AN INTRODUCTION TO EXCHANGEABILITY AND ITS ECONOMIC APPLICATIONS*

by

John J. McCall

University of California, Los Angeles

and

The Hoover Institution

UCLA Working Paper No. 526

September 1988

We are socialized to the belief that there is one true model and that it can be discovered or imposed if only you will make the proper assumptions and impute validity to econometric results that are transparently lacking in power. [R.M. Solow, 1985]

---

*I am indebted to A. Alchian, M. Aoki, S. Berninghaus, R. Clower, B. Ellickson, A. Leijonhufvud, D. Levine, S. Lippman, B. McCall, P. Milgrom, A. Pakes, N. Wallace, and <u>especially</u> K. Velupillai for their encouragement and advice. The careful reading by M. Arnold is appreciated.

ABSTRACT

The symmetry possessed by exchangeability is the key to its usefulness. The methods that are used to comprehend reality will be successful insofar as they contain this symmetry. This is true of theoretical models, empirical methods, and algorithms. Indeed, if the design of mechanisms is by rearrangement of natural matter, their survival value will depend on the degree to which they retain the symmetry of the original matter. This comment encompasses complex machinery and sophisticated political, social, and economic institutions. This paper is simply an elaboration of this observation.

The process begins with designing experiments to decode nature's mysterious language. The success of these experiments hinges on their symmetry. The basic epistemological problem of economics is to build codes that will unscramble the mixed-up messages contained in the single realization of each economic process.

1. Introduction

This paper has several goals. The first is to acquaint economists with some of the recent developments in exchangeability analysis and exemplify several of its many applications in dynamic economics. The second is to demonstrate the unifying power of exchangeability. We claim that this unifying characteristic of exchangeability is strong enough to reverse the seemingly inexorable tendency to Balkinize economics into separate non-communicating subsets. In particular, a link is forged between economic theory and empirical economics. The apparently irreconcilable differences between the New Keynesian and New Classical approaches to macroeconomics are reconciled. The separation between macro and micro economics is narrowed. Historical and institutional economics are recognized as an integral part of evolutionary economic processes.

Exchangeability belongs to the potential probabilistic family whose members include Markov processes, martingales and diffusions with stopping times and conditioning the founding fathers. Thus, with a single economic arrow propelled by an evolutionary bow of symmetric and ancient design, we unify economics and see its connections with institutions, history, the other social and physical sciences and applied mathematics.

Another goal is to show the relation between the dilation ordering of information and the dilation operator. The former corresponds to the familiar "mean preserving spread" and is the natural measure of heterogeneity while the latter resides in function space and simplifies the estimation of complex economic processes. Both the ordering and the operator are exchangeable entities. This is the source of their similarity and endows each with the symmetry that is the basis for their analytical prowess.

It is important to emphasize at the outset that the approach advocated here is not novel. It is a continuation of ongoing economic research and merely changes a few labels and observes several connections that seem to have been overlooked. For example, the contract theory of the firm is classical; the application of stochastic processes to economic activities is the essence of rational expectations and finance; the significance of heterogeneity, misclassification and segmented labor markets is appreciated; and the formation of groups, clubs and coalitions has been studied from institutional and mathematical perspectives.

2. A Survey of Exchangeability[1]

Exchangeability[2] is a powerful and elegant concept that permeates and unifies economic processes. It links price theory and macroeconomics and provides a bridge between these two processes and econometrics. Its apparent conceptual simplicity belies its profundity.

Our task is obvious: to trace the radiations of exchangeability and alert the reader to their economic significance. We begin with a simple description of exchangeability.

Consider the following experiment. A coin is flipped 5 times. If a head (tail) occurs a 1(0) is recorded. What probability shall we assign to the $2^5 = 32$ possible outcomes? De Finetti answers this question by assuming that sequences of length 5 having the same number of 1's are equally likely. Hence, the sequence 11000 has the same probability as the sequence 00101. The precise position of the ones is irrelevant; only their sum enters the probability calculation. This type of symmetry was christened "exchangeability" by de Finetti.

---

[1] A more complete elementary survey is presented in McCall (1988a). The definitive survey is Aldous (1985), rigorous, comprehensive and inspiring.

[2] Exchangeable, symmetric dependence and interchangeable are synonyms. There is much to be said for "interchangeable". It emphasizes the homogeneous aspects of "exchangeability." One of its most important applications in economics and elsewhere is decomposing heterogeneous populations into homogeneous subpopulations. However, "exchangeable" is used most frequently. "Exchangeable" also can be interpreted according to its economic meaning. It is inextricably tied to heterogeneity and classification. The success of the physical sciences is explained partially by the connections discovered among seemingly diverse phenomena. The conversion of the heterogeneous into the homogeneous also has occurred in economics. The introduction of money converted the diverse into a common or exchangeable good with price the stopping rule. Thus trade not only smoothed out the differences among goods via prices, but expanded boundaries. See Hicks (1969) for a nice discussion.

Let us consider two basic exchangeable sequences of random variables: an <u>infinite exchangeable</u> sequence and a <u>finite n-exchangeable</u> sequence.

<u>Definition</u>: The finite sequence $(X_1,\ldots,X_n)$ is said to be <u>n-exchangeable</u> if

$$(X_1,\ldots,X_n) \overset{d}{=} (X_{\pi(1)},\ldots,X_{\pi(n)}),$$

for every permutation π of $\{1,\ldots,n\}$.

An infinite sequence $(X_1,X_2,X_3,\ldots)$ is said to be <u>exchangeable</u> if

$$(X_1,X_2,\ldots) \overset{d}{=} (X_{\pi(1)},X_{\pi(2)},\ldots)$$

where π is a finite permutation of $\{1,2,\ldots\}$.]

1) <u>The Method of Urns</u>

Urn models comprise simple generators of exchangeable sequences.

a. Let an urn contain n balls labelled $x_1,x_2,\ldots,x_n$. An infinite sequence of random samples of size one with replacement generate an infinite exchangeable sequence.

b. On the other hand, a sequence of n random samples of size one without replacement generates an n-exchangeable sequence.

<u>NOTE</u>: Recall that Feller has shown that every Markov chain can be represented as an urn model. All Markov chains are exchangeable.

<u>Definition</u>: An n-exchangeable sequence (X_i) is called <u>m-extendable</u> $(m>n)$ if $(X_1,\ldots,X_n) \overset{d}{=} (\tilde{X}_1,\ldots,\tilde{X}_n)$ for some m-exchangeable sequence $(\tilde{X}_i)$.

Consider an urn containing n balls each numbered, where the numbers are members of $\{1,2,3,\ldots,m\} \equiv S$. A random sample of size j is taken from this urn first with replacement and then without replacement. This produces two distributions on the set of j-triples. Diaconis and Freedman (1980a) show that the variation distance between these two distributions H_j and M_j is

$$\|H_j - M_j\| \le \frac{2c_j}{n},$$

where $c_j < \infty$ is the cardinality of S, H_j is the without replacement distribution (hypergeometric), M_j is the replacement distribution (multinomial), and

$$\|P-Q\| = 2 \sup_A |P(A)-Q(A)|,$$

is the variation distance.

2) The Bayes Method

Suppose there are J probability distributions: $\theta_1,\ldots,\theta_J$, defined on the real line R; suppose $p_j = P(\theta=\theta_j)$, $j = 1,2,\ldots,J$. Finally, let (X_i) be i.i.d. with distribution θ. Now, choose θ at random from $\{\theta_1,\ldots,\theta_J\}$. Then sample from the selected distribution θ giving the exchangeable sequence $X_1,X_2,X_3,\ldots$.

3) The Combinatorial Method

It is remarkable that many combinatorial (deterministic) calculations can be used to obtain theorems for i.i.d. sequences of random variables. Furthermore, most of these theorems remain true when "i.i.d." is replaced by "exchangeable.

Properties of Exchangeability

a) Identically Distributed

Suppose $n = 2$. Then exchangeability means that

$$F_{X,Y}(x,y) = F_{Y,X}(x,y)$$

for all x and $y \in R$. Letting $y \to \infty$ gives

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x,y) = \lim_{y\to\infty} F_{Y,X}(x,y) = F_Y(x).$$

Therefore, exchangeable random variables are identically distributed.[3]

b) Stationarity

Furthermore, it can be shown that for any j, $1 \le j \le n$, the distribution of the subset $(X_{i_1},\ldots,X_{i_j})$ of the set of exchangeable r.v.'s $(X_1,\ldots,X_n)$ is given by $P\{X_{i_1} \in B_1,\ldots,X_{i_j} \in B_j\} = P\{X_1 \in B_1,\ldots,X_j \in B_j\}$ for all $B_1,\ldots,B_j \in \mathcal{B}(R)$, the Borel subsets of the σ-algebra generated by the open subsets of R.

c) Correlation of Exchangeable Random Variables

Suppose $(X_1,\ldots,X_n)$ is an n-exchangeable sequence. The correlation coefficient $\rho(X_{i_j}, X_j)$, $i \neq j$ is given by:

(1) $$\rho \ge \frac{-1}{n-1}$$

with equality when sampling without replacement from an urn containing n balls, i.e., $\Sigma_{i=1}^{n} X_i$ = constant.

As $n \to \infty$, the RHS of (1) approaches 0. Thus in the limit

$$\rho \ge 0.$$

Aldous (1985) also shows that for an infinite exchangeable sequence ρ satisfies

$$0 \le \rho < 1.$$

---

[3]The significance of this observation is a consequence of the following fact: the ergodic hypothesis is true for i.i.d. sequences. Exchangeability is decomposable into two important concepts: conditional independence and identically distributed. See Loeve II (1978).

