cdot)(x)$ from the field $\hat{\mathscr{F}}$ to the σ-field $\mathscr{F}$ generated by $\hat{\mathscr{F}}$. In other words, we would like this property to hold for all $x \in C$ for some set C with $\Pi(C) = 1$. To accomplish this we first restrict the domain of $\mathring{m}(\cdot)(x)$ slightly. Thus, consider only those sets in $\mathscr{F}^t$ which are defined by rectangles with *rational* end points. That is, each rectangle with rational end points A in X^{t+1} (the $t+1$ product of X) defines a set $S \in \mathscr{F}$ as follows:

$$S = \left\{ x \in \Omega \,\middle|\, \begin{array}{l} (x_0, x_1, \ldots, x_t) \in A \text{ where } A \subset X^{t+1} \text{ is a rectangle} \\ \text{with rational end points and } x_j \in \mathbb{R}^K \text{ for all } j > t \end{array} \right\}.$$

The countable collection of such sets in $\mathscr{F}^t$ is denoted by $\mathring{\mathscr{F}}^t$ and we know that $\mathscr{F}^t = \sigma(\mathring{\mathscr{F}}^t)$. The sequence $\{\mathring{\mathscr{F}}^t, t = 0, 1, 2, \ldots\}$ is a generating sequence for $\mathscr{F}$ in the sense that

$$\mathscr{F} = \sigma\left(\bigcup_{t=0}^{\infty} \mathring{\mathscr{F}}^t \right).$$

---

[4] In the term "stationarity" we include all deterministic transformations or decompositions which result in stationarity e.g. stationary increments, stationary ratios, deterministic cycles or, say, two stationary sequences alternating on odd or even dates (see footnote 1).

The field $\mathscr{F} = (\bigcup_{t=0}^{t} \mathscr{F}^t)$ contains a countable number of sets which we denote by $\hat{\mathscr{F}} = \{F_i, i = 1, 2, \ldots\}$ and we call $\hat{\mathscr{F}}$ *the field of cylinders with rational end points.* We can finally restrict $\mathring{m}(\cdot)(x)$ to $\hat{\mathscr{F}}$. Although $\mathring{m}(\cdot)(x)$ is defined on $\hat{\mathscr{F}}$, to avoid confusion we shall not introduce a separate notation for the restriction of this measure to $\mathring{\mathscr{F}} \subset \hat{\mathscr{F}}$. Instead, whenever this restriction is discussed we shall simply specify "$\mathring{m}(\cdot)(x)$ on $\hat{\mathscr{F}}$". Finally note that the countable number of members F_i, $i = 1, 2, \ldots$ of the field $\mathscr{F}$ have the property that $\mathring{m}(F_i)(x)$ are well defined Π a.e. for each $F_i \in \hat{\mathscr{F}}$. With this formalized we have:

Proposition 1: There exists a set $C \in \mathscr{F}$ with $\Pi(C) = 1$ and for each $x \in C$ a probability measure $m(\cdot)(x)$ on $(\Omega, \mathscr{F})$ which is the unique extension of $\mathring{m}(\cdot)(x)$. The dynamical systems $(\Omega, \mathscr{F}, m(\cdot)(x), T)$ are stationary for all $x \in C$.

Proof: From the construction of $\hat{\mathscr{F}}$ there exist sets $C_i \in \mathscr{F}$ such that $\Pi(C_i) = 1$ and $\mathring{m}(F_i)(x)$ is well defined for all $x \in C_i$ and $F_i \in \hat{\mathscr{F}}$. Define

$$C = \bigcap_{i=1}^{\infty} C_i$$

and hence $\Pi(C) = 1$. $\mathring{m}(F_i)(x)$ is well defined for all $F_i \in \hat{\mathscr{F}}$ and $x \in C$. Since $\mathring{m}(\cdot)(x)$ is a finitely additive measure on $\hat{\mathscr{F}}$ for each $x \in C$ and the space $(\Omega, \mathscr{F})$ has the countable extension property, there exists a unique extension $m(\cdot)(x)$ which is a probability measure on the σ-field $\mathscr{F} = \sigma(\hat{\mathscr{F}})$ for each $x \in C$. From the definition of stability it follows that $\mathring{m}(T^{-1}S)(x) = \mathring{m}(S)(x)$ for all cylinders $S \in \hat{\mathscr{F}}$ and this property is inherited by $m(\cdot)(x)$. Hence $(\Omega, \mathscr{F}, m(\cdot)(x), T)$ is stationary.

The uniqueness of $m(\cdot)(x)$ means that although the observing agents can compute only $\mathring{m}(S)(x)$ for $S \in \hat{\mathscr{F}}$, they can analytically deduce the knowledge of $m(\cdot)(x)$ on $(\Omega, \mathscr{F})$ and this knowledge is *common to all of them.* We simplify notation by assuming that $m(S)(x) = 0$ for $x \notin C$ and $S \in \mathscr{F}$, while $m(\Omega)(x) = 1$ for all x. The family $m(\cdot)(x)$ is then well defined for all $x \in \Omega$.

Our assumption that agents learn the stationary measure $m(\cdot)(x)$ from a finite but large set of data needs a few comments. First, it follows from the assumption of stability that with a large enough body of data and for a low dimension cylinder S it is possible to obtain an approximation of the limit $\mathring{m}(S)(x)$ to a very high degree of accuracy. In most economic applications of interest agents discount the future and, to any degree of approximation, are concerned only with events which will occur within a finite horizon. Keeping this in mind, let the σ-field of J horizon events at date t be defined by

$$\mathscr{F}_t^{t+J} = \sigma(x_t, x_{t+1}, \ldots, x_{t+J}).$$

Then, there exists a finite J such that the error in economic values of using sets in $\mathscr{F}_t^{t+J}$ instead of $\mathscr{F}_t^{\infty}$ is negligible. If the length of the data set is large relative to J, $\mathring{m}(S)(x)$ for $S \in \mathscr{F}_0^J$ can be approximated to a high degree of accuracy. Since $m(\cdot)(x)$ is stationary, it is the same on all $\mathscr{F}_t^{t+J}$.

A second point is a methodological one. We should think of the limits $\mathring{m}(S)(x)$ as describing the average or *normal* patterns of the dynamics. The

collection of these limits is what is conceivably knowable by all the agents and therefore it is what all agents could agree upon. In practice, one develops algorithms to approximate $\hat{m}(S)(x)$ and diversity of opinions may arise with respect to the quality of the approximation. The idea of endowing the agents with what they can conceivably learn is a methodological simplification which we are making in order to avoid the complication of approximation.

Finally, the merit of our theory depends upon the tractability of the stationary measure $m(\cdot)(x)$. For example, suppose the stochastic process $\{x_t, t = 0, 1, 2, \ldots\}$ is a Markov process with respect to $m(\cdot)(x)$ or, more generally, suppose the conditional probabilities of future events at date t given the past, under the measure $m(\cdot)(x)$, depend only upon $(x_{t-1}, x_{t-2}, \ldots, x_{t-k})$ where k is modest in length. Then agents can estimate an appropriately defined transition function and construct the measure from it. In this case only low dimensional blocks of data need ever be considered.

Definition 4: A dynamical system $(\Omega, \mathscr{F}, \Pi, T)$ is said to be *weak asymptotically mean stationary (WAMS)* if for all cylinders $S \in \mathscr{F}$, the limits

$$\hat{m}_\Pi(S) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \Pi(T^{-k}S) \text{ exist.} \tag{6}$$

It is *strong asymptotically mean stationary* if the limit in (6) holds for all $S \in \mathscr{F}$.

We mentioned earlier that most of the results available in ergodic theory are applicable only to stationary systems. The exception is the small number of papers written about systems which are strong asymptotically mean stationary. Such processes were studied by Dowker [1951], [1955], Rechard [1956], Gray and Kieffer [1980] and Gray [1988]. Also, some studies in Information Theory have employed such processes (see for example Fontana, Gray and Kieffer [1981] and Kieffer and Rahe [1981]). These authors searched for conditions on non-stationary processes which would imply an ergodic theorem. Their central conclusion is that *strong* stability of $(\Omega, \mathscr{F}, \Pi, T)$ and *strong* asymptotically mean stationarity of this system are equivalent and are both equivalent to the existence of an ergodic theorem for $(\Omega, \mathscr{F}, \Pi, T)$. In this paper we assume only weak stability in order to expand as much as possible the set of processes allowed under our theory. The consequence is that the condition of stability does not imply a *standard* ergodic theorem for our system. The following important proposition, entailing a particular version of the ergodic theorem that applies to our system, is the main technical result which will be at the heart of the Rationality Axioms spcified later.

