Discussion Paper No. 1160

AGGREGATION AND THE LAW OF LARGE NUMBERS IN ECONOMIES WITH A CONTINUUM OF AGENTS

by

Nabil Al-Najjar*

Northwestern University

First draft: March 1996
Revised: June 1996

(*) Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL, 60208.

[Phone: 847-491-5426; Fax: 847-467-1220; e-mail: al-najjar@nwu.edu]

I would like to thank Alejandro Manelli and Rann Smorodinsky for many helpful conversations, and seminar participants at Northwestern for their comments. Any remaining errors and shortcomings are my own.

Abstract:

This paper develops a framework in which a model with a continuum of agents and with individual and aggregate risks can be viewed as an idealization of large finite economies. The paper identifies conditions under which a sequence of finite economies gives rise to a limiting continuum economy in which uncertainty has a simple structure. The state space is the product of *aggregate states* and *micro-states*: aggregate states represent economy-wide random aggregate fluctuations, while micro-states reflect individual shocks which fluctuate independently around aggregate states and have no further discernible structure. In the special case where shocks in the finite economies are exchangeable, the limiting economy satisfies a continuum-version of de Finetti's Theorem. The paper then uses this framework to derive implications for the interpretations of the Strong Law of Large Numbers and the Pettis Integral.

KEY WORDS: *Aggregate states, micro-states, exchangeability, de Finetti's Theorem, Strong Law of Large Numbers.*

1. INTRODUCTION

Models of large economies with individual and aggregate risks are extensively used. A typical setting is that of an economy with a large number of agents where individuals have random characteristics or attributes such as their levels of wealth; valuations for a product; private information about an aggregate variable; preference, technology or cost parameters; or random returns on investments.

From the perspective of economic modeling, economies with a continuum of agents represent a conceptual idealization that can be useful in discarding incidental features of finite economies and highlighting the forces that become predominant as the number of agents increases. The view I take in this paper is that, while the world economic modelers are interested in is always finite, finite models are not necessarily the best tool to understand that world.[1]

The paper develops this interpretation of continuum economies with risk and explores its implications for the issues of aggregation and the Law of Large Numbers. I take as a primitive a sequence of *finite-characteristic economies* $\{f^N\}$, which is meant to capture a modeler's representation of an economic environment with increasing number of agents or risk-types (think, for example, of the case of replicas of a fixed initial economy). Each economy f^N has a space of agents $T = [0,1]$ that is partitioned into N subpopulations such that agents in each subpopulation have identical risk-type (*i.e.*, have identical random characteristic). Finite-characteristic economies can be interpreted as models with large but finite number of agents (See Section 2.1). Conversely, an economy with a finite number of agents can be mapped into a finite-characteristic economy by appropriately choosing the weights of the characteristics. Working with finite-characteristic economies is more convenient because they eliminate the need to use different index sets as the number of agents increases.

Assume, for simplicity, that agents receive binary (0-1) shocks and let $f_t^N(\omega)$ denote the shock received by agent t in the Nth model economy when the state is ω. Since there are only finitely many characteristics, the *empirical frequency* of shocks on a subpopulation of agents A in state ω is simply the fraction of 1's in that subpopulation at state ω and can be expressed as the weighted average

$$\mu^N(A,\omega) = \frac{1}{\tau(A)} \int_A f_t^N(\omega)\, d\tau,$$

where τ denotes the Lebesgue measure on the space of agents [0,1]. We will say that $\{f^N\}$ is a *stable* *sequence* of finite-characteristic economies if, as N increases, the empirical frequencies $\mu^N(A, \omega)$ settle down to well-defined limiting values (which may depend on both A and ω).

The economic content of stability is that correlation patterns found in the finite economies eventually settle down to some limiting pattern. This requirement seems natural if one wishes to use an idealized limiting model to capture corresponding patterns in large finite economies. A more practical justification is that this stability requirement is satisfied in the most familiar and widely used economic applications where models with a continuum of agents with risk are used. For example, stability is satisfied if the component random variables in f^N are exchangeable (*i.e.*, have a symmetric joint distributions) and, in particular, if they are i.i.d. Stability is also satisfied if f^N is generated by replicating a fixed economy.[2]

The paper's main result is that any stable sequence of finite-characteristic economies gives rise to a limiting economy in which uncertainty has the following simple structure. The state space can be expressed as the product of *aggregate states* and *micro-states*. Aggregate states encode all aggregate information about correlation patterns in the limit economy, while micro-states fluctuate independently around the aggregate states and contain details about agents' realized shocks.

In the special case where the component random variables in the finite-characteristic economies are *exchangeable* (*i.e.*, symmetrically dependent), individual shocks in the continuum economy can be represented as i.i.d. fluctuations with an unknown aggregate parameter. This provides a version of de Finetti's Theorem, well-known in statistical decision theory, for economies with a continuum of agents.

The decomposition of uncertainty into aggregate and micro-states provides a clear distinction between what an outside observer can and cannot learn about the state of the economy from observing the realized shocks of an arbitrarily large but finite number of agents. Since all information about systematic correlation patterns in the agents' shocks is contained in aggregate states, this observer will be able to approximately correctly deduce the value of the aggregate state, but not the micro-state of the economy. Implications of this observation on the interpretation of the Law of Large Numbers are discussed next.

Implications for the Law of Large Numbers for Continuum Economies

Despite the extensive use of models of continuum economies with risk, serious doubts concerning their interpretation were raised. The doubts raised center on whether in a continuum i.i.d. economy f where agents shocks have mean α the following interpretation of the Strong Law of Large Numbers (LLN) is valid:

$$\int_{[0,1]} f_t(\omega)\, d\tau = \alpha, \quad \omega - \text{a.s.} \tag{$*$}$$

Judd (1985) and Feldman and Gilles (1985) pointed out that this interpretation of LLN is problematic. Basically, for a typical state of nature ω the realized shock $f_t(\omega)$ is not necessarily measurable as a function of t, so the integral in (*) is not well-defined.

As its stands, the non-measurability of realizations arising in (*) is difficult to interpret because it is not clear how the idealized model economy f relates to large finite economies. Using the framework of this paper, I provide an interpretation of the measurability of realizations in terms of the informational content of observed shocks in large finite economies. With this interpretation, I argue in Section 4 that (*) represents qualitatively a much stronger conclusion than the one obtained in the classical Strong Law for an infinite sequence of random variables and propose an appropriate weakening of it.

Related Papers

The present paper is related to the papers by Uhlig (1996) and Al-Najjar (1995) who used linear methods to bypass the non-measurability of sample realizations. Uhlig's (1996) starting point is a fixed continuum i.i.d. economy viewed as a function defined on the space of agents [0,1] and taking values in the linear space of random variables with the L_2 norm. He proposed to integrate risk by taking Riemann sums[3] and showed that for the i.i.d. case with mean α these sums converge to an essentially constant random variable that is α with probability 1, and that this limit coincides with an integral for vector-valued functions called the Pettis integral (Diestel and Uhl (1977)). While the Pettis integral presents a number of advantages, its interpretation as a law of large numbers remained unclear. In Section 5, I relate the framework of the present paper to Uhlig's (1996) work by proving an exact relationship between the Pettis integral of an idealized limiting economy and the limiting empirical frequencies of shocks for a stable sequence of finite-characteristic economies. This relationship holds for general continuum economies, not just ones with independent shocks.