The practical and theoretical significance of exchangeability have their source in de Finetti's remarkable Representation Theorem:

Theorem 1: (De Finetti)

Let $P_n(r)$ be the probability of a sequence of r heads and $n-r$ tails, where the order is irrelevant, by exchangeability. Then

$$(2) \qquad P_n(r) = \int_0^1 \theta^r (1-\theta)^{n-r} P(\theta) d\theta$$

for some "prior" probability, $P(\theta) \geq 0$ and $\int_0^1 P(\theta) d\theta = 1$.[4]

Note that exchangeable sequences are mixtures[5] of Bernoulli sequences where the mixture is by a distribution over θ, the probability of a head. Recall that probability assessments are altered as new information unfolds. Let A be the event of interest and suppose our initial probability assignment is P(A). An event C occurs and we wish to modify P(A) to incorporate this information. This is accomplished by a mechanism that is fundamental to probability theory -- conditional probability. It is represented as $P(A|C)$ and defined by:

$$P(A|C) = \frac{P(AC)}{P(C)}, \qquad P(C) > 0$$

The conditional probability $P(\cdot|C)$ is a revision of the original or prior probability measure P(A) and contains all the information in C that

---

[4]More formally, let $\{X_i, i = 1,2,\ldots,n\}$ be a sequence of exchangeable Bernoulli r.v.'s. For each $n \ \exists$ a unique probability measure μ on $[0,1] \ni$

$$P\{X_1 = 1,\ldots,X_r = 1, X_{r+1} = 0,\ldots,X_n = 0\} = \\ = \int_0^1 \theta^r (1-\theta)^{n-r} d\mu(\theta).$$

[5]The relations among mixing (shuffling), stopping times and randomness are studied in Aldous and Diaconis (1986).

affects the occurrence of A. If this information is null, we say that A and C are independent events. This is represented by: $P(A|C) = P(A)$. Hence, $P(AC) = P(A)P(C)$ iff A and C are independent events.

Independence is the cornerstone of probability theory. Indeed, some have claimed that its absence would render probability theory indistinguishable from measure theory. Any method that "creates" independence enhances the significance of probability theory. It is remarkable that independence can be generated by using the methods of conditional probability.

Definition: Let A, B and C denote three events. The two events A and B are said to be conditionally independent given C iff

$$P(AB|C) = P(A|C)P(B|C).$$

Note that conditional independence is independence with respect to the conditional measure $P(\cdot|C)$.[6]

Remark 1: Clearly, a sequence of i.i.d. random variables is exchangeable. The de Finetti representation theorem shows that a general infinite exchangeable sequence is a mixture of these i.i.d. sequences. This observation diminishes the surprise associated with the asymptotic behavior of exchangeable sequences. For example, the limit laws that apply to i.i.d. sequences also apply to mixtures of i.i.d. sequences, that is, exchangeable sequences. Of course, the significance of these representation theorems for empirical phenomena relies entirely on the symmetry inherent in the concrete.

Remark 2: An alternative statement of de Finetti's theorem is: Suppose $\{X_n, n=1,2,\ldots\}$ is an infinite exchangeable sequence and $G = \cap_{n=1}^{\infty} G_n$ is

---

[6] An extensive discussion of conditional independence is contained in Pfeiffer (1979).

an exchangeable σ-algebra in that for each n, G_n is impervious to permutations of the first n subscripts. Then conditional on G, $\{X_n, n=1,2,\ldots\}$ is an i.i.d. sequence. This statement leads "immediately" to limit theorems. For example, by conditioning on G, applying standard results to the i.i.d. sequences, and then averaging over G, Buhlman (1960) proved the CLT for infinite exchangeable sequences.

On the other hand, the SLLN can be calculated using conditional expectations, rather than conditional independence. Simply calculate the partial sums, $S_n = \Sigma_{i=1}^n X_i$ and assume $E[|X_1|] < \infty$. It follows via a simple martingale argument that S_n/n converges a.s.

Indeed, Kendall (1967) and others reversed directions and showed that the SLLN implied de Finetti's theorem.

In his important paper, Eagleson (1982) eschews the conditional independence route and uses projection methods to derive his asymptotic distributions. This preference for conditional expectations is that the asymptotic results obtained in this manner apply to finite and partial exchangeable sequences, whereas the conditional independence theorems only apply to infinite exchangeable sequences.

Example 1: In econometrics many estimators have the form:

$$S_n = \sum_{i=1}^{n} a_{ni} X_{ni} \tag{3}$$

where $\{a_{ni}\}$ and $\{X_{ni}\}$ are sequences of constants and random variables, respectively. It is important to show that these estimators are weakly or strongly consistent, that is, converge to the appropriate parameter (weakly) in probability or (strongly) with probability one.

The asymptotic properties of estimators can be derived if S_n also converges in distribution.

It is almost always reasonable to assume that, for each n, the sequence $X_{n1}, X_{n2}, \ldots$ is <u>exchangeable</u>, whereas the same can not be said for <u>independence</u>.

For example, consider the sample mean $\bar{Z}_n$ of the random sample, $Z_1, \ldots, Z_n$. Clearly, $\bar{Z}_n$ can be put in the form of (3) by letting $a_{ni} = 1/n$ and $X_{ni} = Z_i$. Then

(4) $$S_n = \bar{Z}_n = \sum_{i=1}^{n} \left(\frac{1}{n}\right)\left(Z_i\right).$$

In similar fashion, the sample variance $s^2 = 1/(n-1)\ \Sigma_{i=1}^{n}\ (Z_i - \bar{Z}_n)^2$ can be transformed to (4) by letting $a_{ni} = 1/(n-1)$ and $X_{ni} = (Z_i - \bar{Z}_n)^2$, $1 \le i \le n$. The X_{ni} are <u>not</u> independent, but are exchangeable.

<u>Example 2</u>: <u>Regression and Conditional Independence</u>

A firm wishes to explain its annual income Y. The key explanatory variables are believed to be: $X_1, \ldots, X_T$. We wish to choose estimators of the $T+1$ parameters $\beta_o, \beta_1, \ldots, \beta_T$ such that the L^2 norm of $Y - (\beta_o + \beta_1 x_1 + \ldots + \beta_T x_T)$ is minimized. Geometrically, we project Y on $[X] \equiv [1, X_1, X_2, \ldots, X_T]$. The "goodness of fit" is measured by $E[Y-\hat{Y}]^2$, where $\hat{Y} = \hat{\beta}_o + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_T X_T$ is the solution to the minimization problem.

<u>Given</u> these "least squares" estimators, the difference between Y and $\hat{Y}$ is independent of $(X_1, \ldots, X_T)$.

<u>Example 3</u>: A major motive for studying exchangeability is to span the chasm separating modern economic theory from empirical economics. To achieve this goal a theory of inference for stochastic processes must be developed. The

symmetry inherent in exchangeability implies that many economic models are exchangeable processes and that conditional expectations are derived easily for these processes. This means that the well-known limit theorems for martingales can be used to develop an asymptotic theory of inference for exchangeable economic processes. These statistics are applicable to finite exchangeable processes, to partial exchangeable processes and, of course, to the infinite exchangeable processes assumed in de Finetti's original theorem.[7]

Example 4: Convergence to the Poisson Distribution (Eagleson)

Let $(A_{nj}: j = 1,2,\ldots)$ be an infinite sequence of exchangeable events for each n and consider the following sum of indicator functions:

$$T_{nk} = \sum_{j=1}^{k} I_{A_{nj}} .$$

The σ-algebra of events in $\sigma(A_{nj}, j \geq 1)$ is given by G_{nm} and is exchangeable w.r.t. the first m of the second index. Let $G_n = \cap_{m>0} G_{nm}$ be the exchangeable algebra of the n^{th} row and consider:

(5) $$F_{nj}^{(m)} = \sigma(A_{n1},\ldots,A_{nj}) \vee G_{nm}.$$

Notice that the first component of (5) guarantees that $A_{nj} \in F_{nj}^{(m)}$, while the second component lets one calculate conditional expectations, that is

$$P(A_{nj}|F_{n,j-1}^{(m)}) = P(A_{n,j+1}|F_{n,j-1}^{(m)}) = \ldots = P(A_{n,m}|F_{n,j-1}^{(m)}) =$$

$$= (T_{nm} - T_{n,j-1}) / (m-j+1).$$

[7]A complete discussion of the SLLN, CLT, and LIL for martingales is contained in Hall and Heyde (1980). An excellent presentation of functional limit theorems is contained in Greenwood and Shiryayev (1987).

If m is fixed, $F^{(m)}_{n,j-1} \subset F^{(m)}_{n,j}$, all n and j. If n and j are fixed, $\lim_{m\to\infty} F^{(2^m)}_{n,j} = \sigma(A_{n1},\ldots,A_{nj}) \vee G_n$. Now $(T_{nm}-T_{n,j-1})/m-j+1 \overset{a.s.}{\to} P(A_{n1}|G_n)$, so

$$\lim_{m\to\infty} P(A_{nj}|F^{(m)}_{n,j-1}) = P(A_{n1}|G_n).$$

If

$$nP(A_{n1}|G_n) \overset{p}{\to} \lambda, \quad \text{a constant}^{8}$$

then

$$T_{nn} = \sum_{j=1}^{n} I_{A_{nj}} \overset{\mathcal{V}}{\to} P(\lambda),$$

where $P(\lambda)$ is the Poisson distribution with parameter λ.

The Hausdorff Moment Theorem

There are few problems that have generated more practical and theoretical results than the moment problem. Landau observes its influence on functional analysis, operator theory, Fourier analysis, the prediction of stochastic processes, probability, statistics (and, of course, econometrics)), inverse problems and the design of algorithms. The moment problem is: when can a specified sequence of numbers represent successive moments of a nonnegative measure. There are four moment problems: the power, the trigonometric, the Hausdorff and the Stieltjes, corresponding to the measure being defined on the line, the unit circumference, an interval and the half-line, respectively.

---

[8] Kendall (1967) gives a sufficient condition for the hypotheses to hold. See Theorem 3 of Eagleson (1982).

Hausdorff showed that a sequence $m_0, m_1, \ldots$ were moments of a probability measure μ on $[0,1]$ if the following condition obtained:

$$m_k = \int y^k \, \mu(dy). \tag{6}$$

Condition (6) implies $m_{i+1} - m_i < 0$. Let Δ be the difference operator and apply (6) repeatedly to give

$$(-1)^r \, \Delta^r m_k = \int y^k (1-y)^r \, \mu(dx) \geq 0. \tag{7}$$

<u>Moment Theorem (Hausdorff)</u>

A sequence $m_0 = 1, m_1, m_2, \ldots$ has the moment representation (6) iff (7) is satisfied for all k and r. <u>The Hausdorff moment theorem is equivalent to deFinetti's exchangeability theorem</u>.[9]

<u>Partial Exchangeability</u>

When exchangeability is not present sometimes there is sufficient symmetry to induce invariance of a joint distribution under a <u>subgroup</u>[10] of permutations <u>or</u> another group of transformations of the sample space into itself. This yields the partial exchangeability introduced by de Finetti in 1938.