Proposition 2: $(\Omega, \mathscr{F}, \Pi, T)$ is stable if and only if it is weak asymptotically mean stationary.

Proof: To prove that stability implies WAMS let $S \in \mathscr{F}$ be a cylinder. Define

$$1_S(T^k x) = \begin{cases} 1 & \text{if } x \in T^{-k}S \\ 0 & \text{if } x \notin T^{-k}S \end{cases}$$

and consider the following sequence of equalities

$$\frac{1}{n}\sum_{k=0}^{n-1}\Pi(T^{-k}S)=\frac{1}{n}\sum_{k=0}^{n-1}\int_\Omega 1_S(T^k x)\,\Pi(dx)$$

$$=\int_\Omega \frac{1}{n}\sum_{k=0}^{n-1}1_S(T^k x)\,\Pi(dx)$$

$$=\int_\Omega m^n(S)(x)\,\Pi(dx).$$

Now taking limits on both sides and passing to the limit yields

$$\int_\Omega \mathring{m}(S)(x)\,\Pi(dx)=\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\Pi(T^{-k}S)=\mathring{m}_\Pi(S). \tag{7}$$

The condition of stability implies that the integral on the left of (7) exists, hence the limit on the right exists as well. This proves that $(\Omega, \mathscr{F}, \Pi, T)$ is weak asymptotically mean stationary.

The proof that weak asymptotically mean stationarity implies stability requires a proof of an ergodic theorem which is suitable for our situation. In order to do that we adapt the Katznelson and Weiss [1982] proof of the ergodic theorem to our problem. Pick a cylinder S and define

$$\underline{g}_s(x)=\liminf_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}1_S(T^i x)$$

$$\bar{g}_s(x)=\limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}1_S(T^i x).$$

To prove our claim, we need to show that

$$\underline{g}_s(x)=\bar{g}_s(x)\qquad \Pi\text{ a.e.}$$

However, since

$$\underline{g}_s(x)\le\bar{g}_s(x)\ \text{ for all } x$$

and since $\underline{g}_s(x)\ge 0$, it is sufficient to prove that

$$\int_\Omega \bar{g}_s(x)\,\Pi(dx)\le\int \underline{g}_s(x)\,\Pi(dx).$$

By weak asymptotic mean stationarity

$$\lim_{K\to\infty}\frac{1}{K}\sum_{i=0}^{K-1}\Pi(T^{-i}S)=\mathring{m}_\Pi(S)\ \text{ exists.}$$

Hence it is sufficient to prove that

$$\int_\Omega \bar{g}_s(x)\,\Pi(dx)\le\mathring{m}_\Pi(S)\le\int_\Omega \underline{g}_s(x)\,\Pi(dx). \tag{8}$$

Now, since $\bar{g}_s(x) \leq 1$, for any $\varepsilon > 0$ there is an n such that

$$\frac{1}{n} \sum_{i=0}^{n-1} 1_S(T^i x) \geq \bar{g}_s(x) - \varepsilon. \tag{9}$$

We fix $\varepsilon > 0$ and let $n(x)$ be the smallest integer for which (9) holds. Since $\bar{g}_s(x)$ is an invariant function (see Definition 2) we can rewrite (9) as

$$\sum_{i=0}^{n(x)-1} \bar{g}_s(T^i x) \leq \sum_{i=0}^{n(x)-1} 1_S(T^i x) + n(x)\varepsilon. \tag{10}$$

$n(x)$ is finite for all x hence there exists an N for which

$$\Pi(\{x : n(x) > N\}) = \sum_{k=N+1}^{\infty} \Pi(\{x : n(x) = k\}) \leq \varepsilon. \tag{11}$$

The inequality in (11) follows from the fact that the sum goes to 0 as $N \to \infty$. For such an N let

$$B = \{x : n(x) > N\}$$

be the set of "bad" sequences. Note that if $x \in B^c$ then $T^i x \in B^c$ for $i = 1, 2, \ldots, n(x) - 1$. That is, if x is good then the next $n(x) - 1$ shifts are also good. To handle the "bad" sequences in B we define

$$\tilde{1}_S(x) = \begin{cases} 1_S(x) & \text{if } x \notin B \\ 1 & \text{if } x \in B \end{cases} \tag{12a}$$

$$\tilde{n}(x) = \begin{cases} n(x) & \text{for } x \notin B \\ 1 & \text{for } x \in B. \end{cases} \tag{12b}$$

Modifying (10) to account for ((12a)–(12b)) we have

$$\sum_{i=0}^{\tilde{n}(x)-1} \bar{g}_s(T^i x) \leq \sum_{i=0}^{\tilde{n}(x)-1} \tilde{1}_S(T^i x) + \tilde{n}(x)\varepsilon \tag{13}$$

and $\tilde{n}(x) \leq N$. (13) is certainly true for $x \in B$. If $x \notin B$, $T^i x \notin B$ for $i = 1, 2, \ldots, n(x) - 1$ and $\tilde{1}_S(T^i x) = 1_S(T^i x)$, and then (13) follows from (10).

Next choose L large so that $\frac{N}{L} < \varepsilon$ and define $n_k(x)$ inductively by

$$n_0(x) = 0$$

$$n_k(x) = n_{k-1}(x) + \tilde{n}(T^{n_{k-1}(x)} x)$$

and let $k(x)$ be the *largest* k for which $n_k(x) \leq L - 1$. Now, since

$$\sum_{i=0}^{L-1} \bar{g}_s(T^i x) = \sum_{k=1}^{k(x)} \sum_{n_{k-1}(x)}^{n_k(x)-1} \bar{g}_s(T^i x) + \sum_{i=n_{k(x)}(x)}^{L-1} \bar{g}_s(T^i x)$$

we can apply the bound in (13) to each one of the $k(x)$ blocks in the inner sum. That is,

$$\sum_{n_{k-1}(x)}^{n_k(x)-1} \bar{g}_s(T^i x) \leq \sum_{n_{k-1}(x)}^{n_k(x)-1} \tilde{1}_S(T^i x) + (n_k(x) - n_{k-1}(x))\varepsilon. \tag{14}$$

Summing up in (14) we have

$$\sum_{i=0}^{L-1} \bar{g}_s(T^i x) \le \sum_{i=0}^{L-1} \tilde{1}_S(T^i x) + L\varepsilon + (N-1). \tag{15}$$

The term $(N-1)$ accounts for the last (incomplete) block from $n_{k(x)}(x)$ to $L-1$ which is at most $N-1$. Now take expectations in (15) with respect to Π to have

$$\sum_{i=0}^{L-1} \int_\Omega \bar{g}_s(T^i x)\, \Pi(dx) \le \sum_{i=0}^{L-1} \int_\Omega \tilde{1}_S(T^i x)\, \Pi(dx) + L\varepsilon + (N-1). \tag{16}$$

Observe now the following sequence of inequalities

$$\int_\Omega \tilde{1}_S(x)\, \Pi(dx) = \int_{B^c} \tilde{1}_S(x)\, \Pi(dx) + \int_B \tilde{1}_S(x)\, \Pi(dx)$$

$$= \int_{B^c} 1_S(x)\, \Pi(dx) + \int_B 1 \cdot \Pi(dx) \quad \text{by ((12a)–(12b))}$$

$$\le \int_\Omega 1_S(x)\, \Pi(dx) + \Pi(B)$$

$$\le \Pi(S) + \varepsilon \qquad\qquad \text{by (11).} \tag{17}$$

Hence using (17) and inserting into (16) we get

$$\sum_{i=0}^{L-1} \int_\Omega \bar{g}_s(T^i x)\, \Pi(dx) \le \sum_{i=0}^{L-1} \Pi(T^{-i} S) + L\varepsilon + L\varepsilon + (N-1). \tag{18}$$

Now recall that $\bar{g}_s$ is invariant so that

$$\bar{g}_s(T^i x) = \bar{g}_s(x). \tag{19}$$

Recall also that $N/L < \varepsilon$. Now use (19) and divide (18) by L to conclude that

$$\int_\Omega \bar{g}_s(x)\, \Pi(dx) \le \frac{1}{L} \sum_{i=0}^{L-1} \Pi(T^{-i} S) + 3\varepsilon. \tag{20}$$

But since S is a cylinder, weak asymptotic mean stationarity implies that

$$\lim_{L\to\infty} \frac{1}{L} \sum_{i=0}^{L-1} \Pi(T^{-i} S) = \mathring{m}_\Pi(S) \quad \text{exists.} \tag{21}$$

Since L and ε are arbitrary, (20) and (21) prove that

$$\int_\Omega \bar{g}_s(x)\, \Pi(dx) \le \mathring{m}_\Pi(S).$$

This proves the left hand inequality in (8). The right hand inequality is proved by a completely symmetrical argument. ∎

We now seek the final link between the finitely additive measure $\mathring{m}_\Pi$ calculated in (6) from the dynamical system and the family of finitely additive measures $\mathring{m}(\cdot)(x)$ deduced from the data.