In Al-Najjar (1995), I characterized a class of continuum economies (which includes i.i.d. economies as a special case) in terms of the approximation by finite collections of random variables and in terms of the decomposition of risk into idiosyncratic and aggregate components. I derived results with LLN-flavor, namely that an agent with an anonymous payoff function and who regards the continuum model as an idealization of large finite economies will ignore idiosyncratic risk in the limit.

The present paper provides a different perspective on the problem and a number of important improvements. There are two key differences with respect to earlier works.

First, analysis in the present paper is carried out at the level of states (rather than at the level of random variables, viewed as linear objects). This yields results with different flavor and implications. For example, the decomposition of uncertainty into aggregate and micro-states is quite different from the linear decomposition obtained in Al-Najjar (1995) in that the latter only ensures that the residuals are uncorrelated. The difference is critical in applications where the main focus is on the informational content of the shocks. The decomposition provided here guarantees that all aggregate information has been extracted, while the linear decomposition is consistent with some of this information being left in the residuals.

The second main difference is that, rather than starting with a fixed continuum economy, the present paper takes the perspective of an economic modeler studying a class of possible environments each with large but finite number of agents or characteristics. I then ask whether features of the finite economies that persist and eventually predominate as the number of agents becomes large can be captured by some limiting model. This approach provides a straightforward way of interpreting assumptions and conclusion derived for the limiting continuum model. For example, the measurability problem implicit in (*) discussed earlier can now be traced to features of finite economies. This often enhances understanding of what the limiting model is trying to capture.

2. THE MODEL

We start with a space of agents $(T, \mathcal{T}, \tau)$, where $T = [0,1]$, $\mathcal{T}$ is the collection of Borel sets on $[0,1]$ and τ is the Lebesgue measure. To simplify the exposition, I will focus on the case where risk is represented by random variables taking the values of either 0 or 1. This is the case that received most attention in the literature; extensions to the more general case raise no significant new technical or conceptual issues.

An *economy* is a function f which assigns to each agent t a 0-1 random variable f_t representing the random shock received by this agent.

2.1. Finite-Characteristic Economies: Definition and Interpretation

Of particular interest is the class of *finite-characteristic economies*. Any such economy consists of a partition of the space of agents $\pi = \{A_1, \ldots, A_N\}$ and corresponding set of random variables $\phi = \{\phi_1, \ldots, \phi_N\}$ such that $f_t = \phi_n$ for every $t \in A_n$.[4]

In a typical application, a modeler might be interested in an idealized limiting environment where each risk-type has a negligible weight (as, for example, in the case where agents receive i.i.d. shocks). We will think of this as corresponding to a sequence of finite-characteristic economies along which $\delta^N = \max_n \tau(A_n) \to 0$. Here, δ^N measures how diverse agents' characteristics can potentially be, so the requirement that $\delta^N \to 0$ says that there are no restrictions on the diversity of characteristics in the limiting economy. An example where this occurs is when the space of agents is divided into N subintervals of equal length with i.i.d. shocks (in which case $\delta = \frac{1}{N}$). Thus, for very large values of N, this finite-characteristic economy appears as a good candidate to be the sort of finite economy that the continuum i.i.d. economy is meant to capture.

Two interpretations of finite-characteristic economies are as follows. First, a finite-characteristic economy may be viewed as representing an economy with finitely many agents so that a fraction $\tau(A_n)$ has characteristic ϕ_n.[5] Similarly, an economy with finitely many agents can be mapped into a continuum economy by simply assigning agents to subintervals of equal length. This interpretation is particularly relevant in models where equilibria depend only on the distribution of characteristics and is otherwise independent of the number of agents (as is the case with competitive equilibria).

The second interpretation is to imagine a fixed continuum and the partition π as representing a categorization of agents according to the information available about them (*e.g.*, according to socio-economic criteria such as income level, number of children, ... etc). The random variable ϕ_n then represents the random characteristic thought of as being typical for the subgroup A_n. Refinements in the partition π can be interpreted as corresponding to finer categorizations of agents, *e.g.*, by including new criteria that were previously ignored.

2.2. Structure of the Probability Space

In this paper, the structure of the probability space will be of some importance. The theorems below construct limiting economies starting from a sequence of finite-characteristic economy $\{f^N\}$. Each f^N can be defined in terms of a probability space $(\Omega^N, \Sigma^N, P^N)$ with finitely many states. A natural choice for a probability space for the entire sequence is the product $\Omega = \Omega^1 \times \Omega^2 \times \cdots$ so a state $\omega \in \Omega$ is a sequence of functions giving the details of the shocks received by each agent at each stage. Uncertainty in the finite economy f^N is represented by the marginal of P on Ω^N, denoted P^N. Except in special cases (*e.g.*, when all economies are i.i.d.), we cannot take P to be the product of its marginals. The reason is that the distribution of shocks in economy N might contain useful information about the distribution of shocks in economies later in the sequence. This, in fact, is the basis of the aggregation result in Theorem 1 which basically ensures that all such information can be extracted.

While the special product structure of the probability space described above is not essential (and none of theorems depends on it) it can be quite useful to keep in mind in order to better appreciate the intuition. For example, this structure makes it clear that there is a potential cardinality problem in constructing a limiting continuum economy, namely that Ω might not contain enough states (*i.e.*, Ω might not be rich enough) to support non-trivial correlation patterns of shocks.[6] This problem makes it necessary to introduce the following technical device. Define an *expansion* of (Ω, Σ, P) to be a measurable space $(I, \mathcal{I})$ and a probability distribution P^* on the product $(\Omega \times I, \Sigma \times \mathcal{I})$ such that the marginal distribution of P^* on Ω is equal to P.

Expanding the state space is standard technique in many contexts (for example, it is widely used in Decision Theory). One way to interpret this procedure in the present context is that each original state ω is divided into *sub-states* or *micro-states* $\{(\omega, i) : i \in I\}$ so that the distribution on the original states remains unaffected. An expansion amounts to choosing a 'nice' probability space to work with, and so should be thought of as a useful technical device. In particular, any random variable $x(\omega)$ on the original state space can be mapped into the random variable on the expanded space $x'(\omega, i) = x(\omega)$ without affecting any of its relevant properties. Distributions and "a.e. convergence" statements on the original space can be carried out to the corresponding objects in the new space.