Examples of partial exchangeability are weak exchangeability, row × column exchangeability, and mixtures of Markov chains.[11]

---

[9] For a proof and edifying discussion of exchangeability and the moment problem see Diaconis (1987). Obviously (6) is a Choquet representation with (7) the iff positive definite (convexity) condition. Bernstein's theorem is also here. See Feller, vol. 2 (1971).

[10] Let g be a nonempty subset of a group G. Then g is a subgroup iff $a,b \in g \Rightarrow$ iff $a,b \in g \Rightarrow ab \in g$ and $a \in g \Rightarrow a^{-1} \in g$.

[11] The Markov chain analysis is in the important paper by Diaconis and Freedman (1980b). The most important economic application of partial exchangeability is the partitioning that accompanies insurance in particular and all firm behavior in general. Roughly speaking, this generalized notion of exchangeability means that each of the partitions created by the

Let $\{X_i, i = 1,\ldots,n\}$ be random points on the plane whose joint distribution is exchangeable and suppose $g(x,y)$ is a symmetric function. Then the 2-dimensional array of random variables

$$X_{ij} = g(X_i,X_j), \quad i,j = 1,\ldots,n$$

is invariant to permutations applied to each subscript simultaneously. Hence

$$\{X_{ij};\ i,j = 1,\ldots,n\} \stackrel{D}{=} \{X_{\pi(i),\pi(j)};\ i,j = 1,\ldots,n\}$$

These arrays were called weakly exchangeable by Silverman (1976).

Let α, $(\xi_i, i \geq 1)$ and $\gamma_{ij}, i,j \geq 1)$ be mutually independent uniformly distributed (on [0,1]) random variables. A standard weakly exchangeable array is given by

$$X^*_{ij} = f(\alpha,\xi_i,\xi_j,\gamma_{ij})$$

where f is a measurable function. Aldous (1985) shows that infinite weakly exchangeable arrays are distributed like a standard weakly exchangeable array. The Choquet representation theorem for row × column exchangeability is a deep result and has many practical applications including: analysis of variance (Consonni and Dawid 1985), the structure of vision (Diaconis and Freedman 1981), and the design of algorithms, complex systems, and policy. French (1988) observes that symmetry is a distinguishing feature of the production process. Yet "for all its importance, functional design has hardly been studied at all."

---

splitting action of conditional independence can be treated as exchangeable entities.

Extensions of De Finetti's Theorem

Definition: Let $X_1, X_2, \ldots$ be a sequence of random variables. The σ-algebra generated by $(X_1, X_2, \ldots X_n)$ is F_n.

Let F^n be the σ-algebra generated by $(X_n, X_{n+1}, \ldots)$. Then $F_\infty = \cap_{n=1}^{\infty} F^n$ is called the tail σ-algebra. The sets in F_∞ are known as tail events. An F_∞-measurable function is called a tail function. Clearly, a tail event is immune to fluctuations of any finite number of the X's. Two familiar tail functions are: $\lim\sup X_n$ and $\lim\inf X_n$.

A sequence of random variables is called spreading invariant if, for every integer k and every $n_1, < n_2 < \ldots < n_k$, the distribution of $(X_1, \ldots, X_k)$ is identical to the distribution of $(X_{n_1}, \ldots X_{n_k})$.

A stopping time T is a random variable satisfying,

$$\{T \le t\} \in \mathcal{F}_t, \quad \text{for all } t.(*)$$

Roughly speaking, $\mathcal{F}_t$ is the history of the process from 0 till t. The event $\{T \le t_0\}$ means that the process is stopped at or before t_0. Clearly, we want this event to belong to the process history and this is what (*) requires.

Theorem 2: The following statements are equivalent

(i) The sequence of random variables $(X_n) \equiv X_1, X_2, X_3, \ldots X_n \ldots$, is exchangeable, for all n.

(ii) The sequence (X_n) is spreading invariant.

(iii) The sequence (X_n) is i.i.d. given the σ-algebra F_n.

(iv) The sequence (X_n) is i.i.d. conditional on the tail σ-algebra F_∞.

(v) The sequence $(X_n) \stackrel{D}{=} (X_{T_1+1}, X_{T_2+1}, X_{T_3+1}, \ldots)$ for each increasing sequence $0 \le T_1 \le T_2 < T_3 < \ldots$ of stopping times.

(vi) The sequence $(X_n) \stackrel{D}{=} (X_{T+1}, X_{T+2}, \ldots)$ for each stopping time $T \ge 0$.

(vii) There is a random variable X^* such that the sequence (X_n) is i.i.d. conditional on X^*.

Remark 3: Kallenberg refers to property (vi) as strong stationarity. That is, not only is an exchangeable process stationary, but it also has the strong Markov property.

Remark 4: (vii) is based on Olshen's theorem (1973). Let $X_1, X_2, \ldots$ be an exchangeable sequence of random variables taking values in a complete, separable metric space. There is a real-valued r.v. X^* such that: (1) $X_1, X_2, \ldots$ are conditionally i.i.d. given X^*; (2) if W is another r.v. with property (1), then a version of X^* is measurable $\sigma(W)$ and (3) $\sigma(X^*)$ is identical to the tail σ-algebra, $\mathcal{F}_\infty$.

Spherical Symmetry

Let the random vector $Y^n = (Y_1, Y_2, \ldots, Y_n)$ be spherically symmetric ss if $UY^n = Y^n$, where U is an arbitrary $n \times n$ orthogonal matrix. An infinite sequence Y^∞ is ss if Y^n is ss for each n. An i.i.d. $N(0,v)$ sequence is ss and so is a mixture over v of i.i.d. $N(0,v)$ sequences. Now consider

Maxwell's Theorem[12]: An independent ss sequence is $N(0,\sigma^2)$ for some $\sigma^2 \geq 0$.

It should be clear that an ss sequence is exchangeable.

Aldous (1985) shows that the theorems of Maxwell and de Finetti imply the celebrated

Schoenberg Theorem: An infinite ss sequence $\{Y_n, n=1,2,\ldots\}$ is a mixture of i.i.d. $N(0,\sigma^2)$ sequences.

---

[12] A proof is in Feller (1971).

Spatial Exchangeability

Statistical mechanics has made considerable use of de Finetti's Theorem.[13] Much of this work is couched in terms of point processes and these have received attention by the queuing theorists (led by the East German school) and the Scandinavian theorists (Lauritzen, Kallenberg, Johansen, et al.) A substantial statistical methodology has been developed that appears to be a great promise for econometrics. For example, the spatial mobility that accompanies business cycles can be modeled in a manner compatible with the model of temporal fluctuations. These statistical methods can be adapted to estimate the joint spatial-temporal business cycles. They also are applicable to advertising, other forms of information flow and the contagion phenomenon. Under certain conditions there is a critical number, such that the contagion spreads (dies) if the contagion rate exceeds (lies below) this threshold.

Georgii (1979) has extended De Finetti's theorem replacing the homogeneous product measures with Gibbs states. Pitman (1978) does the same for an inhomogeneous product measure. A configuration of particles is described by w, a point in the space $\Omega = \{0,1\}^N$, $N = \{0,1,2,\ldots\}$. A particle (or person) lives at site n if $X_n(w) = w_n = 1$. Hence, $S_n = \Sigma_{i=1}^n X_i$ is the number of individuals living at the first n sites. The collection $\mathcal{P}$ of probabilities on $(\Omega,\mathcal{F})$, where $\mathcal{F}$ is the product σ-algebra, is composed of measures λ. That is, $\lambda \in \mathcal{P}$ assigns a positive probability to each finite dimensional configuration. Let $C(\lambda) \subset \mathcal{P}$ be the associated collection of canonical states: $\mu \in C(\lambda)$ iff for each $n,m \in N$, the conditional distribution of $(X_1,X_2,\ldots,X_n)$ given S_n and $(X_{n+1},\ldots,$

[13]Liggett (1985) and Kallenberg (1986) are excellent sources.

X_{n+m}) is the same under μ as under λ. Thus $\lambda \in C(\lambda)$ and each $\mu \in C(\lambda)$ has a unique integral representation over the set of extreme points $ExC(\lambda)$.

If $\lambda = \lambda p$ for $0 \leq p \leq 1$, where $\lambda p \in \mathcal{P}$ is the homogeneous product probability such that $\lambda p(X_n=0) = 1-p$, it has been shown that $C(\lambda p) \equiv C$, $0 < p < 1$, with C the set of exchangeable probabilities. Applying de Finetti gives

$$Ex(C) = \{\lambda p, 0 \leq p \leq 1\}.$$

<u>Theorem 3</u>: <u>Row-Column Exchangeability</u> (Aldous)

Let X be an exchangeable row-column array. Then there is an arbitrary measurable function f: $[0,1]^4 \to S$ such that $X \overset{D}{=} X*$, where $X*$ is represented by f, with i,j element

$$X^*_{ij} = f(\alpha, \xi^i, \eta^j, \lambda_{ij}).$$

The random variables, $\alpha, \xi^i, \eta^j, \lambda_{ij}$ are independent $U(0,1)$.

<u>Proof</u>: See Aldous (1985).

<u>De Finetti's Theorem for Markov Chains</u>

<u>Theorem 4</u>: Let $Y_0, Y_1, \ldots$ be an exchangeable stochastic process on (Ω, F, P), with (Ω, F) Polish and the Y_i taking values in a compact metric space S. Let $P_\omega(A)$ be a regular conditional probability on F given the tail α-algebra $F^{(\infty)}$ of the Y_i. Then the Y_i are i.i.d., a.s. P_ω.

<u>Proof</u>: See Diaconis and Freedman (1980b).