Proposition 3: The set function $\mathring{m}_{\Pi}(\cdot)$ on $\mathring{\mathscr{F}}$ can be extended uniquely to a probability measure m_{Π} on $(\Omega, \mathscr{F})$ which is stationary with respect to T and such that for all $S \in \mathscr{F}$

(i) $m_{\Pi}(S \mid \mathfrak{I})(x) = m(S)(x) \quad m_{\Pi}$ a.e.

(ii) $m(S) = m_{\Pi}(S)$

where

$$m(S) = \int_{\Omega} m(S)(x) m_{\Pi}(dx).$$

Proof: It is immediate that $\mathring{m}_{\Pi}$ is a countably additive probability measure on $\mathring{\mathscr{F}}$. It follows from the Carathiodory extension theorem that there exists a unique extension of $\mathring{m}_{\Pi}$ to a measure m_{Π} on $(\Omega, \mathscr{F})$. Now define, for $S \in \mathring{\mathscr{F}}$

$$\mathring{m}(S) = \int_{\Omega} \mathring{m}(S)(x) \Pi(dx)$$

and for all $S \in \mathscr{F}$

$$m(S) = \int_{\Omega} m(S)(x) \Pi(dx).$$

The function $\mathring{m}(\cdot)$ on $\mathring{\mathscr{F}}$ is a finitely additive measure and the function $m(\cdot)$ on $\mathscr{F}$ is a countably additive measure which is an extension of $\mathring{m}$. Since by (7)

$$\mathring{m}(S) = m_{\Pi}(S) \quad \text{for all} \quad S \in \mathring{\mathscr{F}}$$

and since the extension is unique it follows that

$$m(S) = \int_{\Omega} m(S)(x) \Pi(dx) = m_{\Pi}(S) \quad \text{for all} \quad S \in \mathscr{F}. \tag{22}$$

From (6) $\mathring{m}_{\Pi}(T^{-1}S) = \mathring{m}_{\Pi}(S)$ for all cylinders $S \in \mathring{\mathscr{F}}$ and this property is inherited by m_{Π} and hence $(\Omega, \mathscr{F}, m_{\Pi}, T)$ is stationary. It then follows from the ergodic theorem that $m_{\Pi}(C_i) = 1$ where C_i are the convergence sets for F_i (see Proposition 1). Consequently

$$\Pi(C) = m_{\Pi}(C) = m(C) = 1.$$

By the stationarity of m_{Π} we can repeat the argument leading to (7) using m_{Π} instead of Π to conclude that

$$\int_{\Omega} m(S)(x) m_{\Pi}(dx) = m_{\Pi}(S) \quad \text{for all} \quad S \in \mathscr{F}$$

and hence

$$m(S) = \int_{\Omega} m(S)(x) m_{\Pi}(dx) \quad \text{for all} \quad S \in \mathscr{F}.$$

Denote by $m_{\Pi}(\cdot \mid \mathfrak{I})(x)$ the conditional probability of m_{Π} given the σ-field of invariant events. Since m_{Π} is stationary, it is strong asymptotically mean

stationary. It then follows that for all $S \in \mathcal{F}$, $m(S)(x)$ is the limit of the relative frequencies m_Π a.e. Hence from Theorem 6.6.1 of Gray [1988] we conclude that

$$m(S)(x) = m_\Pi(S \mid \mathfrak{J})(x) \quad m_\Pi \text{ a.e.} \qquad \blacksquare$$

The case of an ergodic dynamical system is probably the central case of interest. In this case the limits of the relative frequencies are independent of x. We state this as follows:

Corollary to Proposition 3: If $(\Omega, \mathcal{F}, \Pi, T)$ is ergodic then for all $S \in \mathcal{F}$

$$m(S)(x) = m(S) = m_\Pi(S) \quad \Pi \quad \text{and} \quad m_\Pi \quad \text{a.e.}$$

Proof: We need to prove only that if $(\Omega, \mathcal{F}, \Pi, T)$ is ergodic, $m(S)(x)$ is independent of x Π a.e. for all $S \in \mathcal{F}$. Thus pick $S \in \mathcal{F}$ and define

$$A_C = \{x \in \Omega : m(S)(x) = m_\Pi(S)\}$$
$$A_C^+ = \{x \in \Omega : m(S)(x) > m_\Pi(S)\}$$
$$A_C^- = \{x \in \Omega : m(S)(x) < m_\Pi(S)\}.$$

All three sets are invariant and hence $\Pi(A_C) = 0$ or $\Pi(A_C) = 1$. The only case compatible with (22) is $\Pi(A_C) = 1$. $\qquad \blacksquare$

With the results above in place, we now have a strengthening of Proposition 1:

Proposition 4: The dynamical system $(\Omega, \mathcal{F}, m(\cdot)(x), T)$ is stationary and ergodic m_Π a.e.

Proof: Follows from Gray [1988], Theorem 7.4.1.

Assembling our conclusions we see that when the dynamical system is stable *but not stationary*, the agents who try to learn Π know that it is not learnable; they end up learning the stationary probability $m(\cdot)(x)$. If the system is ergodic $m(\cdot)(x) = m$ independent of x but $m \neq \Pi$. The condition of stability provides the agents additional knowledge which consists of three parts:

(a) A stable Π induces a unique stationary measure m_Π on $(\Omega, \mathcal{F})$ which we shall call "the *stationary measure* of Π."
(b) $m(S)(x) = m_\Pi(S \mid \mathfrak{J})(x)$ for all $S \in \mathcal{F}$ m_Π a.e.
(c) The dynamical system $(\Omega, \mathcal{F}, m(\cdot)(x), T)$ is stationary and ergodic m_Π a.e.

These conclusions are central to our development to follow. However, they also indicate that we have been conducting the analysis under the presumption that the learning agents may observe only one single realization $\{x_0, x_1, x_2, x_3, \ldots\}$ of the dynamical system. Proposition 4 says that even if $(\Omega, \mathcal{F}, \Pi, T)$ is not ergodic the single realization x together with all its iterates $T^t x$ *will all belong to only one ergodic component of Ω* which is an invariant set. But this

means that if F is the ergodic component which contains x, at no time in human history would we ever observe anything outside of F! For practical purposes, we should think of F as the basic space rather than Ω. Defining $\mathscr{F}_F$ to be the restriction of $\mathscr{F}$ to F and Π_F to be the restriction of Π to $\mathscr{F}_F$ the real dynamical system is therefore $(F, \mathscr{F}_F, \Pi_F, T)$ and this dynamical system is ergodic. We then claim that if the data available is a single realization of the dynamical system then from the analytical point of view we may as well assume that the system is ergodic and in that case

(a) $m(S)(x) = m(S)$ for all $S \in \mathscr{F}$, independent of x,

(b) $m(S) = m_{\Pi}(S)$ for all $S \in \mathscr{F}$.

For the sake of generality and analytic unity we shall assume that the true dynamical system $(\Omega, \mathscr{F}, \Pi, T)$ is not necessarily ergodic. However, we shall also assume that the agents know the stationary measure m and, in addition, they know that $m = m_{\Pi}$.

4 The structure of rational beliefs

In most economic applications the concept of "rationality" must be understood with respect to statements about conditional probabilities. However, our approach is based on considerations of ergodic theory rather than on Bayesian statistics and for this reason it is not convenient to specify criteria for the selection of rational *conditional* probabilities. Instead we shall specify two rationality axioms which we apply to the selection of *unconditional* probabilities. We shall then provide a characterization of rational beliefs in terms of unconditional as well as conditional probabilities.

In the previous sections we have endeavored to show that the common empirical knowledge of all the agents is entirely represented by the stationary probability measure m and all rational beliefs should be required to be compatible with this knowledge. Thus, suppose an agent considers if a given stable system $(\Omega, \mathscr{F}, Q, T)$ can conceivably be the true one (which is empirically known to have generated m). From Definition 4 and the extension argument in Proposition 3 we know that $(\Omega, \mathscr{F}, Q, T)$ induces a "theoretical" stationary measure which we denoted by m_Q. Now denote by $\mathscr{P}(\Omega)$ the space of all probabilities on $(\Omega, \mathscr{F})$. This leads to the following:

Definition 5: We say that a probability $Q \in \mathscr{P}(\Omega)$ is *compatible with the data* if $(\Omega, \mathscr{F}, Q, T)$ is stable with a stationary measure m. That is, for all cylinders $S \in \mathscr{F}$

$$m_Q(S) \overset{\mathrm{d}}{=} \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} Q(T^{-k}S) = m(S).$$

We define the *agent's acceptable set* $B(\Pi)$ to be

$$B(\Pi) = \{Q \in \mathscr{P}(\Omega) : Q \text{ is compatible with the data.}\}$$

Proposition 5: $B(\Pi)$ is a convex subset of $\mathscr{P}(\Omega)$ which is a complete and separable metric space.