3. EMPIRICAL FREQUENCIES AND THEIR STABILITY

3.1. Stability

For a finite-characteristic economy f^N, define the *empirical frequency* over a subpopulation of agents $A \subset T$ with $\tau(A) > 0$ by:

$$\mu^N(A,\omega) = \frac{1}{\tau(A)} \int_A f_t^N(\omega)\, d\tau.$$

This is the fraction of agents in A receiving shock 1 when the state is ω. Note that for a finite-characteristic economy, $f_t^N(\omega)$ is a step function so the integral in the definition is just an appropriately weighted sum. Note also that the fraction $\mu^N(A,\omega)$ will typically be random, as the distribution of shocks fluctuates with the underlying state ω even when the underlying random variables are independent.

DEFINITION: $\{f^N\}_{N=1}^{\infty}$ *is a stable sequence of finite-characteristic economies if for almost every state ω and every subset of agents A, the sequence $\mu^N(A,\omega)$ converges.*

Stability is the central condition underlying the analysis of this paper. It says that even though the number of characteristics N continues to increase without bound, the correlation patterns of agents' shocks, as reflected in the empirical frequencies, eventually settle down to a well-defined limiting pattern. Alternatively, stability ensures that the limiting behavior of the sequence of economies can be captured (approximately) by looking at a finite economy f^N for a large enough N. This means that the correlation patterns that can be identified by looking at a finite economy f^N for large N will not be drastically altered when the information contained in the tail $\{f^{N'}\}_{N'>N}$ is taken into account.

To further clarify the informational natural of stability, restrict attention to the weaker concept that requires convergence of empirical frequencies on subintervals of agents.[7] One can think of stability as reflecting the perspective of an observer who can make a good assessment of the average shocks affecting agents over subintervals. Crucially, however, stability does not require this observer to keep track of shocks for any particular individual agent. Thus, stability makes sense in situations in which an observer has information that is coarse relative to the increasingly finer differences in the agents' random idiosyncratic characteristics.

3.2. Main Theorem

Let G be the set of all measurable functions $g : T \to [0, 1]$.[8] The conclusion of Theorem 1 below will justify the interpretation of elements of G as *aggregate states*. We will also be interested in G-valued random functions $g(\omega)$ which we will interpret as a random aggregate state.

THEOREM 1. *If $\{f^N\}$ is a stable sequence of finite-characteristic economies, then there is a random aggregate state $g(\omega) \in G$ and a limiting economy f on an expanded state space such that*

1- *With probability 1, conditional on knowing $g(\omega)$, the shocks $\{f_{t_1}, \ldots, f_{t_M}\}$ are independent for every subset of agents $\{t_1, \ldots, t_M\}$, and $E[f_t \mid g(\omega)] = g_t(\omega)$ for almost every agent t;*

2- *For every subpopulation of agents A of positive measure (i.e., $\tau(A) > 0$) we have*

$$\lim_{N\to\infty} \mu^N(A, \omega) = \frac{1}{\tau(A)} \int_A g_t(\omega) d\tau, \quad \omega - \text{a.s.}.$$

The Theorem constructs a limiting model economy f that mimics corresponding correlation patterns in large finite economies and in which uncertainty has a particularly simple structure. Condition (1) says that all correlations between the agents' shocks can be completely reduced to the dependence of their shocks on the random aggregate state $g_t(\omega)$. Thus, the random aggregate state contains all information about systematic patterns in the distribution of shocks in the sense that once the aggregate state is known, remaining fluctuations are independent and have no discernible pattern. Condition (2) says that the limiting economy encodes all relevant information about the limiting empirical frequencies and justifies viewing this economy as a faithful representation of the distribution of shocks for a large finite economy f^N.

The construction in Theorem 1 performs both a reduction as well as an expansion of the state space. Thus, on the one hand, some elements of the original state space Ω are merged into aggregate states. This effectively *reduces* the complexity of the original state space as states inducing similar limiting distributions of shocks are lumped together. On the other hand, to ensure that the limiting economy has randomness that mimics the one found in finite economies, Theorem 1 *expands* the state space by creating *micro-states* that encode all detailed (micro) information about the nature of shocks received by particular agents.

Stability has no force for individual agents, so the implication $\int_A Ef_t^N d\tau \to \int_A Ef_t\, d\tau$ in Condition (2) cannot be strengthened to $Ef_t^N \to Ef_t$ in general.[9] On the other hand, in many cases the sequence of finite economies considered has more structure than stability. For example, if the component random variables in the sequence $\{f^N\}$ are either i.i.d., exchangeable or arise from replicating a fixed given economy, then we also have:

Corollary *Suppose that, in addition to stability, $\{f_t^N\}$ converges in distribution for almost every agent t. Then the limiting economy f constructed in Theorem 1 has the property that $\lim_{N\to\infty} f_t^N = Ef_t$ for almost every t.*

3.3. Economies with Exchangeable Risks and de Finetti's Theorem

An important special case to which Theorem 1 applies is the class of exchangeable sequences:

DEFINITION: A finite characteristic economy $f = (\pi, \phi)$ has exchangeable risks (or *exchangeable*, for short) if the component random variables $(\phi_1, \ldots, \phi_N)$ are exchangeable (*i.e.*, they have symmetric joint distribution).

Our next theorem is in the spirit of the classical de Finetti's Theorem for infinite sequences of random variables (see Hewitt and Savage (1956)). Theorem 2 says that for an exchangeable sequences of economies the random aggregate state $g(\omega)$ is now a random parameter $\theta(\omega) \in [0, 1]$ so that agents' shocks fluctuate in an i.i.d. manner with mean $\theta(\omega)$:

THEOREM 2. *Suppose that $\{f^N\}$ is a stable sequence of exchangeable finite-characteristic economies such that, for each N, π^N is a partition of [0,1] by equal subintervals. Then there is a limiting economy f on an expanded probability space and a random parameter $\theta(\omega) \in \Theta = [0, 1]$ such that:*

1- *The conditional shocks $[f_t \mid \theta(\omega)]$ are i.i.d. with mean θ;*

2- *For every subpopulation of agents A of positive measure (i.e., $\tau(A) > 0$) the limiting empirical frequencies satisfy:*

$$\lim_{N \to \infty} \mu^N(A, \omega) = \theta(\omega).$$

An infinite sequence of random variables $\{x_1(\omega), x_2(\omega), \ldots\}$ is *exchangeable* if $\{x_1, \ldots, x_N\}$ have a symmetric joint distribution for every N. One of the most natural ways to generate an exchangeable sequence is to first choose the probability of heads θ at random from the parameter space $\Theta = [0, 1]$ according to some distribution ν, then generate the sequence $\{x_1, x_2 \ldots\}$ as an infinite independent tosses of a coin with probability of heads θ. Conditional on knowledge of the true probability of the coin θ, tosses are i.i.d. with mean θ and distribution P^θ on $\{x_1, x_2 \ldots\}$. However, without knowledge of the value of θ (but knowing that θ was generated according to ν) the distribution P of $\{x_1, x_2 \ldots\}$ is symmetric and the sequence is exchangeable.