Zaman (1986) developed a finite version of this Markov theorem. He considered a finite, stationary, sequence $X_1, \ldots, X_n$ which is Markov exchangeable. By an ingenious combination of urn analysis and extreme point theory, Zaman shows that any portion of k consecutive elements, $k < n$, is an approximate mixture of Markov chains where the metric is the variation norm. Hence, Zaman follows roughly the same procedure as Diaconis and Freedman (1980b) in obtaining a finite version of de Finetti's theorem. Both proofs are constructive, enlightening and general enough to apply to all previous versions.

The Markov chain analysis began with de Finetti's conjecture that all infinite length Markov sequences are mixtures of Markov chains. Freedman showed this to be true for stationary sequences, Diaconis and Freedman (1980b) replaced stationary by recurrent, and Zaman (1986) replaced infinite by finite. Using entirely different methods Kallenberg proved that Markov processes were exchangeable. Kallenberg considers the continuous analogue of partial sums of exchangeable sequences to obtain his fundamental concept: processes with interchangeable increments. Aldous suggests that interchangeable increment processes be viewed as integrals $X_t = \int_0^t Z_s ds$ of some underlying exchangeable "general process" Z.

Kallenberg studies Levy processes -- those processes having stationary independent increments. Remember that for a Levy process X, X_1 possesses an infinitely divisible distribution. Recall that the random variable X_1 is infinitely divisible (i.d.) iff for each n, X_1 has the same distribution as the sum of n i.i.d. random variables. Hence, X_1 is i.d. if for each positive integer n it is the nth power of some characteristic function $c_n(t)$.

Infinitely divisible distributions are extremely important.[14] They were discovered in 1929 by de Finetti, who proved the following.

---

[14] Feller's volume II devotes considerable space to these distributions. Feller argues "simply that if certain phenomena can be described probabilistically, the theory will lead to infinitely divisible distributions".

<u>Theorem 5</u>: A characteristic function is infinitely divisible iff it has the form

$$c(t) = \lim_{j\to\infty} \exp\{p_j[h_j(t) - 1],$$

where the p_k's are positive real numbers and the h_j's are <u>characteristic functions</u>.

While de Finetti proved this theorem, it was Levy and Khinchine who developed the theory of infinite divisibility. For any i.i.d. distribution there is a Levy process such that X_1 has that distribution. The two most distinguished Levy processes are:

(8) $$X_t = at + bB_t,$$

where B_t is Brownian motion, a and b constants and

(9) $$Y_t = N_{\lambda t},$$

where N_t is a Poisson process with rate 1 and λ a constant.

The process X_t has continuous paths, whereas Y_t has jumps and is called a counting process. Bühlman (1960) proved that a process on R is exchangeable iff it is a mixture of Levy processes.

3. The Unifying Power of Exchangeability[15]

This section illustrates the versatility of exchangeability: (1) in its ability to comprehend seemingly contrary methodological positions of the two major schools of Economics: the "Rational" and the "Behavioral;"[16] (2) in its technical prowess to close the breach between economic theory and empirical economics, and finally (3) in its demonstration of the equivalence of superficially distinct solution methods it forces the economist to rely on economic criteria in determining the best solution technique. While our instincts tell us that the elegant, symmetrical solution also will be the most pragmatic, these intuitions must be checked at every stage of the problem-solving activity.

(1) The Arrow-Lucas Debate

It would be difficult to find more lucid and forceful defenses of the Behavioral and Rational economic positions than the definitive papers by Arrow (1986) and Lucas (1986).[17]

The approach to dynamic economics espoused by Lucas concentrates on steady-state analysis.

---

[15]In the recent literature on Bayesian econometrics (for example see Poirier (1988)) there is a tendency to select those aspects of "Bayesianism" that are divisive and controversial. The practical and unifying contributions of exchangeability to psychology, physics, biology and statistics are ignored. The equivalences among versions of de Finetti's theorem and the SLLN, Hausdorff's moment theorem, and semimartingales are facts independent of metaphysical profundities. They are the source of this unifying force.

[16]The labels "rational" and "behavioral" are used to divide economists in a simple but imprecise manner. The "rational" includes those who emphasize market efficiency, rational expectations, steady-states, and homogeneity, whereas the "behavioral" concentrates on fairness, disequilibrium, segmented markets and heterogeneity, the new classical macroeconomics, "bounded" rationality and market imperfections.

[17]The books by Arrow (1974) and Lucas (1987) as well as the papers by Simon (1984), Winter (1986), and Zechauser (1986) are recommended.

> Technically, I think of economics as studying decision rules that are steady states of adaptive process, decision rules that are found to work over a wide range of situations and hence are no longer revised appreciably as more experience accumulates.

On the other hand,

> experimental psychology has traditionally focused on the adaptive process by which decision rules are replaced by others. In this tradition, the influence of the subject's...preferences are kept simple by choosing outcomes that are easily ordered (rewards vs. punishments), and the focus is on the way that behavior is adapted over time toward securing better outcomes.

Lucas' description of the Walrasian auctioneer is especially pertinent to this paper:

> If the trial price equates demand and supply, it does prevail and trade is consummated. If it does not "clear the market", all bets are off and the auctioneer selects a new trial price. Under some quite reasonable conditions, this adaptive process (<u>though it is only the auctioneer who does any adapting</u>) converges to the market clearing price. [emphasis added]

This mechanism "has the virtues of being concrete, of relying on simple adaptive capacities, and of being, under a wide range of circumstances, stable".

In his trenchant critique of "rational" economics, Arrow (1986) claims that "it is possible to devise complete models of the economy on hypotheses other than rationality ... the use of rationality ... is ritualistic, not essential." Furthermore,

> The attainment of equilibrium requires a disequilibrium process. What does rational behavior mean in the presence of disequilibrium? Do individuals speculate on the equilibrating process? ... Since no one has market power, no one sets prices; yet they are set and changed. There are no good answers to these questions ...
>
> In particular, the homogeneity assumption seems to me to be especially dangerous. It denies the fundamental assumption of the economy, that it is built on gains from trade arising from individual differences ... it takes attention away from ... the distribution of income ...
>
> The main implication of this extensive examination of the use of the rationality concept in economic analysis is the extremely severe strain on information-gathering and computing abilities.

Recall Smale's (1976) critique of general equilibrium analysis:

> the theory has not successfully confronted the question, "How is equilibrium reached?" Dynamic considerations would seem necessary to resolve this problem. Another weakness is the reliance of the theory on long range optimization ... economic agents make one life-long decision, optimizing some value. With future dating of commodities, <u>time has almost an artificial role</u>. The model is reminiscent of John von Neumann's game theory ... the very best chess players don't analyze very many moves and <u>certainly don't make future commitments</u> .... <u>Between one economic decision and another there has been a real passage of time</u>, circumstances have changed, and the new decision takes place in this environment. [emphasis added]

Finally, "long-run optimization would be impractical ... because of barriers of complexity .... Dynamic models based on some kind of behavioral strategy could meet these obligations."

Returning to Arrow: "Walras' arguments can only be rescued by assuming a stationary state".

In his insightful commentary, Winter (1986) strongly objects

> to an alternative image of economic rationality that seems to have ever-increasing influence among theorists. This is the image of the economic actor as superoptimizer ... Subtle inferences from observations to underlying conditions are always correctly made, as in models of fully revealing rational expectations equilibrium or in recent models of reputation effects .... Above all, the superoptimizer has unlimited access to free information-processing capacity ....

This characterization is "totally inappropriate in a science concerned with the social implications of resource scarcity".

Zeckhauser sagaciously observes that both Arrow and Lucas have seized the high ground in this "turf battle" between the rational and behavioral schools and from these heights are able to select battlefields where victory is most likely.[18] However,

---

[18] This struggle exemplifies nicely the scope of exchangeability analysis. In the mosaic comprising the intellectual disciplines, conflict (interdisciplinary research) occurs near the boundaries. Thus the partitioning of knowledge into disciplines and schools is analogous to the

> From time to time there will be mutually agreed-on skirmishes. Major recent ones have centered on macroeconomics, where the evidence remains exceedingly controversial and inconclusive, and finance, where markets work exceedingly well but not perfectly -- an outcome sufficiently ambiguous to enable both sides to claim victory. In the future, I suspect, the behavioralists will continue mounting experiments or micro evidence of non-rational choice ... The rationalists will take succor from the overwhelming power of their model ... and the absence of any equivalently powerful competitor.

In terms of exchangeability the intellectual basis of this rational/behavioral conflict resides in the distinction between a <u>contractual</u> and an <u>efficiency</u> view of society. This divergence is embedded in a philosophy of society that has the contract school focusing on the welfare costs of change, especially labor mobility, externalities and asymmetries of broken promises. The efficiency school applauds resource mobility and individual initiative and responsibility. Guided by rational expectations decisionmakers prepare for adversity and seize opportunities as the future unfolds. Interference with this rational behavior creates more costs than benefits.

From an exchangeability perspective the key decisionmakers in an economic system are the firm and the individual. Each is a collection of stochastic contracts. The stochastic contract is governed by a birth and death process. This view of the economy is compatible with <u>both</u> the rational and behavioral approaches to economics. The stability of economic behavior resides in the institutions, customs and routines that govern civilized behavior.

---

partition of the globe into countries. Homogeneity reigns within borders. Borders possess the martingale property. Analysis of conflict can be studied in the benign environment of computer networks, where the competition is for access to transmission nodes. (I owe this network analogy to Bob Hall.) Furthermore, these conflicts can be diminished by <u>dilating</u> the nodes. See Padmanabahn and Netravali (1987).

In this subjective, de Finetti system there is no business cycle -- each individual, firm and industry has its own set of sample paths. Fluctuations are rainbow phenomena.[19]

Stability emerges when critical economic properties are found to be invariant to normal fluctuations. Heterogeneity and stability work together. For example, the insurance company partitions customers so that the SLLN (the "best" ergodic theorem) can be applied to each partition. Behavior is then predictable and the premiums charged allow the insurance company to survive. Formally, partial exchangeability applies to heterogeneous sets. The method is similar to the two stage process used by Bühlman, et al. in the actuarial literature. First, split the space into conditional independent subsets. Then apply the SLLN to each subset and calculate the "fair" premium. With de Finetti's partial exchangeable approach, all of this is accomplished in a single elegant stroke.