Proof: Since X^z is a complete and separable metric space, the space $\mathscr{P}(\Omega)$ endowed with the topology of weak convergence is also a complete and separable metric space (see Parthasarathy [1967], Chapter II.6). The convexity of $B(\Pi)$ follows from the definition of stability.   ∎

4.1 Forming beliefs

We presume that each agent makes a probability assessment of events in Ω. Taking a direct approach first, the formation of a belief by an agent about the dynamical system $(\Omega, \mathscr{F}, \Pi, T)$ must result in his selecting a probability P and then presuming that the dynamical system is $(\Omega, \mathscr{F}, P, T)$. This means that he will assign the probability $P(S)$ to each event $S \in \mathscr{F}$. An alternative view point holds that if $\mathscr{P}(\Omega)$ is the set of all probabilities on Ω, then the object of uncertainty is $\mathscr{P}(\Omega)$ itself and hence *forming a belief* about Π would necessitate *the agent's selecting a probability P^* on $\mathscr{P}(\Omega)$*. That is, if $\mathscr{P}(\mathscr{P}(\Omega))$ is the space of all probabilities on $\mathscr{P}(\Omega)$ then the agent must select $P^* \in \mathscr{P}(\mathscr{P}(\Omega))$. P^* is an agent specific, or "subjective," probability over the set of all possible probabilities which the agent may adopt for his decision-making. We would then define the expectations to be

$$P = \int_{\mathscr{P}(\Omega)} \mu P^*(d\mu)$$

(if such an integral makes sense) and say that the agent *forms the belief P^** with hich the probability P is selected. Given this we would also say that *the agent believes that the dynamical system is $(\Omega, \mathscr{F}, P, T)$*. We stress that this selection procedure is *fixed* and is not updated with additional data. This is the case since P^* itself is already the limit of the updated beliefs given all the wealth of data which we have provided the agent in the first place.

We now claim that if the support of P^* is $B(\Pi)$ rather than $\mathscr{P}(\Omega)$, then the two procedures above are equivalent. Since the basic axiom of rationality (to be presented below) requires the agents to select beliefs only from $B(\Pi)$, this will provide us a *representation result* for rational beliefs. Let $C(\mathscr{P}(\Omega))$ be the space of continuous and bounded real valued functions f: $\mathscr{P}(\Omega) \to \mathbb{R}$. Since $B(\Pi) \subset \mathscr{P}(\Omega)$ then for each $P^* \in \mathscr{P}(\mathscr{P}(\Omega))$ the integral

$$P^*(f) = \int_{B(\Pi)} f(\mu) P^*(d\mu)$$

is a well defined linear functional on $C(\mathscr{P}(\Omega))$. We say that P^* *represents* $P \in B(\Pi)$ if

$$P^*(f) = f(P) \quad \text{for all} \quad f \in C(\mathscr{P}(\Omega)).$$

Now define the collection of real valued functions f^S: $\mathscr{P}(\Omega) \to \mathbb{R}$ by

$$f^S(\mu) = \mu(S), \quad \mu \in \mathscr{P}(\Omega), \quad S \in \mathscr{F}. \tag{23}$$

Proposition 6: (*Representation of Acceptable Beliefs*). For any $P^* \in \mathscr{P}(B(\Pi))$ let

$$P(S) = \int\limits_{B(\Pi)} f^S(\mu)P^*(d\mu) \quad \text{all} \quad S \in \mathscr{F}. \tag{24}$$

Then $P \in B(\Pi)$; conversely, if $P \in B(\Pi)$ there exists $P^* \in \mathscr{P}(B(\Pi))$ such that P^* represents P and the support of P^* is the set of extreme points in $B(\Pi)$.

Proof: Let $P^* \in \mathscr{P}(B(\Pi))$. If $S = \Omega$ we have $f^\Omega(\mu) = \mu(\Omega) = 1$ and if $S = \varnothing$ then $f^\varnothing = \mu(\varnothing) = 0$. Hence $P(\Omega) = 1$ and $P(\varnothing) = 0$. To check for σ-additivity let $S = \bigcup\limits_{i=1}^{\infty} S_i$ with $S_i \cap S_j = \varnothing$ for $i \neq j$ and consider $P(S)$. Since $f^S(\mu) = \lim\limits_{n \to \infty} \sum\limits_{i=1}^{n} \mu(S_i) = \lim\limits_{n \to \infty} \sum\limits_{i=1}^{n} f^{S_i}(\mu)$ and since these limits exist P^* a.e., it follows from the bounded convergence theorem that

$$P(S) = \lim\limits_{n \to \infty} \int\limits_{B(\Pi)} \sum\limits_{i=1}^{n} f^{S_i}(\mu)P^*(d\mu) = \sum\limits_{i=1}^{\infty} P(S_i).$$

This proves that P is a probability. To prove that $P \in B(\Pi)$ let S be a cylinder. Then

$$\frac{1}{n} \sum\limits_{k=0}^{n-1} P(T^{-k}S) = \int\limits_{B(\Pi)} \left(\frac{1}{n} \sum\limits_{k=0}^{n-1} f^{T^{-k}S}(\mu) \right) P^*(d\mu). \tag{25}$$

But now from (23) and using the fact that $\mu \in B(\Pi)$ we have

$$\lim\limits_{n \to \infty} \frac{1}{n} \sum\limits_{k=0}^{n-1} f^{T^{-k}S}(\mu) = \lim\limits_{n \to \infty} \frac{1}{n} \sum\limits_{k=0}^{n-1} \mu(T^{-k}S) = m(S) \qquad P^* \quad \text{a.e.}$$

Again, by the bounded convergence theorem

$$m(S) = \int\limits_{B(\Pi)} \left(\lim\limits_{n \to \infty} \frac{1}{n} \sum\limits_{k=0}^{n-1} f^{T^{-k}S}(\mu) \right) P^*(d\mu) = \lim\limits_{n \to \infty} \int\limits_{B(\Pi)} \left(\frac{1}{n} \sum\limits_{k=0}^{n-1} f^{T^{-k}S}(\mu) \right) P^*(d\mu). \tag{26}$$

Combining (25) and (26) we conclude that for all cylinders S

$$\lim\limits_{n \to \infty} \frac{1}{n} \sum\limits_{k=0}^{n-1} P(T^{-k}S) = m(S)$$

and this shows that $P \in B(\Pi)$.

To prove the converse let $P \in B(\Pi)$. By the properties of $B(\Pi)$ proved in Proposition 5 the integral representation theorem of Choquet (see Choquet [1969] Proposition 26.3 or for an exposition of the theory, see Phelps [1980] pp. 115–157) becomes applicable. It implies that there exists $P^* \in \mathscr{P}B(\Pi))$ such that P^* represents P and has the claimed property. ∎

4.2 Axioms of rationality

Before introducing our axioms of rationality define, for each $S \in \mathscr{F}$, the set

$$B_s = \{P \in B(\Pi): P(S) > 0\}.$$

The axioms specify conditions which a rational choice of P^* must satisfy.

Axiom 1 (Compatibility with the data): An agent forms a belief P^* with $B(\Pi)$ as its support.

Axiom 2 (Continuity with respect to the data): If for $S \in \mathscr{F}$ $m(S) > 0$ then $P^*(B_S) > 0$.

Discussion. Axiom 1 is the crucial axiom of rationality. It requires of any agent to form a belief P^* which places probability 1 on $B(\Pi)$. It then follows from Proposition 5 that this is equivalent to a selection of $P \in B(\Pi)$. Axiom 2 further restricts the allowable P's within $B(\Pi)$. We think of Axiom 2 as a continuity axiom. To see this note that if S is a cylinder and $m(S) > 0$ Axiom 1 together with the definition of stability imply that P^* assigns positive probability to the set of measures $P \in B(\Pi)$ which satisfy a particular requirement. That is, that there exists an $\varepsilon > 0$ such that for an infinite set of dates, k, constituting a positive fraction of all the positive integers $P(T^{-k}S) \geq \varepsilon > 0$. Apart from being intuitively appealing, the main function of Axiom 2 is to prevent the degeneracy of belief to a probability mass at a single point when m gives the point zero probability.