De Finetti's Theorem roughly says that any exchangeable sequence is generated in this way: Any distribution P of an exchangeable sequence is generated by a random draw of a parameter θ according to some distribution ν on θ followed by an infinite independent tosses of a coin with probability θ. This implies that, with probability 1, the limiting empirical frequency converges to the realized value of the random parameter $\theta(\omega)$. Conclusion (2) in Theorem 2 above has the same content for the continuum model.

The conclusion of de Finetti's Theorem can be formulated more abstractly (Hewitt and Savage (1956)) as the statement that: There is a distribution ν on the set of parameters Θ such that P is a convex combination of the i.i.d. coin tosses P_θ with weights given by ν. This formulation has the advantage that it can be defined directly for continuum economies without reference to a specific approximating sequence. Call a continuum economy exchangeable if for every finite subset of agents $\{t_1, \ldots, t_M\}$, the random variables $\{f_{t_1}, \ldots, f_{t_M}\}$ have a symmetric joint distribution. The next result shows that the economy constructed in Theorem 2 is exchangeable and satisfies a continuum analog of the conclusion of de Finetti's Theorem:

THEOREM 3. *Let f be the economy constructed in Theorem 2 on the expanded probability space $(\Omega^*, \Sigma^*, P^*)$. Then there is a random parameter $\theta(\omega) \in \Theta = [0, 1]$ with distribution ν and a collection of probability measures $\{P_\theta\}$ on Ω such that:*

1- *P_θ is i.i.d. with mean θ for almost every θ;*

2- *P^* is a ν-convex combination of the P_θ's:*

$$P^*(S) = \int_\Theta P_\theta(S) d\nu, \qquad \textit{for every } S \in \Sigma.$$

Remarks:

1- Theorems 2 and 3 are not trivial consequences of de Finetti's Theorem. It is well-known that de Finetti's Theorem does not hold for a finite set of random variables. In our context, each finite-characteristic economy f^N has no more than finitely many random variables, and the joint distribution is allowed to change across economies, so we never have more than finitely many exchangeable random variables to use in the proofs. Rather, stability of the empirical frequencies enters in an essential way in the proof of these theorems.

2- The reader might find it useful to note that the collection of probability distributions $\{P_\theta\}$ form a *disintegration* of the original distribution of P with respect to the σ-algebra generated by the random parameter $\theta(\omega)$ (see Dellacherie and Meyer (1978) for the basic theory and Stinchcombe (1990) for an economic application).

3- An interesting feature of Theorem 3 is that uncertainty in the limiting economy has a "compound lottery" form where first a probability law θ is drawn, then the final shocks are generated i.i.d. across agents using this law. In fact, the construction of the general case (Theorem 1) also has this feature where overall uncertainty is disintegrated into first choosing a probability law at random (the function $g_t(\omega)$ in the general case), then using that law to generate shocks in an independent manner.

4- A learning framework can be readily adapted to our static context to address the question: What parts of the structure of the continuum economy can be learned from finite observations generated using some reasonable mechanism? (this will be relevant for remark 5 below) To illustrate this, consider (for simplicity) the economy constructed in Theorem 2 for the exchangeable case and imagine an outside observer who is able to inspect the realized shocks received by K randomly sampled agents $\{t_1, \ldots, t_K\}$ (see Al-Najjar (1995) for a more extensive discussion of such sampling model). What would this observer be able to learn about the state of the economy based on this limited information? Intuitively (and one can rigorously prove), for large enough K this observer will be able to approximately correctly learn the aggregate state of the economy. However, no amount of finite information of this sort will enable the observer to learn the true micro-state, regardless of how large K is.

5- The decomposition of uncertainty into aggregate states and micro-states is related to a recent work of Jackson, Kalai and Smorodinsky (1996, JKS) on learnability in a Bayesian learning context. JKS consider an infinite sequence of random variables with joint distribution P. They identify a condition on P that makes it possible to break it down into a set of simpler, more atomic measures P_{θ^*} parametrized using some abstract space of parameters Θ^*. Breaking down here means that P can be written as a convex combination of measures P_{θ^*} with the property that, roughly, no additional observations can (asymptotically) help to uncover any further systematic (predictable) structure to it.

It is interesting to note that some aspects of the frameworks of the present paper and that of JKS have the same flavor. Theorems 1 and 2 also break down the joint distribution of a continuum of random variables into a set of aggregate states that cannot be further broken down, in the sense that conditional on knowledge of their value, residuals are independent and have no discernible pattern. Taking finite random sampling described in Remark 4 above as our mechanism for generating information used as a basis for inference, we also conclude only aggregate states are learnable, while the information contained in the micro-states is not. It should also be noted that the sense is in which learning aggregate states in our contexts leaves no further learnable structure is stronger because we can ensure that the residuals left after extracting aggregate states are independent. In JKS the comparable notion of asymptotic mixing is weaker as it requires only asymptotic independence between past observations and realizations in the distant future.

4. INTERPRETATION OF THE LAW OF LARGE NUMBERS

This section discusses the problem of defining and establishing a Strong Law of Large Numbers for continuum economies, taking the point of view that these economies are theoretical idealizations whose purpose is capturing the limiting behavior in large finite economies. My focus will be on the conceptual issues, so the treatment will be somewhat informal.

Benchmark: Standard Strong Law of Large Numbers

Fix an infinite sequence of i.i.d. random variables $\{x_1, x_2, \ldots\}$ which, for simplicity, we assume to take the values 1 with probability α, and 0 with probability $1-\alpha$. For each state ω and initial N-segment of the sequence we have a sample realization $\{x_1(\omega), \ldots, x_N(\omega)\}$ of 0's and 1's detailing the distribution of shocks across agents.

The classical Strong Law of Large Numbers can be thought of as consisting of:

1- Converting sample realizations $\{x_1(\omega), \ldots, x_N(\omega)\}$ into a sequence of random empirical frequencies

$$S^N(\omega) = \frac{x_1(\omega) + \cdots + x_N(\omega)}{N}, \quad N = 1, 2, \ldots$$

Each empirical frequency $S^N(\omega)$ represents the percentage of agents in the initial N-segment economy whose realized shock was 1.

2- Showing that these random empirical frequencies settle down almost surely to a unique limiting value that is equal to the population average:

$$\lim_{N\to\infty} S^N(\omega) = \alpha, \quad \omega - \text{a.s.} \qquad (**)$$

The key observation is that the conclusion of the classical LLN (**) is formulated in terms of implications on the behavior of a highly aggregated summary statistic S^N that keeps track only of the information contained in the one-dimensional variable representing the *fraction* of agents whose realized shock is 1. This involves considerably less information than the sample realization $\{x_1(\omega), \ldots, x_N(\omega)\}$ of 0's and 1's, an object that keeps track of the 2^N-dimensional information vector containing details about the actual shock received by any particular agent.