This general theory of partitions explains the formation of other economic institutions whose range encompasses <u>the street gang</u> and <u>the prestigious club</u>.[20] A vibrant healthy society is composed of segmented groups with permeable boundaries. When the partitions become rigid and constrain individual mobility the society <u>stultifies</u>. It is absolutely essential that all of these segmented groups be embedded in the larger societal network. The institutions that bind society's members must in some

---

[19]This approach to business cycles is consistent with the empirical results reported in Lucas (1987): <u>the cycle is insignificant compared to the trend</u>. This theory also is compatible with de Finetti and the remarkable book by Pigou (1926). For further discussion see Velupillai (1988) and McCall and Velupillai (forthcoming).

[20]These partitions or clubs correspond to those in the seminal work by Buchanan (1965). See Cornes and Sandler (1986) for a fine survey of the subsequent literature.

vague, but essential sense dominate the bounds that satisfy the instinctual urge to belong to a segmented group. This is achieved by conditional independence and the corresponding conditional expectation and conditional martingale. They split and then project from the larger space into the smaller spaces with conditional martingales preserving the "conservation of fairness" across partitions.

The equilibrium concept associated with the exchangeable, economic process is based on the "no arbitrage principle" -- a martingale. An economy abiding by this principle is automatically "efficient" and "fair". The martingale also is equivalent to the conservation of fairness.[21]

This equilibrium concept is the stochastic version of the Arrow-Debreu general equilibrium. However, the "Levy demon" replaced the Walrasian auctioneer, the *deus ex machina* corresponding to the "Maxwell demon".[22] The "Levy demon" is the natural extension of the Lucas-Sargent steady-state.[23] Both originate in function space and are efficient relative to the fluctuating phenomena. Now, however, endowments, tastes and technology are random processes and there are many sample paths. Optional stopping times enforce

---

[21] The fairness espoused by Baumol (1986) et al. is related to this conservation principle. See McCall and Velupillai (forthcoming).

[22] I am indebted to K. Velupillai for this and many other insights. Loeve (1978) has a nice description of these demons.

> Statistical physics has a familiar -- the "Maxwell demon" who travels along the individual paths of particles subject to deterministic laws of mechanics; his clock is the same along all paths. In sample Analysis there is now also a familiar -- the "Levy demon" who travels along individual sample paths of r.f.s (random functions) and his "random time" clock varies with the paths. In fact, the Maxwell demon is but a degenerate form of the Levy demon. [Loeve (1977)]

[23] The "Levy demon" is related to the analysis of Willinger, Taqqu, Aldous and Metevier, who study the fine structure of "general processes". Stopping times and information structures are fundamental to this analysis.

the "conservation of fairness" along each path. The exchangeable/martingale equilibrium analysis accommodates many stochastic processes including the recent research on dynamical economics.[24]

In 1926 L. Szilard exorcised Maxwell's demon by observing that a demon so "well-informed" (to track every molecule's sample path) required so much entropy that the corresponding perpetual motion machine was impossible.[25]

In their pathbreaking paper Anderson and Sonnenschein (1985) "feel strongly that positing the existence of a market-maker, who is able to infer the joint signal, is both unrealistic and antithetical to the spirit of decentralized competitive analysis". A demonic translation: they don't like Maxwell's demon, but this doesn't imply they would embrace Levy's demon -- they must be willing to substitute Meyer for Brouwer, that is, trade a fixed point for a floating crap game.

(2) <u>The Bridge From Economic Theory to Empirical Economics</u>

Most "moderate"[26] economists and many "extremists" agree that progress in economics is tied to sophisticated and steady empirical analysis. Perhaps the most important feature of the exchangeability approach is its ability to sustain such an endeavor. We will see that the dilation operator, an exchangeable entity, can be used to estimate dynamic economic models. It also has many other applications in designing and estimating economic theory in function spaces.

---

[24]For discussions see Brock (1986) and Day (1984).

[25]The economic implications of these demons are discussed extensively in McCall and Velupillai (forthcoming).

[26]Zeckhauser (1986) reflects this moderate position: "Empirical support should be our criterion ... we moderates should be happy ... knowing where rational models hold sway and where behavioral explanations" apply."

The spacings between order statistics are exchangeable. This observation has immediate implications for estimating tournament theory, auction theory, Stiglerian search theory and other economic theories that ask: whose on first (second, third, ...), how long have they been there, etc.

Finally, the U-statistic must be mentioned.

<u>The U-Statistic</u>[27]

The U-statistic for a sample $X_1, X_2, \ldots, X_n$, $n \geq m$ is given by

$$U_n^m = \binom{n}{m}^{-1} \sum_c h(X_{i_1}, \ldots, X_{i_m}),$$

where Σ_c is the sum over the m distinct combinations $(i_1, \ldots, i_m)$ from $\{1,2,\ldots,n\}$, h is a symmetric kernel. The sample is independently drawn from F, the objective being an unbiased estimator of θ.

Let $\theta = \theta(F)$ be the parameter function:

$$\theta(F) = E_F[h(X_1, \ldots, X_m)] = \int \ldots \int h(x_1, \ldots, x_m) dF(x_1), \ldots, dF(x_m).$$

The unbiased estimator of θ is obtained by averaging the symmetric kernel with respect to the observations. This is exactly how U_n^m is constructed <u>a la</u> de Finetti. The sequence of r.v.s $X_1, \ldots, X_n$ is exchangeable.

Let $\mathcal{F}_n$ be the σ-algebra generated by U_j^m, $j \geq n$. Then for $n+1 > i_m \geq \ldots \geq i_1 \geq 1$,

$$E[U_n^m | \mathcal{F}_{n+1}] = E[U_{n+1}^m | \mathcal{F}_{n+1}] = U_{n+1}^m.$$

Now let $U_n^* = U_{-n}^m$ and $\mathcal{F}_n^* = \mathcal{F}_{-n}$, $n \leq -m$ and $\{U_n^*, \mathcal{F}_n^*\}$, $n \leq -m$ is a

---

[27]The U-statistic was found by Halmos (observe the similarity to his dilation operator), developed by Hoeffding, Serfling (1980) and Denker et al. (1985). W. Brock introduced U to me. Its economic applications are considered more thoroughly in Brock et al. (1987).

martingale closed on the right.

Many familiar estimators are U-statistics. They include the sample mean and variance, Gini's mean difference, the empirical distribution, and the sample moments.

Example: Exponential Bounding

Three of the major methods for bounding the net return from a decision are based on moments, marginals, and the exponential distribution. Recall that if F has a constant hazard rate, it is the solution to:

$$1 - F(s+t) = (1-F(s))(1-F(t))$$

and

$$\theta(F) = \left(\int_0^\infty (1-F(t))dF(t)\right)^2$$

Let $\hat{h}(x,y,z) = 1_{(x-y-z>0)}$; then

$$\theta = \int \hat{h}(x,y,z)dF(x)dF(y)dF(z).$$

Finally let

$$h(x,y,z) = \frac{1}{3}\,[\hat{h}(x,y,z) + \hat{h}(y,x,z) + \hat{h}(z,x,y) + \hat{h}\ (z,x,y)].$$

Substitution yields an unbiased symmetric estimator for θ and the corresponding U-statistic can check for exponentiality.

Example: Cumulants, Mobius Theory and the U-Statistic

In a series of papers, Speed has probed the deep interplay between combinatorics and moment problems. The Möbius function of a partition lattice is ingeniously applied to moments, cumulants[28] and k-statistics.

---

[28] Let μ be a measure on R^n and suppose $MG(\mu)$ is the moment generating function of μ. the cumulants $\{K_r\}$ are defined by:

Several equivalences are derived and both the cumulant and the moment are generalized. Symmetry and exchangeability assume a prominent role in this analysis. The extension to analysis of variance and multiply-indexed arrays is developed and is related closely to the joint work with Lauritzen et al. on contingency tables. It is impossible to do this work justice here, but several points are too tantalizing to postpone.

(1) The generalized moments are not the same as those used in econometrics. Why not?

(2) Fisher's k-statistics are exchangeable!

(3) The Möbius function has been exploited in the spatial analysis (stereology) of Matheron, Serra, Kendall and Stoyan. These applications are in the form of dilations.[29]

(4) The spatial version of the de Finetti representation theorem has been developed by Pitman.

In his brief and insightful review of Speed's research, Diaconis displays great enthusiasm: (Application of the Möbius function yields a formula (2.9) linking moments and cumulents). "Formula (2.9) is brilliantly presented and developed by Speed who uses this combinatorial approach to prove all of the standard facts about cumulants in a unified way".

Finally, the flow of deterministic methods into and out of stochastic processes is displayed lucidly in Speed's analysis. The transitions from moments to polynomials, to reverse martingales to cumulants, etc. are indicative of the unifying methodology.

$$MG(\mu) = \Sigma_r K_r \theta_1^{r_1} \theta_2^{r_2} \ldots \theta_n^{r_n} / r_1' \ldots r_n'.$$

[29] Serra (1982) shows that a compact set is infinitely divisible w.r.t. dilation iff it is convex. This connects four of the unifying principles: dilation/exchangeability (Halmos/de Finetti), infinite divisibility (Feller), convexity (Strang, 1988) and compactness.

(3) Combinatorics, Algorithms and Network Theory[30]

One of the major sources of exchangeable processes is combinatorial analysis. Andersen, Baxter, de Finetti, and Feller were among the first to show that a deterministic combinatorial analysis at some point becomes probabilistic. This is a deep duality and has sweeping practical consequences. It is possible to choose many equivalent methods for solving and/or estimating stochastic economic processes. A stubborn combinatoric problem may be tractable when posed as a Markov process[31] or point process. Alternatively, a point process may simplify when characterized as a network flow and interpreted as an economic equilibrium. For example, in a study of labor mobility in a particular market the potential at a node corresponds to a price (wage), whereas the potential differences across arcs correspond to wage differences. In arc j there is a general relation Γ_j which roughly requires the flows of labor to obey the "conservation of fairness". In some applications it is more useful to treat Γ_j as demand and supply relations. This is a simple application of Rockefeller's monotropic optimization (1984) to the dual linear programming problem. The power of the equivalences among network theory, Brownian motion and harmonic analysis are revealed in function space and has not been exploited fully by economics.