Before proceeding to our next proposition we dispose of a few technicalities. First, add to the notation in Sections 2 and 3 the notation indicating the history *up to* t: $x_{(t)} = (x_0, x_1, \ldots, x_{t-1})$. Next, we write the conditional probability of a future event $S \in \mathscr{F}_t^x$ given $x_{(t)}$ in the form $P^t(S | x_{(t)})$. We remark that the technical issue of selecting regular conditional probabilities P^t and Q^t should be entirely disregarded here since $\Omega = X^\infty$ is a complete and separable metric space (see Blackwell and Dubins [1975] and Ash [1972] page 265). In addition, we are interested in $P^t(\cdot | x_{(t)})$ for large t and this motivates

Definition 6: Regular conditional probabilities $P^t(\cdot | x_{(t)})$ and $Q^t(\cdot | x_{(t)})$ on $(\Omega, \mathscr{F})$ are said to agree for Q almost all histories if

$$\limsup_{\tau \to \infty} {}_{S \in \mathscr{F}_\tau^x} |P^\tau(S | x_{(\tau)}) - Q^\tau(S | x_{(\tau)})| = 0 \quad Q \quad \text{a.e.}$$

and we write $P^t \approx Q^t$ Q a.e. If they agree Q a.e. and P a.e. then we write $P^t \approx Q^t$ Q a.e., P a.e.

The notation $P \ll Q$ is used to indicate that a probability measure P is absolutely continuous with respect to a probability measure Q: $P \perp Q$ to indicate that they are singular and $P \ll Q$ as well as $Q \ll P$ to indicate that they are equivalent. We now state the important result:

Proposition 7: (Blackwell and Dubins [1962]). Suppose that P and Q are probability measures on $(\Omega, \mathscr{F})$ and $Q \ll P$. Then for each t and for every regular conditional probability P^t of the future given the past there exists a corresponding conditional probability Q^t such that

$$P^t \approx Q^t \quad Q \quad \text{a.e.}$$

To see the significance of Proposition 7 for us, suppose that two agents hold probability beliefs P and Q with $P \neq Q$ but P and Q are equivalent. Since

these two probabilities are required to agree only on the null sets, their numerical values on other sets may be drastically different and hence it appears that ample ground is left for diversity of opinions. In fact, it is almost standard in the economics literature to assume the equivalence of subjective probability beliefs whenever heterogeneity of beliefs is introduced (see, for example, Harrison and Kreps [1979] where this assumption is crucial). In our context ample past data is available and therefore it follows from Proposition 6 that, although $P \neq Q$, $P^t(S|x_{(t)})$ is essentially equal to $Q^t(S|x_{(t)})$ for all $S \in \mathscr{F}_t^x$. In our context mathematical equivalence implies an essential economic equivalence[5].

4.3 The main theorem

Main Theorem: Given a dynamical system $(\Omega, \mathscr{F}, \Pi, T)$ let an agent form a rational belief P^* which satisfies Axioms 1 and 2. Then $P^* \in \mathscr{P}(B(\Pi))$ and there exists a probability $P \in B(\Pi)$ which is the expectation under P^* in the sense of (24). Moreover, there exist probabilities P_a and P_0 on $(\Omega, \mathscr{F})$ and a constant $0 < \lambda_p \leq 1$ such that

(i) P has a unique representation

$$P = \lambda_p P_a + (1 - \lambda_p) P_0 \qquad (27a)$$

where P_a and m are equivalent while P_0 and m are singular; that is, there exist sets A and B with $A \cap B = \varnothing$, $A \cup B = \Omega$ such that

$$m(A) = 1 \quad \text{and} \quad P_0(B) = 1.$$

(ii) $(\Omega, \mathscr{F}, P_a, T)$ and $(\Omega, \mathscr{F}, P_0, T)$ are stable with stationary measures $\bar{P}_a$ and $\bar{P}_0$ such that for all $S \in \mathscr{F}$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P_a(T^{-k}S) = \bar{P}_a(S) \qquad (27b)$$

and for all cylinders S

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P_0(T^{-k}S) = \overset{\circ}{\bar{P}}_0(S) \quad \text{and} \quad \bar{P}_0 \text{ is the extension of } \overset{\circ}{\bar{P}}_0. \qquad (27c)$$

In addition, $\bar{P}_a$ and m are equivalent

$\bar{P} \ll m$.

If $(\Omega, \mathscr{F}, \Pi, T)$ is ergodic and $0 < \lambda_p < 1$ then we have $\bar{P}_a = \bar{P}_0 = m$.

---

[5] In a model of learning in games, Kalai and Lehrer [1993] and Nyarko [1992], [1993] assume the absolute continuity of the equilibrium probability measure with respect to the belief of every player. Feldman [1987] makes a similar assumption in a market context. From our perspective this amounts to assuming that players know *at the outset* what the model sets for them to learn.

(iii) There exist regular versions of the conditional probabilities P^t, m^t and P_0^t and densities

$$\psi_m = \frac{dm}{dP} \quad \psi_0 = \frac{dP_0}{dP}$$

which satisfy m a.e. for $S \in \mathscr{F}_t^x$

$$P^t(S | x_{(t)}) \approx \lambda_p m^t(S | x_{(t)}) \psi_m(x_{(t)}) + (1 - \lambda_p) P_o^t(S | x_{(t)}) \psi_0(x_{(t)}). \quad (27d)$$

And conversely: For any P, P_a and P_0 satisfying conditions (i)–(iii) there exists a rational belief $P^* \in \mathscr{P}(B(\Pi))$ which represents P.

Proof of Main Theorem

Let P^* be a rational belief. By Axiom 1 $P^* \in \mathscr{P}(B(\Pi))$ and by Proposition 6 there exists $P \in B(\Pi)$ which P^* represents.

Next we consider the decomposition of P. It follows from the Lebesgue Decomposition Theorem (See Royden [1988], page 278) that there exist probabilities P_a and P_0 on $(\Omega, \mathscr{F})$, sets $A \subset \Omega$ and $B = \Omega - A$, and a constant $0 \leq \lambda_p \leq 1$ such that

$$P = \lambda_p P_a + (1 - \lambda_p) P_0$$

where $P_a \ll m$, $P_0 \perp m$, $\lambda_p = P(A)$, $m(A) = 1$ and $m(B) = 0$. For any set $S \in \mathscr{F}$, if $P(A) > 0$ and $P(B) > 0$ we have

$$P_a(S) = \frac{P(A \cap S)}{P(A)} \quad P_0(S) = \frac{P(B \cap S)}{P(B)}.$$

By Axiom 2, $m(A) = 1$ implies $P(A) > 0$ and hence $0 < \lambda_p \leq 1$. We do not exclude $\lambda_p = 1$ and $P(B) = 0$.

We shall now show that P_a and m are equivalent. From the Lebesgue decomposition theorem we already have that $P_a \ll m$. To prove that $m \ll P_a$ suppose that it is false. Thus let $S \in \mathscr{F}$ and $P_a(S) = 0$ while $m(S) > 0$. From Axiom 2 it follows that $P(S) > 0$. But then we have

$$0 < P(S) = \lambda_p P_a(S) + (1 - \lambda_p) P_0(S) = (1 - \lambda_p) P_0(S)$$

$$0 < m(S) = m(S \cap A) + m(S \cap B) = m(S \cap A).$$

Hence $m(S \cap A) > 0$ and $P_0(S) > 0$. Now consider $\hat{S} = S \cap A$. Clearly $m(\hat{S}) > 0$ but $\hat{S} \subset S$ implies $P_a(\hat{S}) = 0$ and $\hat{S} \subset A$ implies $P_0(\hat{S}) = 0$ (P_0 and m are singular). Hence $P(\hat{S}) = 0$ and this contradicts Axiom 2.