This distinction is important for the arguments presented below, so I further clarify it with the following simple example. Consider an exchangeable sequence of random variables $\{x_1, x_2, \ldots\}$ whose distribution is generated as follows. There is first a lottery which picks one of the two parameter values $\alpha_1 = 0.25$ or $\alpha_2 = 0.75$ with equal probability. Once α_i is selected, the sequence is i.i.d. with mean α_i. Let Ω denote the set of all infinite sequences of 0's and 1's and $A = \{\alpha_1, \alpha_2\}$ denote the set of parameters, so the state space is $A \times \Omega$ with a joint distribution defined in the obvious way. An outside observer accumulates massive information in the form of the actual realizations of the first K random variables, where K is arbitrarily large. An implication of the Strong Law is that the empirical frequencies S^N will settle down to the true value of the underlying parameter as N goes to infinity.

The classical Strong Law implies that this outside observer will be able, with high probability, to correctly infer the unknown parameter value α_i, and consequently infer the *distribution* of future realizations $\{x_{N+1}(\omega), x_{N+2}(\omega), \ldots\}$. It is all too obvious, but worth emphasizing nevertheless, that the Strong Law does not imply that this observer can infer what the future realization is going to be. In the learnability language discussed earlier, LLN implies that an observer using finite (but possibly very large) sets of observations will learn the A component of the state space, but not the Ω component. Knowing the future realization amounts to knowing the value of the true state. Clearly, one should not expect the Strong Law to deliver such knowledge because there just isn't enough information in finite realizations to make such inference.

Measurability as an Informational Requirement

A common view is that the Strong Law of Large Numbers for continuum economies should be interpreted as the statement:

$$\int_{[0,1]} f_t(\omega)\, d\tau = \alpha, \quad \omega - \text{a.s.} \tag{*}$$

As we discussed earlier, the problem with (*) is that the realizations $f_t(\omega)$ will typically be non-measurable, so the integral in (*) is not well-defined.

Does the seemingly technical nature of this measurability problem have interesting *economic* content? I can provide such content based on the view of the limiting model economy f as an idealization of large finite economies in the way described in Theorem 1.

For concreteness, fix a stable sequence of finite-characteristic economies $\{f^N\}$ where the Nth economy f^N is obtained by partitioning the space of agents into N subintervals of equal length. For the moment, we will not make any additional assumptions on the distributions of shocks in the economies $\{f^N\}$. Let f be the limiting economy constructed in Theorem 1. For a fixed state of nature ω and each N, we have a realization $f_t^N(\omega)$ that is a step function defined on the interval [0,1] providing full details about which individuals received which shocks in the Nth economy in state ω.

I will argue that, in essence, the measurability problem in (*) centers on the following question: How much (and what type of) information does knowledge of the state of the finite economy f^N provide about the state of 'finer' economies $f^{N'}$, $N' > N$ later in the sequence? More concretely, suppose an observer knows everything there is to be known about the state of an N-replica economy f^N for an arbitrarily large N. Using this information, what will this observer be able to infer about the properties of finer replicas $f^{N'}$ for N' much greater than N?

Consider the following two (informal) statements. For a typical ω, knowing the actual realization of shocks $f_t^N(\omega)$ in the Nth economy provides approximately correct information about

(A) The *realized shocks* $f_t^{N'}(\omega)$ in all finer economies $N' > N$.

(B) The *empirical frequencies* $\mu^{N'}(A,\omega)$ in all finer economies $N' > N$.

The relationship between the measurability problem in (*) and these two statements is expressed in the following proposition:

Proposition: *Suppose that the sample realizations of f are measurable almost surely (hence (*) is well defined almost surely). Then for every $\epsilon > 0$, there is N such that, with probability 1, for every economy $f^{N'}$*

$$\tau\{t : f_t^N = f_t^{N'}\} > 1 - \epsilon.$$

The conclusion of the proposition is a formalization of statement (A) above. The Proposition says that a necessary condition for the measurability problem not to appear is that knowledge of the state of the economy f^N enables an observer to approximately accurately predict the state of finer economies, down to the details of which agent gets which shock. Statement (A) is considerably stronger than statement (B) because it requires keeping track of the actual shocks received at the individuals' level. Statement (B), on the other hand, restricts attentions to aggregated information as it only keeps track of shocks at a macroscopic level (*e.g.*, the level of non-degenerate subintervals of agents). Statement (A) is also very different in spirit from the conclusion derived in the context of the classical Strong Law of Large Numbers (**) because, as I argued earlier, the conclusions of the latter are in terms of the limiting behavior of ***highly aggregated information*** in the form of empirical frequencies S^N, rather than in terms of the detailed description of future realizations.

A Restatement of the Law of Large Numbers for Continuum Economies

The earlier discussion suggests that one defines the Strong Law of Large Numbers for continuum economies in terms of the limits of empirical frequencies: Take as a primitive a sequence of finite characteristic economies $\{f^N\}$ whose component random variables are i.i.d. with mean α, and proceed in two steps:

1- Convert realizations of the finite-characteristic economies f^N to random empirical frequencies

$$\mu^N(A,\omega) = \frac{1}{\tau(A)} \int_A f_t^N(\omega)\, d\tau.$$

2- The Strong Law is the statement that these empirical frequencies settle down almost surely to a unique limiting value that is equal to the population average:

$$\mu^N(A,\omega) \to \alpha, \qquad \omega - \text{a.s.} \tag{+}$$

Conclusion (+) is in the spirit of statement (B) above and appears to provide a natural continuum-analog of the classical Strong Law. This is in the sense that (+) ignores details about the way realized 0-1 shocks are distributed across agents and instead limits attention to the asymptotic behavior of an aggregate statistic of the sample realizations.

What role does the idealized limiting economy f constructed in Theorem 1 play in all this? Here the object of fundamental economic interest is the sequence of finite economies $\{f^N\}$; the role of f is just as an analytically convenient device that provides a compact and easy-to-manipulate summary of relevant information in the finite economies.

5. INTERPRETATION OF THE PETTIS INTEGRAL

Uhlig (1996) suggested that a LLN for continuum i.i.d. economies can be obtained using an integral known in Functional Analysis as the Pettis Integral. To define it, let L_2 denote the space of square integrable (equivalence classes of) random variables on Ω with the usual inner product $(f \mid f') = \int_\Omega f(\omega)f'(\omega)\,dP$ and corresponding norm.[10] An economy f is *weakly measurable* if $(\phi \mid f_t)$ is measurable in t for every $\phi \in L_2$. A weakly measurable economy is *Pettis integrable* if for every measurable subset of agents A there is a random variable $\int_A f_t\,d\tau$ such that

$$\left(\phi \;\middle|\; \int_A f_t\,d\tau\right) = \int_A (\phi \mid f_t)\,d\tau, \quad \text{for all } \phi \in L_2$$

where the integral to the right is the usual Lebesgue integral; see Diestel and Uhl (1977). The random variable $[\int_A f_t\,d\tau](\omega)$ is called the (indefinite) *Pettis integral of f*.