---

[30] The interested reader may wish to consult Riordan (1978), Grotschel et al. (1988), Greene and Knuth (1986) and Rockefeller (1986).

[31] One of the first applications of L.P. to solve Markov decision processes was Manne (1965). Stochastic processes reciprocated when Khachian used (exchangeable) dilations to construct his celebrated algorithm. Sinden (1962) was earlier than Khachian. See Rustem and Velupillai (1987) for a discussion and nice application of Khachian's algorithm and the first economic application of complexity. Now hashing, sorting, and merging all have probabilistic equivalents. These are discussed in Ramakrishna (1987). The Möbius transformation was first used by Rota (1962) to characterize combinatorial methods.

These probabilistic processes founded on stopping times and conditioning are equivalent to harmonic analysis (potential theory).[32] All of the algorithmic and combinatoric methods used in economic processes[33] can be converted using probabilistic potential techniques.[34]

These methodologies imply that computers can be designed according to exchangeable principles.[35] Furthermore, the organizational structure of the firm, the design of its product and sales policies may benefit from symmetry transformations. The decisive observation is the absence of economics in many of the symmetry transformations. Most economists would agree that market survival is higher for processes possessing economic symmetry than for those that are only technically symmetric.

---

[32]The pathbreaking paper by Ressel (1985) exploits this equivalence to obtain "harmonic" representation theorems.

[33]An excellent survey of these methods is presented in the appropriate format by Arniel et al. (1988).

[34]Excellent introductions to this important literature are contained in Dynkin and Yushkevich (1969) and Doyle and Snell (1984). The remarkable paper by Samuelson (1952) "sees" the equivalences among electrical networks, general spatial equilibrium, linear programming and graph theory!

[35]Masani (1978) hints at this while Kailath (1987) is more explicit.

4. Dilations and Their Economic Consequences: A Brief Overview[36]

In the early 1950s Blackwell, Girschick, Shapely, LeCam, et al. began asking how one chooses among experiments on the basis of their "information content". The partial ordering of majorization emerged from this sustained analysis.[37] So de Finetti, Savage, Karlin and Wald emphasized the indispensible link between decision theory and any measure of the value of information. At about the same time, sufficiency was being incisively analyzed by Halmos and Savage, Bahadur et al. Perhaps the most influential paper on this subject was by Strassen (1965). It not only "solved" the stochastic ordering problem, but also demonstrated the significance of the marginal problem, lifting and much more.

A dilation is essentially a norm preserving stretch and it is equivalent to exchangeability. The essence of the dilation operator is to magnify a problem so that the basic complex structure is preserved but "simplified through enlargement". Roughly put, this clarifies the problem, enlightens the researcher, and improves the analysis. The applications of the dilation methodology are manifold, ranging from the estimation of complicated stochastic processes to assessing the degree of heterogeneity in an income distribution.[38]

---

[36] A more complete description of dilations, delation operators, and their economic applications is contained in McCall (1988). The connection between dilation ordering of stochastic processes and the dilation operator is the major contribution.

[37] A recent analysis is conducted by Cheng et al. (1987).

[38] The dilation pervades Fourier analysis and group theory. Gantmacher maintains that every linear transformation in R^n is obtainable by a series of rotations and dilations.

Exchangeability, Stochastic Ordering and Dilations

The influence of "mean preserving spreads" (MPS) on probabilistic economics has been decisive. Without ordering relations stochastic economics would have foundered. Generalizations of MPS's have been developed to compare dynamic stochastic models.

Majorization

Let x and y be vectors in R^n. The vector y majorizes x if

$$\sum_{i=1}^{k} x_{[i]} \le \sum_{i=1}^{k} y_{[i]}, \quad \text{for } k = 1,\ldots,n-1 \text{ and}$$

$$\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]},$$

where $x_{[1]} \ge x_{[2]} \ge \ldots \ge x_{[m]}$.

Thus x is less diffuse or spread out than y. This is symbolized by

$$x < y.$$

Theorem 6: The following statements are equivalent

(10) $x < y$

(11) $x' = Py'$ for some $P \in \mathcal{P}^n$,

the set of all $n \times n$ doubly stochastic matrices.[39]

---

[39] P sometimes is called the Schur transformation. Let M be the set of double stochastic $n \times n$ matrices. M is convex with the permutation matrices its extreme points. (A permutation $n \times n$ matrix has a single nonzero entry in each row and each column). The convex hull of the n! permutation matrices is identical to M. See Marshall and Olkin (1979) and Appendix 2.

(12) $x \in \text{conv}\{y_\pi \mid \pi \text{ a permutation of } \{1,2,\ldots,n\}\}$[40]

(13) $$\sum_{i=1}^{n} f(x_i) \le \sum_{i=1}^{n} f(y_i) \quad \text{for every convex function } f\colon R \to R.$$

Proof: See Marshall and Olkin (1979).

Schur Functions

Let $A \subseteq R^n$. The real valued functions that preserve the order of majorization are called Schur-convex: $f\colon A \to R$ is Schur-convex if x and $y \in A$ and $x \le y \Rightarrow f(x) \le f(y)$. A function f is termed Schur-concave if it reverses the order.

Let μ_1 and μ_2 be two measures defined on a set s in R^n. Suppose

(14) $$\int_s f d\mu_1 \le \int_s f d\mu_2$$

for all bounded functions f contained in the convex cone $\mathcal{F}$. When $\mathcal{F}$ contains all convex functions, μ_2 is referred to as a dilation of μ_1.[41]

Clearly, a dilation in R^n corresponds to majorization and MPS.

The following urn analysis relates dilations to exchangeability.[42] Let S_n and S_n^* be the partial sums of n random draws from an urn without and with replacement, respectively. Then it can be shown, that roughly,

$$S_n^* > S_n \quad \text{for each } n \ge 1.$$

[40] conv ≡ convex hull.

[41] It is also called Choquet order. See Kingman, (1978) for a martingale analysis and Phelps (1966) for a Choquet analysis.

[42] Of course, exchangeability has already entered as an integral part of a one-dimensional dilation. The dilation is a conditional expectation, and a contraction mapping. Every contraction has a unitary dilation. This connects de Finetti and Halmos!

The rationale for this is clear: the information derived from an n sample without replacement exceeds that from an n sample with replacement.

Observe that ergodicity (invariance to permutations) is the defining property of exchangeability, whereas entropy is used to compare exchangeable processes.

Some of the most important contributions of exchangeability to economics reside in the following analysis: identify those elementary probability measures and express all others as convex combinations of these. Often this set of probability measures lives on a Choquet simplex, with the extreme points representing the elementary measures. This is the approach taken by de Finetti using symmetry arguments and also by Dynkin (1978) using sufficiency arguments. The most elegant approach using this methodology is Lauritzen (1982). This deep analysis reveals the relations among exchangeability, martingales, sufficiency, Markov processes, the interplay between stochastic and deterministic (e.g., linear programming) methods, and much more.[43]

The Dilation Operator

The functional-theoretic path to the dilation operator began with Sz. Nagy and Halmos. Let the Hilbert space H be a subspace of the Hilbert space $\tilde{H}$. A bounded operator $\tilde{T}$ on $\tilde{H}$ is said to be a <u>dilation</u> of an operator t and H if the following projection relation obtains:

$$T^n = \pi_H T'^n | H \;\forall n \geq 1,$$

where π_H is the orthogonal projection in H' with H as the range

---

[43] A more complete discussion is contained in the Appendix.

space.[44]

Sarason (1987) claims that Halmos introduced the dilation operator to learn more about its then "mysterious structure," and that this intention has been "realized repeatedly". All isometries and coisometries possess unitary dilations. In addition, Sz-Nagy showed that every contraction has a unique dilation up to an isomorphism provided it is minimal -- not the direct sum of two unitary dilations. Sarason reminds us that the dilation operator can solve a number of classical moment and interpolation problems by its "lifting" ability. For example, let S_b be a shift operator relative to the inner function $b = e^{i(N+1)\theta}$; then

$$S_b e_n = \begin{cases} e_{n+1}, & n < N \\ 0, & n = N. \end{cases}$$

The exponentials form an orthonormal basis and the matrix associated with S_b is Toeplitz. In order that the Caratheodory-Fejer interpolation problem is solved, operators must be devised that commute with S_b. By its lifting ability, the dilation operator is able to commute with S_b.[45]

In recent years there has been an enormous theoretical literature on various features of the dilation operator. The trend of this literature has been toward simplification. For example, Masani develops an argument that portrays the dilation operator as a "simple" generalization of Kolmogorov's theorem on positive definite kernels. The dilation operator has been used to solve moment problems, the realization (identification) problem, and prediction (extrapolation and interpolation problems). An important example

---

[44]Halmos' original definition set $n = 1$. The dilation described in the text usually is called the power dilation.

[45]The details are important and definitely not trivial. See Sarason (1987) and references therein.

is the following: Let X_t, $t \in R'$ be a stochastic process with values in a specified Hilbert space H. The process is harmonizable provided

$$X_t = \int_R e^{-itu} \, d\Phi(u),$$

where Φ is a countably-additive Hilbert valued measure on R of bounded semi-variation. The class of <u>harmonizable processes</u> includes projections of <u>stationary processes</u>. Miamee and Salehi (1978) show that these classes are equivalent -- <u>a stochastic process is harmonizable iff it possesses a stationary dilation</u>. That is, <u>the harmonizable process is a projection of a stationary process taking values in some larger Hilbert space</u>. There is an immediate connection between the dilation operator in this setting and <u>ergodicity</u>.

Suppose X_t is a Hilbert-valued harmonizable process. Then

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T X_t \, e^{iu_0 t} \, dt = \Phi(u_0).$$

<u>Proof</u>: See Miamee and Salehi (1978).

There is a string of connections that ties dilation theory to almost every stochastic economic process. These include: identification, ergodicity, entropy, the economics of information, prediction, interpolation, approximation (splines, moments, and marginal problems), classification, projections and the corresponding estimation processes in both cross-section and time-series studies in econometrics. There is also a close connection between dilation operators and dynamical systems.[46]

---

[46] These are explored extensively in McCall and Velupillai (forthcoming).