We now demonstrate that $(\Omega, \mathscr{F}, P_a, T)$ and $(\Omega, \mathscr{F}, P_0, T)$ are stable with stationary measures $\bar{P}_a$ and $\bar{P}_0$ and that $\bar{P}_a$ and m are equivalent. For $S \in \mathscr{F}$ we have that

$$\frac{1}{n} \sum_{k=0}^{n-1} P(T^{-k} S) = \lambda_p \frac{1}{n} \sum_{k=0}^{n-1} P_a(T^{-k} S) + (1 - \lambda_p) \frac{1}{n} \sum_{k=0}^{n-1} P_0(T^{-k} S). \quad (28)$$

Since $P \in B(\Pi)$ the left hand side of (28) converges to $m(S)$ for all cylinders

$S \in \mathscr{F}$. Now since $P_a \ll m$ and since m is stationary it follows from Theorem 2 of Gray and Kieffer [1980] that P_a is strong asymptotically mean stationary and hence for all $S \in \mathscr{F}$ the limit

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P_a(T^{-k}S) = \bar{P}_a(S) \text{ exists.} \tag{28'}$$

Combining (28) and (28') leads to the conclusion that for all cylinders $S \in \mathscr{F}$ the limit

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} P_0(T^{-k}S) = \overset{\circ}{\bar{P}_0}(S) \text{ exists.}$$

Now extend $\overset{\circ}{\bar{P}_0}$ to a measure $\bar{P}_0$ on $(\Omega, \mathscr{F})$. This implies that for all $S \in \mathscr{F}$

$$m(S) = \lambda_p \bar{P}_a(S) + (1 - \lambda_p) \bar{P}_0(S). \tag{29}$$

This shows that both $(\Omega, \mathscr{F}, P_a, T)$ and $(\Omega, \mathscr{F}, P_0, T)$ are stable dynamical systems with stationary measures $\bar{P}_a$ and $\bar{P}_0$. It is immediate from (29) that $\bar{P}_0 \ll m$ and $\bar{P}_a \ll m$. We need to prove that and $\bar{P}_a$ and m are equivalent. To prove that $m \ll \bar{P}_a$ assume the contrary and select $S \in \mathscr{F}$ with $\bar{P}_a(S) = 0$ but $m(S) > 0$. Since P_a is equivalent to m, $P_a(S) > 0$. Now define

$$\hat{S} = \bigcap_{n=0}^{\infty} \bigcup_{k=n}^{\infty} (T^{-k}S)$$

$$= \limsup_{k \to \infty} (T^{-k}S).$$

We claim that $\bar{P}_a(\hat{S}) = 0$. This is so since

$$\bar{P}_a(\hat{S}) = \lim_{n \to \infty} \bar{P}_a\left(\bigcup_{k=n}^{\infty} T^{-k}S \right) \leq \bar{P}_a\left(\bigcup_{k=1}^{\infty} T^{-k}S \right) \leq \sum_{k=1}^{\infty} \bar{P}_a(T^{-k}S) = 0.$$

But $\hat{S}$ is an invariant set and since P_a is strong asymptotically mean stationary it follows from Lemma 6.3.1 of Gray [1988] that $\bar{P}_a(\hat{S}) = P_a(\hat{S}) = 0$. Hence $m(\hat{S}) = 0$. By Fatou's Lemma

$$m(\hat{S}) = m\left(\limsup_{k \to \infty} T^{-k}S \right) \geq \limsup_{k \to \infty} m(T^{-k}S) = m(S) > 0$$

and this is a contradiction hence $m \ll \bar{P}_a$. This concludes the proof that $\bar{P}_a$ and m are equivalent.

To prove that in the ergodic case $\bar{P}_a = \bar{P}_0 = m$ it is sufficient to prove that $\bar{P}_a = m$ since the conclusion follows from (29) and $0 < \lambda_p < 1$. To prove $\bar{P}_a = m$ recall that $P_a \ll m$ hence there exists an m-integrable function g such that for all $S \in \mathscr{F}$

$$P_a(S) = \int_S g(\omega) m(d\omega).$$

Hence

$$P_a(T^{-k}S) = \int_{T^{-k}S} g(\omega) m(d\omega).$$

By the change of variables thorem and the stationarity of m we have that

$$\int_{T^{-k}S} g(\omega)m(d\omega) = \int_s g(T^k\omega)(m\,T^{-k})(d\omega) = \int_s g(T^k\omega)m(d\omega).$$

It then follows that

$$\bar{P}_a(S) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P_a(T^{-k}S) = \int_s \left(\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(T^k\omega) \right) m(d\omega).$$

Since P_a is strong asymptotically mean stationary it follows from Theorem 7.2.1 of Gray [1988] and the assumption that $(\Omega, \mathscr{F}, P_a, T)$ is ergodic that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g(T^k\omega) = E_m g = 1.$$

Hence

$$\bar{P}_a(S) = m(S) \quad \text{all} \quad S\in\mathscr{F}.$$

We finally turn to the conditional probabilities. Since Ω is a complete and separable metric space it follows from Proposition 7 that for all $S\in\mathscr{F}^t$

$$P_a^t(S\,|\,x_{(t)}) \approx m^t(S\,|\,x_{(t)}) \quad \text{a.e.} \quad m \text{ and } P_a.$$

A standard argument leads to

$$P^t = \lambda_p P_a^t \psi_m + (1 - \lambda_p) P_0^t \psi_0$$

where $\psi_m = \dfrac{dm}{dp}$ and $\psi_0 = \dfrac{dP_0}{dP}$. We thus conclude that

$$P^t \approx \lambda_p m^t \psi_m + (1 - \lambda_p) P_0^t \psi_0 \quad \text{a.e.} \quad m \text{ and } P_a.$$

To prove the converse let P, P_a, P_0 and λ_p be as in conditions (i)–(iii). Clearly P is stable hence $P\in B(\Pi)$. It then follows from Proposition 6 that there exists $P^*\in\mathscr{P}(B(\Pi))$ such that P^* represents P. This proves Axiom 1. Next let $S\in\mathscr{F}$ satisfy $m(S) > 0$. Since $P_a \ll m$ and $m \ll P_a$ it follows that $P_a(S) > 0$. Since $\lambda_p > 0$ it follows that $P(S) > 0$. This implies that $P^*(B_S) > 0$ thus proving Axiom 2. ∎

5 Some comments and examples

In the introduction we asked why should two rational agents with the same information end up with different probability beliefs. The significance of the Main Theorem is that it shows not only *why* the two agents may have different beliefs but also *how* these beliefs will differ. To explain this we refer to the representation (27d) in terms of conditional probabilities in which case we may as well replace P_a with m:

(a) First, the agents may have different degrees of confidence in the validity of the empirically generated stationary probability m. These levels of con-

fidence are measured by the subjective parameters λ_p. Thus, the degree of belief in the stationarity of the environment is translated into a degree of confidence in the empirical evidence m.

(b) Second, agents may have drastically different singular measures P_0. Since these measures are not stationary they represent beliefs in what we commonly call "structural changes." This means that important but infrequent events may be assigned significant probabilities under P_0 even if their probabilities under m are very small (or even zero). Although these singular measures may be drastically different for different agents, the condition of rationality requires that they satisfy common asymptotic stability conditions which are compatible with the stationary measure m.

The Main Theorem also highlights the main consequences of our departure from the Rational Expectations framework. It says that when agents do not have structural knowledge and we use only statistical regularity as a foundation for a rational theory of belief, then *objective* rationality criteria can provide only asymptotic restrictions. These restrictions generate sets like $B(\Pi)$. Consequently, the selection by each agent of a particular member of $B(\Pi)$ must be based on *subjective* criteria which represent *individual* "theories" about the environment. In applications we would identify a "society" by the distribution of beliefs in $B(\Pi)$ and such distributions could be as important as the distribution of preferences for the explanation of economic performance.

A word about the advantage of our approach over other learning theories such as Bayesian learning or least squares learning. We take a non-parametric approach and thus avoid the artificial construct of a parameter space. More important is the fact that we demand from a rational agent to form a belief only after he has, so to speak, "mined" the data. By doing so we declare as *irrational* any Bayesian who selects a prior such that his limit posterior is not in $B(\Pi)$. For example, if a prior stipulates that the x_t are i.i.d. all posteriors will continue to maintain such a stipulation. Such a posterior would be contradicted by a stationary measure m if under it x_t and x_{t+k} are correlated for some $k > 0$. Hence, the Main Theorem insists that rational *frequentist* criteria be imposed on the selection of a prior. The Main Theorem also suggests that when an econometrician formulates a stationary least squares learning model then the condition of stability will ensure the convergence of the empirical moments. However, rational agents know that such a procedure enables them merely to learn something about the stationary probability m. Since m and Π may not be the same, these agents may doubt the validity of any least squares learning model and adopt as their beliefs an object as in the Main Theorem. In a general equilibrium context this would invalidate any stationary model while, at the same time, permit the estimated parameters of such a least squares stationary model to converge.

We conclude by considering three examples. These aim to further clarify certain aspects of the concept of "rational beliefs."