Uhlig (1996) showed that the Pettis integral for an i.i.d. economy with mean θ is the degenerate random variable which equals θ with probability 1. While this and other applications to economies with simple correlation structure suggest that the Pettis integral might provide intuitively plausible answers to questions of aggregation, the value of this integral as a basis for a LLN for continuum economies remained somewhat unclear. For one thing, while LLN is fundamentally a probabilistic statement, the Pettis integral is an analytic concept derived using the linear structure of L_2 and so does not appear to have an immediately obvious probabilistic interpretation.

The framework of this paper provides a probabilistic interpretation of the Pettis integral: When the continuum economy under study arises as the limit of a stable sequence of finite-characteristic economies, the Pettis integral is the limit of the empirical frequencies of shocks.

THEOREM 4. *Let f be the continuum economy constructed in Theorem 1 and $g(\omega)$ be the corresponding random aggregate state. Then for every subpopulation of agents A of positive measure (i.e., $\tau(A) > 0$), the Pettis integral satisfies:*

$$\frac{1}{\tau(A)}\left[\int_A f_t\,d\tau\right](\omega) \;=\; \lim_{N\to\infty} \mu^N(A,\omega) \;=\; \frac{1}{\tau(A)}\int_A g_t(\omega)d\tau.$$

Roughly, Theorem 4 says that the Pettis integral is the limit of the average of the agents' shocks along the sequence of finite-characteristic economies. In economies with independent shocks, the Pettis integral provides a Law of Large Numbers only in the sense of Section 4, namely that averages shocks become essentially constant in the limit. However, it would be incorrect in the present framework to claim that the Pettis integral delivers a Law of Large Numbers when this law is interpreted as in (*).

Finally, Theorem 4 makes it relatively easy to calculate the Pettis integral in many situations (and often help in guessing its value). For example, if the economy is exchangeable then the Pettis integral is $\theta(\omega)$, which is simply the random parameter representing the unknown probability of the coin.

APPENDIX

PROOFS

I begin with some notation that will be needed below. All related definitions and results used in the proofs can be found in Diestel and Uhl (1977). Let $L_2(\Omega)$ and $L_2(T)$ denote, respectively, the spaces of real-valued square integrable functions on Ω and T with the corresponding L_2 norms. Let G to be the subset of $L_2(T)$ of measurable functions taking values in $[0,1]$ endowed with the (subspace) weak topology on $L_2(T)$. This makes G into a separable metric space. We will use $\mathcal{G}$ to denote the induced Borel σ-algebra on G.

PROPOSITION A.1: *There is an $L_2(T)$-valued random function $g : \Omega \to G$ such that for every set of agents of positive measure, $\frac{1}{\tau(A)} \int_A g_t(\omega) dt = \mu(A, \omega)$.*

Proof: For each N, define a vector measure ν^N with domain $(T, \mathcal{T}, \tau)$ and taking values in the space L_2 of random variables on Ω by setting $\nu^N(A) = \int_A f_t^N(\omega) d\tau$. Stability implies that there is a set-function $\nu : \mathcal{T} \to L_2(\Omega)$ such that, for any given $A \in \mathcal{T}$, $\nu^N(A) \to \nu(A)$, ω – a.e., and hence in L_2 norm. By the Vitali-Hahn-Sacks Theorem, the set-function ν is a countably additive $L_2(\Omega)$-valued vector measure. It is also possible to show that ν has bounded variation. Since $L_2(\Omega)$ has the Radon-Nikodym property, there is a Bochner integrable function $g : T \to L_2(\Omega)$ such that for every A, $\nu(A) = \int_A g_t \, d\tau$ (where the integral is the Bochner integral). By the Pettis Measurability Theorem, g is essentially separably valued, so there must be a countable orthonormal set $\{\delta_k\} \subset L_2(\Omega)$ such that $\sum_{k=1}^{\infty} \beta_{kt} \delta_k(\omega)$ for almost every t, and where each $\beta_k : T \to I\!R$ is a measurable function. Thus, for each k and any state ω, the function $\beta_{tk}\delta(\omega)$ is measurable. Since g_t is (ω – a.e.) the countable sum of such functions, we conclude that $t \mapsto g_t(\omega)$ is a measurable function ω – a.e.[11]

We must now show that g, interpreted as a function $g : \Omega \to G$ is measurable when these spaces are endowed with the σ- algebras Σ and $\mathcal{G}$ respectively. Define $\mathcal{G}^*$ to be the collection of all sets of functions of the form $\{x \in G : \int_A x \, d\tau \ > \alpha\}$ where α ranges over the rational numbers and A ranges over all subintervals of $[0,1]$ with rational endpoints. Recall that $L_2(T)$ is the closed linear space spanned by all indicator functions of subintervals in T with rational endpoints. Thus, the collection of sets $\mathcal{G}^*$ defined above forms a (countable) subbase for the weak topology on $L_2(T)$, in the sense that any open set can be expressed as a countable union of finite intersections of sets in $\mathcal{G}^*$. This in turn implies that the Borel σ-algebra $\mathcal{G}$ is obtained by performing a succession of countable operations of unions and intersections starting with sets in $\mathcal{G}^*$. Thus, to prove that g is measurable, we only need to show that for any A and α as above, the set $D_1 = \{\omega : \int_A g_t(\omega)dt > \alpha\} = g^{-1}\{\int_A x\, d\tau \ > \alpha\}$ is in Σ.

Fix such A and α and define $B_k = \int_A \beta_{kt}$. Note that $\int_A g_t(\omega) = \int_A \sum_{k=1}^{\infty} \beta_{kt}\delta_k(\omega) = \sum_{k=1}^{\infty} B_k\delta_k(\omega)$. Since each δ_k is a random variable by construction, sets of the form $\{\omega : B_k\delta_k(\omega) > b_k\}$ are measurable for any rational number b_k. Let $n = 1, 2, \ldots$ be an enumeration of all sequences $\{b_k^n\}$ such that: (1) all of the b_k^n's are rational numbers; (2) no more than finitely many of them are different from zero; and (3) $\sum_k b_k^n > \alpha$. Then, the set $D_2 = \cup_{n=1}^{\infty}\{\omega : B_k\delta_k(\omega) > b_k^n \text{ for all } k\}$ is the countable union of countable intersections of measurable sets and is therefore itself measurable. It is now easy to see that $D_1 = D_2$, establishing the measurability of the former as required.

Q.E.D.

The reader might wonder about the possibility of replacing the proof of Proposition 1 above by a simpler argument based on the fact that stability guarantees that for each state we can get a limiting measure $\mu(\omega)$ on [0,1], and that the proof of Proposition 1 above ensures that $g(\omega)$ is just the derivative of μ_ω with respect to τ. This suggests that one can simply extract aggregate states directly from the sequence. The main complication is that we also want to ensure that aggregate states are Σ-measurable, *i.e.*, they contain no more information that was available in the primitive sequence of economies. This is the hardest part of the argument above. Thus, while it is conceivable that the more direct argument mentioned earlier might work, it is by no means as straightforward or simple as it might initially appear.