5. Exchangeability, Semimartingales and the Mean Ergodic Theorem

In a recent article Kallenberg (1988) shows that exchangeability and semimartingales are equivalent.[47] Furthermore, he demonstrates that exchangeability is implied by the mean ergodic theorem. The first equivalence means that exchangeability can use the powerful analysis of the Strausberg School. It also verifies a long held conjecture: anyone who uses semimartingales must be a Bayesian.

Kallenberg notes that his theorems depend crucially on the stationarity assumption. We will see shortly that the dilation operator can transform some non-stationary processes into stationary processes. Kallenberg's theorems hold for finite and infinite sequences as well as for continuous processes. These theorems generate many counter intuitive examples.

When assessing the proof that the ergodic theorem implies exchangeability, recall that de Finetti proved that his exchangeability theorem implied the strong law of large numbers and Kendall (1967) demonstrated that the SLLN implied de Finetti's theorem. Also, in his recent survey article Dacunha-Castelle (1982) proves the following theorem.

First, he defines splitting in the sense of McKean. That is, suppose B is a subσ-algebra of A and $X_1,\ldots,X_n,\ldots$ are random variables on (Ω,A,P). Then $X_1,X_2,\ldots$ are conditionally independent w.r.t. B iff $\forall k$ the bounded measurable functions $f_1,f_2,\ldots,f_k$ are split by B, i.e.,

$$E[(f_1(X_1),\ldots,f_k(X_k))|B] = E[f_1(X_1)|B] \ldots E[f_k(X_k)|B].$$

Theorem 7: The following statements are equivalent:

(1) The random variables $X_1,X_2,\ldots$ are exchangeable above B.

---

[47] See Appendix I for an introduction to semimartingales.

(2) The random variables $X_1, X_2, \ldots$ are <u>spreading invariant</u> above B, that is, the distribution of $(Y_1,\ldots,Y_m,X_1,\ldots,X_n)$ is the same as the distribution of $(Y_1,\ldots,Y_m,X_{n_1},\ldots,X_{n_n})$, where $n_1 < \ldots < n_h$ $\forall h$ and $Y_1,\ldots,Y_m$ are B-measurable.

Exchangeability is roughly the same as a stretching or dilation transformation.

<u>The Mean Ergodic Theorem</u>:

Let J be a contraction on a Hilbert space H. Then the sequence of operators, $\{1/n \, \Sigma_{i=0}^{n-1} J^i\}$ on H is strongly convergent.

<u>Proof</u>: (Sz Nagy).

Apply the <u>dilation</u> operator.[48]

Hence, while the mean ergodic theorem implies the de Finetti theorem, an elegant proof of the former exploits an "exchangeability" operator.

Kallenberg (1988) observes that all of his results depend on the stationarity assumption. Indeed, one of his most important theorems shows that exchangeable sequences are strongly stationary in the sense that the sequence is invariant to random time shifts.[49]

Recently, it has been demonstrated that application of the dilation operator transforms a class of special nonstationary sequences to a stationary sequence. Thus there may be additional practical implications of Kallenberg's important paper, modulo dilation theory.

---

[48] See Halmos (1982, p. 327).

[49] Observe that this is an ergodic theorem and it is closely related to the dilation operator via the shift transformation.

APPENDIX 1

Exchangeability and Martingale Theory

The sudden and explosive growth of exchangeability analysis was possible because of its intimate relation with martingale theory. The latter has had a sustained period of development under the leadership of Doob, Meyer and the Strausberg School. While extremely abstract, martingale theory has been applied in almost all of the disciplines, especially physics, biology and financial economics.

Martingales are created rather easily. Let $\{X_n, n \geq 0\}$ be a sequence of random variables with $E[X] < \infty$. The function

$$M(X) = \left\{ \sum_{i=1}^{n} (X_i - E[X_i | X_{i-1}, \ldots, X_1]) \right\}$$

is a martingale relative to the sequence of σ-algebras generated by X_i, $i \leq n$.

Definitions: Let $\{X_i, i \geq 0\}$ be a sequence of random variables that are conditionally independent (c.i.) given the sigma algebra G. Suppose $E[X_i | G] \uparrow a$ in i a.s., $i \geq 1$. If $S_n = \Sigma X_i$ and $F_n = \sigma(G \cup \sigma(X_1, \ldots, X_n))$, then $\{S_n, F_n, n \geq 1\}$ is a submartingale.

If c.i. is replaced by exchangeable and $E(X_i, X_j) = 0$ all i, j, then $\{S_n, F_n, n \geq 1\}$ is a martingale.

If $E[X_{n+1} | F_n] = 0$, the sequence $\{X_n, F_n, n \geq 1\}$ is said to be a martingale difference.

It is now routine to go from the submartingale to the principal martingale convergence theorem: The law of large numbers.

Let $S_1, S_2, \ldots$ be a submartingale and let K denote $\sup_n E[|S_n|]$.

If $K < \infty$, $S_n \to S$ w.p.1, where S is a random variable such that $E[|S|] \leq K$.

The central limit theorem and the law of iterated logarithm are also easily obtained. Finally, estimators can be calculated (most of these are martingales), parameters of interest in the exchangeable process can be estimated, based on the observed temporal behavior.[50]

Semimartingales[51]

A stochastic process $X = (X_t, F_t)$, $t \in T$ is a semimartingale if it is decomposible into a local martingale X_1 and a process X_2 with locally bounded variation.

The process X_1 is said to be a local martingale with respect to F if there are stopping times σ_n increasing without limit so that $X_{t \wedge \sigma_n}$ is a martingale with respect to $\{F_{t \wedge \sigma_n}, t \geq 0\}$.

The local martingale corresponds to the rapidly changing aspect of the semimartingale.

The process $X_2 \equiv (X_{2t}, F_t; t \geq 0)$ has locally bdd. variation if

$$\int_0^t |dX_{2S}(\omega)| < \infty \ \forall t \quad \text{and} \quad \omega \in \Omega.$$

The semimartingale is a rich class of stochastic processes encompassing all discrete processes, all exchangeable processes, and all processes with stationary independent increments.

Protter (1986) observes that stochastic integration, until recently, referred to the Ito integral, that is, integration with respect to Brownian motion. For example, many diffusions can be depicted as solutions to

---

[50] A fine discussion of the limit theorems for exchangeable processes is contained in Hall and Heyde (1980).

[51] An excellent discussion of the semimartingale's economic significance is presented in Sims (1984).

systems of stochastic differential equations represented by:

$$X_t = X_0 + \int_0^t \sigma(s,X_s)dB_s + \int b(s,X_s)ds,$$

where B is Brownian motion.[52]

Now it is "well-known" that the Ito integral with respect to Brownian motion is inappropriate both for the practitioner and the theoretician. The more general semimartingale integral is the correct form of stochastic integration.

There are two reasons why the attachment to Ito integration is difficult to sever. First, the new semimartingale approach is abstract and exceedingly difficult. Second, Brownian motion is a very attractive process. Indeed, it, among all stochastic processes, must be singled out for special consideration, being a member of three families: the martingales, the <u>strong</u> Markov processes, and the Gaussian processes. It is also exchangeable and self-similar. Yet Doob insisted on the importance of the martingale. He saw that it was the key to the Ito integral and also showed that the submartingale is decomposible into a martingale and a nondecreasing process. Meyer followed the decomposition trail and encountered many subtle obstacles. Given a filtered probability space $(\Omega, F, \{F_t\}_{t\geq 0}, P)$ with $F_s \subseteq F_t$, when $s \leq t$, $F_t = \cap_{u>t} F_u$, and F_o containing all P-null sets, consider the stochastic process H_t. H_t is a function mapping $R_+ \times \Omega$ to R. The predictable σ-algebras $\tilde{P} \equiv \sigma\{H\colon H_t \in F_{t^-}, \forall t$ and $t \to H_t$ is left continuous$\}$. The optional σ-algebras $\tilde{O} \equiv \sigma\{H\colon H_t \in F_t$ and $t \to H_t$ is right continuous$\}$, where $\sigma(\ )$ is the smallest σ-algebra on $R_+ \times \Omega$ conferring measurability on all these stochastic processes.

---

[52]For an extensive analysis see Karlin and Taylor (1982).

Now the puzzle was that for Brownian motion, $\tilde{P} = \tilde{O}$. This follows from the strong Markov property together with continuity. Yet for general martingale integrals (with respect to Brownian motion) the integrand was restricted to predictable processes. Deep reflection by Meyer saw that this limitation was natural. He proceeded to extend the theory of stochastic integration to local martingales including the generalized Ito's change of variable formula. (Ito's original formula had been generalized by Kunita and Watanabe. They showed that any submartingale can be decomposed into a local martingale and an increasing predictable process.) The semimartingale now occupies center stage in the theory of stochastic integration. It is composed of a local martingale and an increasing process, where the latter can be viewed as the difference of two increasing processes.

The semimartingale is the most general differential for which an integral exists: a local martingale integral for the martingale and a path by path Lebesque-Stieltjes integral for X_2. It has been shown that if a right-continuous process X is a linear differential satisfying a weak dominated convergence theorem, then X is a semimartingale. The semimartingale X can be decomposed into $M + A$, where M is a locally square integrable local martingale and A is a process whose paths have finite variation on the compacts. This reduces the problem of stochastic integration in that elementary Hilbert space methods greatly simplify the theory.

Protter (1986) has observed that the theory of stochastic integration is now a mature field in stochastic theory. The two key objects -- semimartingale and predictability -- have been identified. Furthermore, the law of large numbers and the functional central limit theorem have been obtained for semimartingales by Lipster and Shiryayev.[53] Thus, semimartingale theory is ready for applications. Out contention is that stochastic economics has and will continue to reap enormous rewards from applications that combine this robust and elegant mathematical edifice with the de Finetti-Savage methodology as portrayed by Lauritzen (1982).

---

[53]These are discussed in the fine survey article by Shiryayev (1981).

APPENDIX 2

Comments on the Integral Representation Theorem

At first it may seem odd that all of Bayesian inference can be represented by an integral of mixtures with exchangeability being the driving assumption.

Comment 1: The integral representation theorem is a deep result in functional analysis. Indeed, it is the generalization of two profound theorems: those by Minkowski and Krein-Millman.