5.1 Coin tossing (1): Singularity of stable Q_0 and m.

We give an example of a measure Q_0 which is stable, singular with m and has m as its stationary measure. Thus, consider the sequence of i.i.d. random variables x_t such that under the true measure $P = m$

$$x_t = \begin{cases} 1 & \text{with probability } 1/2 \\ 0 & \text{with probability } 1/2. \end{cases}$$

Now consider a belief Q_0. Let $D = \{t_1, t_2, t_3, \ldots\}$ be an infinite sequence of "remote" dates such that $t_1 \geq 2$, $t_n \geq 2t_{n-1}$ for all $n > 1$. Define the random variables z_t under Q_0:

If $t \notin D$
$$z_t = \begin{cases} 1 & \text{with probability } 1/2 \\ 0 & \text{with probability } 1/2. \end{cases}$$

If $t \in D$
$$z_t = \begin{cases} 1 & \text{with probability } 1/3 \\ 0 & \text{with probability } 2/3. \end{cases}$$

Note that for any cylinder set S the future set $(T^{-k}S)$ will fall on fewer and fewer dates in D and since these distances are more than geometric it follows that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} Q_0(T^{-k}S) = m(S).$$

It is then clear that $(\Omega, \mathscr{F}, Q_0, T)$ is stable with m as its stationary measure. Nevertheless we shall now indicate that Q_0 and m are singular. To see this consider the following B set:

$$B = \left\{ x: \begin{array}{l} \text{the frequency of "1" at the infinite dates in} \\ D \text{ is } 1/3. \end{array} \right\}.$$

By the strong law of large numbers $m(B) = 0$ and $P_z(B) = 1$. Now consider the set

$$A = \left\{ x: \begin{array}{l} \text{the frequency of "1" in the sequence } (x_0, x_1, x_2, \ldots) \\ \text{is } 1/2 \text{ and the same frequency holds among the infinite} \\ \text{number of dates in } D \end{array} \right\}.$$

Here we must have $m(A) = 1$, $Q_0(A) = 0$ and this violates Axiom 2. However, any belief Q of the form $Q = \lambda_p m + (1 - \lambda_p)Q_0$ with $\lambda_p > 0$ is compatible with the Main Theorem.

5.2 Coin tossing (2): Singularity with m

The example above hints how delicate the condition of singularity of measures on sequence spaces is. To highlight this point consider a set of dates D as above and define the stochastic process $\{x_t, t = 0, 1, 2, \ldots\}$, where $x_t \in \{0, 1\}$, by the conditions

$$P\{x_t = 1\} = \begin{cases} \alpha + \varepsilon & \text{if } t \notin D \\ \alpha & \text{if } t \in D. \end{cases}$$

This defines the probability P. The stationary measure m is the probability on i.i.d. sequences where the probability of 1 is α and it is then seen that for any $\varepsilon > 0$ and any set of remote dates D, P is singular with m.

5.3 Coin tossing (3): Persistent non-stationarity

We consider now an example where persistent non-stationarity is represented by a process of switching regimes. Since such a process of regime switching must satisfy the stability conditions we employ a procedure of unobserved parameters. To do that pick an i.i.d. process $\{y_t, t = 0, 1, 2, \ldots\}$ of random variables in $\{0, 1\}$ with probability of 1 being, say, 1/4. Now generate a realization $(y_0^*, y_1^*, y_2^*, \ldots)$ and define the process $\{x_t, t = 0, 1, 2, \ldots\}$ where $x_t \in X = \{0, 1\}$ by

$$P\{x_t = 1\} = \begin{cases} \alpha & \text{if } y_t^* = 1 \\ \beta & \text{if } y_t^* = 0 \end{cases} \quad \text{independently over time.} \tag{30}$$

The *non-stationary* process $\{x_t, t = 0, 1, 2, \ldots\}$ identifies a probability Π_{y^*} on $(X^\infty, \mathscr{F}(X^\infty))$ and a stable dynamical system denoted by $(X^\infty, \mathscr{F}(X^\infty), \Pi_{y^*}, T)$ with a stationary measure m represented by the i.i.d. process $\{v_t, t = 0, 1, 2, \ldots\}$, $v_t \in X$ with $P\{v_t = 1\} = (1/4)\alpha + (3/4)\beta$. Each realization y^* is a sequence of structural parameters which selects, at each date, either the α or the β "coin". This regime switching generates the non-stationarity of the process $\{x_t, t = 0, 1, 2, \ldots\}$.

A belief in the stationary measure m represents the view that the data is generated by a single coin which is tossed as an i.i.d. process with the probability of 1 being $(1/4)\alpha + (3/4)\beta$. This is wrong but it is a rational belief. There are, however, other rational beliefs. To construct one define a process of *private signals* $\{z_t, t = 0, 1, 2, \ldots\}$ of i.i.d. random variables on $\{0, 1\}$ with probability of 1 being ζ. Using a realization z^* of private signals with frequency of $z_t^* = 1$ being ζ an agent can now define the *perceived process* $\{x_t', t = 0, 1, 2, \ldots\}$ by

$$P\{x_t' = 1\} = \begin{cases} \alpha' & \text{if } z_t^* = 1 \\ \beta' & \text{if } z_t^* = 0 \end{cases} \quad \text{independently over time.} \tag{31}$$

(31) defines Q: it is a rational belief relative to Π_{y^*} if $\zeta\alpha' + (1 - \zeta)\beta' = (1/4)\alpha + (3/4)\beta$.

In the example above it is important to see that the realizations z_t^* are strictly *private* signals and not objective "data": the agent is the only one who perceives and understands these signals and the signals z_t^* are not part of the public record of "data". On the other hand if at date t^0 the agent adopts a rational belief about the process $\{x_t, t \geq t^0\}$, then the future realization of the process could be used to statistically test the validity of the agent's belief. For example, in (31) the agent can consider future dates in his own life when $z_t^* = 1$ and discover that the relative frequency of the x's may not be α'. With finite data at his disposal, the agent cannot be certain that his belief is "wrong".

Moreover, in any economic model in which an agent lives for a finite life and in which the number of regimes is large (so that the relative frequency of any one of them is small), *it is a fact* that by the time an agent with a positive discount rate has any statistical evidence with which to question his own belief, it is too late since all of his important decisions have already been made. On the other hand, using long time series of past data both agents as well as econometricians can study, in retrospect, the empirical *effects* of "wrong" past beliefs on the performance of markets.

We finally remark that the possibility that agents with rational beliefs may hold "wrong" beliefs is not unique to our theory and is analogous to situations where Bayesian agents hold "wrong" posterior beliefs. Just consider a Bayesian agent who starts with a prior which leads to posterior beliefs which either do not converge or that converge to the wrong posterior.

5.4 Other deterministic patterns

The handling of deterministic patterns requires an examination of subsequences of x. Suppose that $\{x_{t_k}, k = 0, 1, \ldots\}$ is a subsequence. We can compute relative frequencies of the subsequence and if they converge, they may contradict the stationary measure m. We stress that, in general, relative frequencies on subsequences may not converge. Moreover, even if they converge, no general theory exists about the behavior of the system on subsequences and consequently two rational agents may disagree about the interpretation of such patterns of behavior. Keep in mind that given a sequence of non-degenerate i.i.d random variables with probability one we can find subsequences which will contradict m. It is then clear that the only relevant subsequences are those which contradict m and are also *predictable*. That is, a subsequence has a *predictable index set* (in short "predictable") if there exists a time invariant function f such that for all $k = 1, 2, \ldots$

$$t_{k+1} = f(t_k) \quad \text{hence} \quad t_k = f^k(t_1)$$

where f^k is the k iterate of f. If a rational agent computes relative frequencies on predictable subsequences and these contradict m, he would supplement m with this knowledge.

The main class of predictable subsequences is the one used to study seasonality and deterministic cycles and for this class we have a complete theory. The class is characterized by a function f of the linear form

$$t_{k+1} = t_k + n.$$

In this case we consider the shift transformations T^n instead of T. For example, quarterly data with seasonality should be studied as a sequence of quadruples $((x_0, x_1, x_2, x_3), (x_4, x_5, x_6, x_7), \ldots)$ under the four-shift T^4 rather than under T. It can be shown that if $(\Omega, \mathscr{F}, \Pi, T)$ is stable with a stationary measure m then $(\Omega, \mathscr{F}, \Pi, T^n)$ is stable for all n with the stationary measure $m_{(n)}$. The

relationship between m and $m_{(n)}$ is

$$m(A) = \frac{1}{n} \sum_{i=0}^{n-1} m_{(n)}(T^{-i}A) \qquad A \in \mathcal{F}.$$

If $m = m_{(n)}$ for all n then the data contains no seasonality or deterministic cycles. If $m \neq m_{(n)}$ then rational agents will use the structure of $m_{(n)}$ instead of m. In this paper we assume that the data has been adjusted for such patterns.