Proof of Theorem 1: I first define an appropriate expansion of the state space Ω. Let $(I, \mathcal{I})$ be the measure space where I is the set of all 0-1 functions defined on the space of agents and $\mathcal{I}$ the σ-algebra generated by all finite cylinder sets (Shiryayev (1984), p. 149). A cylinder set consists of a finite set $\sigma = \{t_1, \ldots, t_M\}$ called a *base*, and a function $x_\sigma : \sigma \to \{0, 1\}$. Such function defines the subset of I consisting of all function that agree with x_σ.

States in the expanded space $\Omega \times I$ will be denoted $\omega^* = (\omega, i)$. Define the continuum economy f

by $f_t(\omega, i) = i(t)$ (recall that i is a function $i : T \to \{0, 1\}$). I now define a probability distribution P^* on $\Omega \times I$ such that its marginal on Ω is P and such that the conclusion of the Theorem is satisfied.

Fix a state ω and a base of agents σ. Define a probability distribution P_ω^σ on all cylinders x_σ with base σ so that the random variables $f_{t_m}(\omega, i)$ are independently distributed and such that $Ef_{t_m}(\omega, i) = g_{t_m}(\omega)$. Clearly, the P_σ's satisfy the consistency condition in Kolmogorov's Extension Theorem (Shiryayev (1984), pp. 164-5), so there is a unique probability measure P_ω under which the agents' shocks f_t are independent with mean $g_t(\omega)$. (For future reference, note that, by construction, the support of P_ω is $\{(\omega, i) : i \in I\}$.)

We must now ensure that the various distributions P_ω can be threaded together into a measure on $\Omega \times I$. This is guaranteed (Shiryayev (1984), p. 247) provided the following condition is satisfied: For every measurable set $B \in \mathcal{I}$, the real-valued function $\omega \mapsto P_\omega(B)$ is measurable (*i.e.*, a random variable on Ω). To begin, consider a cylinder set with a singleton base, *e.g.*, $x_t = \{i : i(t) = 1\}$ for some fixed t. By construction, $P_\omega(x_t) = Ef_t(\omega, i) = g_t(\omega)$. From Proposition A.1, $g_t(\omega)$ is a Bochner integrable Radon-Nikodym density, and so it is a random variable, hence $P_\omega(B_t)$ is a measurable function. For general cylinder set x_σ, independence implies that $P_\omega(x_\sigma) = \prod_{t\in\sigma} Ef_t(\omega, i) = \prod_{t\in\sigma} g_t(\omega)$, which is measurable in ω (for fixed σ) by the same argument mentioned earlier. The extension to general sets in $\mathcal{I}$ is now immediate. Thus, there is a unique probability measure P^* on $\Sigma \times \mathcal{I}$ such that its marginal on Σ is P.

By construction, the distribution of random variables on any finite base is independent given $g(\omega)$. To prove condition (2) in the Theorem, note that $\frac{1}{\tau(A)} \int_A g_t(\omega)dt = \mu(A, \omega)$ and by stability $\mu^N(A, \omega) \to \mu(A, \omega)$ (both statements hold $\omega -$ a.s.).

Q.E.D.

Proof of the Corollary: From the construction of the limiting economy f, we have $\int_A f_t^N(\omega)d\tau \to \int_A f_t(\omega)d\tau$ for every subset of agents A of positive measure. Taking expectations and reversing the order of integration, we get that $\int_A Ef_t^N d\tau \to \int_A Ef_t(\omega)d\tau$. Since Ef_t^N is a step function for every N and since it converges pointwise by assumption, then the limit must be a measurable function. But then the convergence of the means implies that the limit must be Ef_t, $t-$ a.e..

Q.E.D.

Proof of Theorem 2: Apply Proposition 1 to get a random aggregate state $g : \Omega \to G$. The key step is to show that, with probability 1, $g(\omega)$ is constant $t -$ a.e. Given this, we would then be able to replace the function $g(\omega)$ by the parameter $\theta(\omega)$ representing its (a.e.) value at ω and the conclusions of Theorem 2 are then simple restatements of those of Theorem 1.

To prove the claim that $g(\omega)$ is constant $t-$a.e., it is enough to show that for any pair of disjoint (non-degenerate) subintervals A and B and every $\epsilon > 0$, the probability $P\{\,|\,\mu^N(A,\omega) - \mu^N(B,\omega)\,|\, > \epsilon\} \to 0$ as $n \to \infty$. To simplify notation (but this is clearly not important for the argument) assume that the two subintervals have the same length and define $C = A \cup B$. Let K denote the (random) number of 1's in f^N and reindex the sequence of economies to ensure that in f^N the sets A and B have precisely N component random variables. For large enough, N, the effect on empirical frequencies of random variables which fall only in part within A or B may be ignored.

The key implication of exchangeability we need in the proof is that it guarantees that states with precisely K 1's all have equal probability, thus reducing the problem of calculating probabilities to that of counting states. Let Z_K^N denote the set of draws in C at stage N with precisely K 1's. Let L denote the random variable representing the number of 1's that fall in A. Then, conditional on Z_K^N and using this implication of exchangeability, the distribution of L is hypergeometric with probability distribution

$$\frac{\binom{N}{L}\binom{N}{N-L}}{\binom{2N}{K}}$$

and variance that is essentially equal to

$$\frac{K}{4}\left(1-\frac{K}{2N}\right).$$

By Chebyshev's inequality, we have

$$P\left\{\left|\frac{L}{N}-\frac{K}{2}\right| > \epsilon\right\} = P\left\{\left|L - N\frac{K}{2}\right| > N\epsilon\right\}$$

$$< \frac{\frac{K}{4}\ (1 - \frac{K}{2N})}{N^2\epsilon^2}$$
$$< \frac{N}{N^2\epsilon^2} = \frac{1}{N\epsilon^2}.$$

The last term converges to 0 as $N \to \infty$ for any fixed ϵ, and note that this bound is uniform in K.

Q.E.D.

Proof of Theorem 3: That f is exchangeable follows from the fact that g_t is a.e. constant. The two conditions in the conclusion of the theorem follow directly from the construction in Theorem 2.

Q.E.D.

Proof of The Proposition: From the construction of Theorem 1, if realizations are measurable (with probability 1), then they must equal the aggregate states. Condition (2), on the other hand, implies that the sequence of integrals $\int_A f_t^N(\omega)d\tau$ converge to the integral $\int_A g_t(\omega)d\tau$ for every measurable subset of agents A. This is easily seen to imply that the sequence of realizations must converge in measure. Combining this with the fact that shocks take only the values 0 or 1 yields the conclusion of the proposition.

Q.E.D.

The next proof gives a good illustration of the role played by the product structure of the state space in the limiting economy:

Proof of Theorem 4: The second equality follows from Proposition A.1. Since the Pettis integral is uniquely defined (when it exists and up to sets of measure zero), we only need to show that for every set of agents A and every linear functional $\phi \in L_2(\Omega^*)$,

$$\int_A (\phi \mid f_t(\omega^*))\ d\tau = \left(\phi \;\middle|\; \left[\int_A g_t(\omega)\right]\right)\ d\tau.$$

Since the random variables g_t, $t \in A$ were obtained as the values of a Bochner integrable function (see the proof of Proposition A.1), the RHS is equal to $\int_A (\phi \mid g_t(\omega))$ so it suffices to show that $(\phi \mid f_t(\omega^*)) =$

$(\phi \mid g_t(\omega))$ for t – a.e. (for this statement to make sense. I interpret g_t trivially as a function on Ω^* by setting $g_t(\omega, i) = g_t(\omega)$, so ϕ is a well-defined functional relative to g_t).

Since L_2 is spanned by the indicator functions of measurable sets, we only need to prove the equality for ϕ that is an indicator function of a set of the form $C \times D$ where $C \subset \Omega$ and $D \subset I$ is a cylinder set with countable base. It can also be verified that we only need to examine the case where D is a cylinder set x_σ with finite base.

For a fixed t we have:

$$\begin{aligned}
(\phi \mid f_t(\omega^*)) &= \int_{C \times x_\sigma \subset \Omega^*} f_t \; d\omega^* \\
&= P^*(\{f_t = 1\} \cap x_\sigma \cap C) \\
&= \int_C P_\omega(\{f_t = 1\} \cap x_\sigma) \; d\omega && (a) \\
&= \int_C P_\omega(\{f_t = 1\}) \cdot P_\omega(x_\sigma) \; d\omega && (b) \\
&= \int_C g_t(\omega) \cdot P_\omega(x_\sigma) \; d\omega, && (c)
\end{aligned}$$

where (a) follows from the disintegration property of P^* implied by the construction while (b) and (c) follow from the fact that individual shocks under P_ω are independent with mean $g_t(\omega)$ by construction. On the other hand, and using similar steps as before, we have:

$$\begin{aligned}
(\phi \mid g_t) &= \int_{C \times x_\sigma} g_t \; d\omega^* \\
&= \int_C g_t(\omega) \cdot P_\omega(x_\sigma) \; d\omega.
\end{aligned}$$

Q.E.D.

ENDNOTES

1- This 'finitistic' view of continuum models is economically well-motivated and, in fact, quite commonly cited as a basic desiderata for a sound theoretical understanding of this problem (see, for example, Judd (1985) and Uhlig (1996)).

2- By way of example, suppose there are two agents, each with a shock that assumes one of two possible values (their types). Thus, resolutions of uncertainty can be represented by a random variable taking values in the four point set $\{0,1\}^2$ (reflecting, in particular, the possibility of correlation between the types of the two agents). The analysis of this paper, which is carried out here only for the 0-1 shock case can be easily generalized to the finite-valued case (see Al-Najjar (1995) p. 1209 and Section 6).

3- This process can be roughly described as partitioning the space of agents into finitely many subintervals; selecting an agent from each subinterval then averaging their random shocks. As the partition becomes finer, the Riemann sums converge in L_2 to a constant random variable which is the population mean α with probability 1.

4- Throughout the paper, the only subsets of agents considered will be assumed to be measurable (*i.e.*, belong to $\mathcal{T}$).

5- If the weights $\tau(A_n)$ are rational numbers, then we can always find M large enough so that this is true. If the weights are irrational, then they can be approximated arbitrarily closely by a large finite economy with rational weights.

6- To illustrate this problem, note that supporting a finite characteristic-economy f^N does not require the state space Ω^N to have more than 2^N states. This means that a sequence of economies can be supported by the countable product of finite sets $\Omega = \Omega^1 \times \Omega^2 \times \cdots$, and this has the cardinality of the continuum. On the other hand, supporting the limiting economy may require a state space of the cardinality of the power set of the continuum. To see this, consider the continuum i.i.d. economy and note that there has to be as many states as there are characteristic functions of subsets (not necessarily measurable) of the continuum [0,1].

7- This is no restriction as Borel measurable subsets of agents are in fact generated by subintervals.

8- Throughout, G will be given the topology of weak convergence and the measurable structure obtained by taking the Borel σ-algebra generated by that topology. Both are formally defined in the Appendix.

9- Consider, for example, a sequence in which the space of agents in economy N is divided into equal subintervals indexed $1, \ldots, 2N$. Intervals with odd index are populated by agents with mean shock 0.75, while agents in even subintervals have a mean shock 0.25. Shocks are otherwise independent. It is easy to see that this sequence is stable, so Theorem 1 applies yielding a limiting economy with a single aggregate state at which all agents receive i.i.d. shocks with mean .5. The limiting economy correctly reflects the limiting frequencies aggregated over measurable subpopulation, and these indeed converge to .5. The additional structure we find in each finite economy in the sequence (namely that there are two distinct subgroups whose characteristics differ) is lost because the two subgroups are too finely mixed in the limit to distinguish between them based on aggregate information. This is reflected in the fact that the sequence of functions $\{Ef_i^N\}$ does not converge to a measurable limit.

10- Abusing notation, we will use the same notation for random variables and their equivalence classes.

11- A technical issue of some importance is that we are often not explicit about the fact that the δ's, which are members of $L_2(\Omega)$, are equivalence classes of random variables. Thus, strictly speaking, in evaluating these functions we are selecting representatives d_k of the equivalence class δ_k. If $\{d_k\}$ and $\{d'_k\}$ are two such selections, then d_k and d'_k agree on a set of measure zero $C_k \subset \Omega$. Since the span of the δ's is separable, the two selections will agree outside $\cup_k A_k$, which is a set of measure zero. Here, separability is crucial in making sure that sample realizations are well defined.

REFERENCES

Al-Najjar, N. I. (1995): "Decomposition and Characterization of Risk with a Continuum of Random Variables," *Econometrica*, **63**, 1195-1224.

Dellacherie, C. and P.-A. Meyer (1978): *Probabilities and Potential*. Amsterdam: North-Holland.

Diestel, J.J. and Uhl, Jr. (1977): *Vector Measures*. Providence, R.I.: AMS (Mathematical Surveys no. 15).

Feldman, M. and C. Gilles (1985): "An Expository Note on Individual Risk without Aggregate Uncertainty," *Journal of Economic Theory*, **35**, 26-32.

Hewitt, E. and L. J. Savage (1956): "Symmetric Measures on Cartesian Products," *Transactions of the American Mathematical Society*, **80**, 470-501.

Jackson, M., E. Kalai and R. Smorodinsky (1996): "Patterns, Types, and Bayesian Learning," MEDS Department, Kellogg GSM, Northwestern University.

Judd, K. L. (1985): "The Law of Large Numbers with a Continuum of IID random Variables," *Journal of Economic Theory*, **35**, 19-25.

Shiryayev, A. N. (1984): *Probability*. New-York: Springer-Verlag.

Stinchcombe, M. (1993): "Bayesian Information Topologies," *Journal of Mathematical Economics*, **19**, 233-53.

Uhlig, H. (1996): "A Law of Large Numbers for Large Economies," *Economic Theory*, **8**, 41-50.