There could be predictable subsequences which are not linear shifts. For example, let $\{\xi_t, t = 0, 1, 2, \ldots\}$ be an i.i.d. sequence on the interval $[0, 1]$ and let $\{\eta_t, t = 0, 1, 2, \ldots\}$ be an i.i.d. sequence on the interval $[2, 3]$. Now consider the composition of these two sequences

$$x_t = \begin{cases} \eta_t & \text{for} \quad t = 2^k \\ \zeta_t & \text{for} \quad t \neq 2^k. \end{cases} \tag{32}$$

Note first that the distinguishing characteristic of (32) compared with any of the linear shifts above is that the realizations of x_t along the predictable subsequence $t = 2^k$ *has no effect on the stationary measure m*. It then appears that the restriction imposed by (32) on the data is not reflected in m. From the formal point of view the study of the subsequence of data at dates $t = 2^k$ can be transformed into the study of the linear shifts T^n in the following way. Let

$$y_\tau = \begin{cases} \eta_{2^k} & \text{for} \quad \tau = 2k \qquad k = 0, 1, 2, \ldots \\ \zeta_{2^k} & \text{for} \quad \tau = 2k + 1 \qquad k = 0, 1, 2, \ldots \end{cases}$$

Hence

$$y = (\eta_{2^0}, \zeta_{2^0}, \eta_{2^1}, \zeta_{2^1}, \ldots, \eta_{2^j}, \zeta_{2^j}, \ldots).$$

Treating y as a stochastic sequence we study its asymptotic properties on even and odd dates by comparing the outcomes under T and T^2. Under the i.i.d. assumptions made above m and $m_{(2)}$ are different and their interpretation is clear. However, in general, suppose that relative frequencies are calculated along all dates and compared to those calculated along the dates $t = 2^k$. If both converge and yield different stationary measures then an agent may obtain useful information from this difference. From the formal viewpoint this is exactly the same as in the case of linear shifts. The difference is that without the i.i.d. assumption above, we do not know how to interpret the asymptotic frequencies of y along the dates $t = 2^k$ and consequently, two rational agents would be allowed – under our theory – to have different views about the meaning of the observations. We think that there are two ways to consider the issue raised by example (32).

We start with the purely theoretical perspective. We have already pointed out that the central characteristic of the predictable subsequences such as in (32) is that the dates $t = 2^k$ are "remote" in the sense that the realizations of x_t at these dates have no effect on the stationary measure and for this reason the measure m will place zero probability on the events defined by (32) at dates

$t = 2^k$. Note that we may extend this example. For example, modify (32) and have the process take *different* values at dates $t = 2^{k^k}$ which is a predictable subsequence of $t = 2^k$ and *further different* values on a predictable subsequence of these dates, etc. Assuming that the relative frequencies converge on each of these infinite number of subsequences we shall get a hierarchy of probability measures each having the property of being a probability measure on subsets of zero measure of the probability which precedes it in the order. More specifically, the probability measure calculated from the data realized at the predictable subsequence of dates $t = 2^k$ is defined on subsets of zero m measure and the probabilities calculated from data realized at $t = 2^{k^k}$ is defined on subsets of zero measure of the second measure in the order. Thus, the theoretical solution which responds to example (32) is to impose on the measure m the added empirical restrictions implied by the entire hierarchy of measures – singular with m – which the data on predictable subsequences may imply.

One may, however, consider a different view and we offer it by making three observations. First, our theory is based on the perspective that a rational belief is formulated *given a long history of the data*. It then follows that any predictable subsequence like (32) which has no effect on the stationary measure, has no significance by the time we reach date t at which belief are formed. To see this note that since by assumption, the starting date $t = 0$ is far in the past, for large t the frequency of *future dates* $\tau > t$ at which x_τ satisfies $x_\tau = \eta_\tau$ is, practically speaking, zero. Keep in mind that the stationary measure is calculated with a finite but very large data set and even if we have the extraordinary amount of one million observations we would have among them only 20 past observations at which $x_j = \eta_j$! Moreover, conditional on observing at t that $x_t = \eta_t$, we need to wait another two million dates for the occurrence of one additional such observation. Thus, predictable subsequences that have no effect on the stationary measure m also have the property that they may be neglected for the purpose of forming a belief about the future.

Our second observation notes that the stationary measure is calculated, after all, with a large but finite number of observations. We have stressed in our discussion above (see Section 3) that we made the assumption that agents know the stationary measure for *analytical convenience* with the view to avoid turning the theory into a complex machinery of statistical approximations. We may also note that the usefulness of our theory depends upon the stationary measure being a reasonably tractable object. Hence, even under the assumption that we have a vast set of data and the stationary measure can be reasonably approximated, the theory is justified in concluding that rational agents may doubt that a pattern that was observed only 20 times in a million dates is *sure* to be a deterministic pattern which will be continued in the future. More generally, with finite data rational agents may be reasonable to doubt the certainty of deterministic patterns *on any remote subsequence which has no effect on the stationary measure*. This is compatible with the basic view of our theory which holds that rational agents may or may not place positive probabilities on sets which are assigned zero probability by the stationary measure.

Our third and final note is rather simple. We have ample evidence that some economic time series have deterministic patterns which occur along linear shifts such as seasonality and other cyclical patterns. These deterministic patterns would be treated as outlined above. However, in all other applications of our theory one should view non-stationarity as a model of *unpredictable* but important structural change. Relative to such dynamical processes, deterministic patterns on non-linear shifts which have no effect on the statistics of the economy are ignored since they have little interest to the usual applications of the theory.

References

Ash, R. B.: *Real analysis and probability*. New York, Academic Press, 1972

Aumann, R. J.: Agreeing to disagree. *Annals of Statistics*, **4**, 1236–1239 (1976)

Aumann, R. J.: Correlated equilibrium as an expression of bayesian rationality. *Econometrica*, **55**, 1–18 (1987)

Billingsley, P.: *Convergence of probability measures*. New York: Wiley, 1968

Blackwell, D., Dubins, L.: Merging of opinions with increasing information. *Annals of Mathematical Statistics*, **33**, 882–886 (1962)

Blackwell, D., Dubins, L.: On existence and non-existence of proper, regular, conditional distributions. *The Annals of Probability*, **3**, 741–752 (1972)

Choquet, G.: *Lectures on analysis, Volume II*. New York: W. A. Benjamin, 1969

Diaconis, P., Freedman, D.: On the consistency of bayes estimates. *The Annals of Statistics*, **14**, 1–26 (1986)

Dowker, Y. N.: Finite and sigma-finite invariant measures. *Annals of Mathematics*, **54**, 595–608 (1951)

Dowker, Y. N.: On measurable transformations in finite measure spaces. *Annals of Probability*, **8**, 962–973 (1955)

Feldman, M.: Bayesian learning and convergence to rational expectations. *Journal of Mathematical Economics*, **16**, 297–313 (1987)

Feldman, M.: On the generic nonconvergence of bayesian actions and beliefs. *Economic Theory*, **1**, 301–321 (1991)

Fontana, R. J., Gray, R. M., Kieffer, J. C.: Asymptotically mean stationary channels. *IEEE Transactions on Information Theory*, **27**, 308–316 (1981)

Gray, R. M., Kieffer, J. C.: Asymptotically mean stationary measures. *Annals of Probability*, **8**, 962–973 (1980)

Gray, R. M.: *Probability, random processes, and ergodic properties*. New York: Springer-Verlag, 1988

Harrison, M., Kreps, D.: Martingales and arbitrage in multiperiod security markets. *Journal of Economic Theory*, **20**, 381–408 (1979)

Harsanyi, J. C.: Games with incomplete information played by "Bayesian" players, Parts I, II, III. *Management Science*, **14**, 159–182 and 320–334 and 486–502 (1967–1968)

Kalai, E., Lehrer, E.: Rational learning leads to nash equilibrium. *Econometrica*, **61**, 1231–1240 (1993)

Katznelson, I., Weiss, B.: A simple proof of some ergodic theorems. *Israel Journal of Mathematics*, **42**, 291–296 (1982)

Kieffer, J. C., Rahe, M.: Markov channels are asymptotically mean stationary. *Siam Journal of Mathematical Analysis*, **12**, 293–305 (1981)

Kurz, M.: On the structure and diversity of rational beliefs. *Economic Theory*, **4**, 877–900 (1994a)

Kurz, M.: On rational belief equilibria. *Economic Theory*, **4**, 859–876 (1994b)

Neveu, J.: *Mathematical foundations of the calculus of probability*. San Francisco: Holden-Day, 1965

Nyarko, Y.: Bayesian learning without common priors and convergence to nash equilibria. Working paper. New York University, 1992

Nyarko, Y.: Bayesian learning leads to correlated equilibria in normal form games. *Economic Theory*, **4**, 821–841 (1993)

Parthasarathy, K. R.: *Probability measures on metric spaces*. New York, London: Academic Press, 1967

Phelps, R. R.: Integral representation for elements of convex sets. In: Bartle, R. G. (ed.) *Studies in Functional Analysis*. Vol. 21, pp. 115–157. MAA Studies in Mathematics. 1980

Rechard, O. W.: Invariant measures for many – one transformation. *Duke Journal of Mathematics*, **23**, 477–488 (1956)

Royden, H. L.: *Real analysis*. Third Edition. New York, London: Macmillan Publishing, 1988