A subset K of a real linear space is called convex iff the following condition obtains: Let $\{y_i\}$ $i = 1,2,\ldots,k$ be points in K; then K is convex iff

$$\sum_{i=1}^{k} \gamma_i y_i \in K; \; \gamma_i \geq 0, \quad i = 1,\ldots,k \quad \text{and} \quad \sum_{i=1}^{k} \gamma_i = 1.$$

Any point $y \in K$ is called an extreme point of K (ext K) if y can only be represented in a trivial manner, that is,

$$y = \frac{1}{2} x + \frac{1}{2} z, \quad x,z \in K$$

with $x = z = y$.

Minkowski's Theorem: Let K be a compact convex subset of a finite dimensional vector space. Then every point in K can be represented as a convex combination of K's extreme points. That is, if $K \subset S^n$, an n-dimensional space, then each element of K is a convex combination of n+1 or fewer extreme points.[71]

---

[71]For proofs and discussions of this important theorem see Phelps (1980) and Mirsky (1963).

In probability theory, economics, statistics and, of course, functional analysis K is frequently infinite dimensional and Minkowski does not suffice.

The <u>Krein-Milman theorem</u> does approximate each point in infinite dimensional convex sets by convex combinations of extreme points. However, it does not provide a representation theorem for this more abstract setting.

To remedy this consider H, a separable real Hilbert space with inner product $(\, , \,)$ and norm $(x,x) = \|x\|^2$, K a nonempty compact convex subset of H, X, an arbitrary compact Hausdorff space, $C(X)$, the Banach space of all real valued functions on X possessing the sup norm, $\|f\| = \sup\{|f(x)|: x \in X\}$, $f \in C(X)$.

From probability theory we know that the Borel subsets of X comprise the smallest σ-algebra of subset of X containing the closed subsets. A probability measure μ on X is a nonnegative measure on the Borel sets such that

$$\mu(B) = \sup\{\mu(A): A \text{ compact}, A \subseteq B\} \quad \text{and} \quad \mu(X) = 1.$$

By the Riesz representation theorem, if L is a linear functional on $C(X)$ such that $L(f) \geq 0$ whenever $f \geq 0$ and $f \in C(X)$, there exists a <u>unique</u> measure μ on X satisfying

$$L(f) = \int_X f d\mu, \quad f \in C(X).$$

Phelps notes that the Stone-Weierstrass theorem implies that if X is a compact metric space, $C(X)$ is separable.

<u>Choquet Theorem</u>: Let K be a compact convex subset of the separable Hilbert space H. For any $x \in K$, there is a probability measure μ on K such that:

(a) $\int_K f d\mu = f(x)$, for each $f \in A(K)$, the subspace of all continuous real-valued affine functions on K, and

(b) $\mu(\text{ext } K) = 1$.[72]

It is this analysis that accounts for the generality of the modern versions of de Finetti's theorem.

Comment 2: In the sophisticated econometric research on estimation,[73] mixture methods have been used to cope with heterogeneity and other specification problems. Is this work related to the de Finetti analysis? Not only is it related directly via heterogeneity, but it also reveals powerful connections among exchangeability, the design of experiments and convex programming.

Comment 3: The partial ordering of majorization in R^n and the Schur functions are closely related to heterogeneity.[74]

Let $\mathcal{R}$ be an open convex subset of R^n, where R^n has exchangeable coordinates. A function $\phi: \mathcal{R} \to R^n$ is called Schur convex if it is nondecreasing relative to the parted ordering of majorization ($<$) on $\mathcal{R}$. Schur convex functions are always exchangeable. For example,

$$\phi(x_1,\ldots,x_n) = \sum_{i=1}^{n} f(x_i) \quad \text{where} \quad f: R \to \mathcal{R} \text{ is convex.}$$

Now $x = (x_1,\ldots,x_n)$ is said to be majorized by $y = (y_{11},\ldots,y_n)$ or that y is a dilation of x, written $x < y$ if

[72] A detailed and illuminating proof is given by Phelps (1966).

[73] A fine example is Heckman and Singer (1984).

[74] A brief and incisive description of dilations is given in Kemperman (1981).

$$\sum_{i=1}^{n} f(x_i) = \int f d\mu \leq \int f dv = \sum_{i=1}^{n} f(y_i)$$

obtains for each convex function f on R, where μ and v are the appropriate measures.

A necessary and sufficient condition for (1) is

$$\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]}, \quad k = 1,2,\ldots,n-1, \tag{2}$$

and

$$\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]},$$

where $[i]$ denotes the order of the observation in the n-sample.

When the analysis is conducted in abstract spaces the dilation is the appropriate ordering. Let S be a compact metrizable space and C a convex cone of continuous functions on S. Then a measure $\mu < r$ iff

$$r(B) = \int k(x,B)\mu(dx),$$

where $k(x,B)$ is the kernel.

The dilation is also called Choquet ordering and has the martingale property. A rigorous analysis will be presented in a subsequent paper.

By indicating the more diffuse probability measure, the dilation is a natural measure of homogeneity (and entropy). By integrating over measures via mixing, heterogeneity is treated explicitly in the estimation process.

<u>Comment 4</u>: The equivalence between de Finetti's theorem and the Hausdorff moment problem clearly demonstrates the connections among exchangeability, the moment problem, interpolation and approximation theory,

Wiener-Masani prediction theory and operator theory. The relation between dilations and exchangeability reveals the connection between exchangeability and Strassen's Marginal theorem. That both the class of moment problems -- the typical example is given some subset in R, determine the probability measure on R having specified moments, perhaps the 1st, 2nd, and 3rd, so that a criterion function is optimized -- and the class of marginal problems -- the typical example is given n marginal distributions on $S_1, S_2, \ldots, S_n$ what is the joint measure on $S_1 \times \ldots \times S_n$ that optimizes some criterion function[75] -- are glued to de Finetti, means that the fundamental research on convex programming of Rockefeller and Wets and the Ford-Fulkerson network analysis are all intertwined in what may be an economically productive web.

Comment 5: The equivalence of the Bayesian representation theorem and Hausdorff's moment theorem means that the robustness analysis and approximation theory associated with sophisticated studies of complex systems is not only compatible with the Bayesian methodology, but required by it. These approximations are discussed in McCall (forthcoming).

Comment 6: The admissibility problem first detected by Stein has been studied extensively by Brown (1986) and Johnstone (1986). Johnstone (1986) shows that a unique reversible birth and death process on the positive integers is associated with every generalized Bayes estimator.[76] Johnstone estimates a Poisson mean λ based on one observation x with a quadratic loss function $(d(x)-\lambda)^2/\lambda$, with d(x) the estimator. An adroit

[75] Whitt (1976) has a clear analysis of the bivariate marginal problem with an excellent historical account of its genesis. The moment problem and the marginal problem are treated extensively in the important book by Klein-Haneveld (1986).

[76] The "generalized" estimator need not have a proper prior. For a full discussion see Berger (1985).

application of harmonic analysis shows d(x) is admissible iff a "zero energy" conservation principle obtains. In economic terms this corresponds to a "no arbitrage" condition. In a subsequent piece, we will show that hitting the boundary of a Choquet simplex corresponds to solving the admissibility problem. In both cases the equilibrium condition is a martingale. Coherence of an inductive process also is equivalent to a martingale condition. Thus, the "no arbitrage" condition is a fundamental equilibrium condition.

Comment 7: Consider the Choquet simplex in a locally convex Hausdorff space. Well-behaved functionals like Markov decision processes and dynamical systems achieve their maximum (minimum) value at an extreme point of the simplex.

REFERENCES

Aldous, D.J., "Some Inequalities for Reversible Markov Chains," J. London Math. Soc., 2 (1982), 564-76.

__________, Exchangeability and Related Topics, Ecole d'Eté St. Flour XIII, Lecture Notes in Mathematics 1117, pp. 298-409, Springer-Verlag, 1985.

Aldous, D.J., and P. Diaconis, "Shuffling Cards and Stopping Times," Amer. Math. Monthly, 93 (1986), 333-48.

Anderson, E.S., "An Algebraic Treatment of Fluctuations of Sums of Random Variables," Proc. of 5th Berkeley Symp. on Math. Stat. and Prob., 2 (1967), 423-29.

Anderson, R.M., and H. Sonnenschein, "Rational Expectations Equilibrium with Econometric Models," Rev. Econ. Stud., 52 (1985), 359-69.

Arniel, M., W.E. Diewert, S. Schaible, and I. Zang, Generalized Concavity, Plenum, 1988.

Arrow, K.J., "Rationality of Self and Others in an Economic System," Journal of Business, 59 (1986), S385-99.

__________, The Limits of Organization, W.W. Norton, 1974.

Aumann, R.J., "Correlated Equilibrium as an Expression of Bayesian Rationality," Econometrica, 55 (1987), 1-18.

Baumol, W.J., Superfairness, MIT Press, 1986.

Berger, J.O., Statistical Decision Theory and Bayesian Analysis, 2nd ed., Springer-Verlag, 1985.

Borch, K., "The Theory of Risk," Journal of the Royal Statistical Society B, 29 (1967), 432-67.

Brock, W.A., "Distinguishing Random and Deterministic Systems: Abridged Version," J. Econ. Th., 40 (1986), 168-95.

__________, W.D. Dechert and J.A. Scheinkman, "A Test for Independence Based on the Correlation Dimension," Univ. of Wisconsin Working Paper, April 1987.

Brown, L.D., "Fundamentals of Statistical Exponential Families," Inst. Math. Statist., 9 (1986).

Buchanan, J.M., "An Economic Theory of Clubs," Economica, 39 (1965), 1-14.

Buhlman, H., Mathematical Methods in Risk Theory, Springer-Verlag, 1960.

Cheng, H.-C., M.J.P. Magill and W.J. Shafer, "Some Results on Comparative Statics Under Uncertainty," Int. Econ. Rev., 28 (1987), 493-507.

Chow, Y.S., and and H. Teicher, Probability Theory, New York: Springer-Verlag, 1978.

Consonni, G., and A.P. Dawid, "Invariant Normal Bayesian Linear Models and Experimental Designs," in J.M. Bernardo, et al. (eds.), Bayesian Statistics 2, North Holland, 1985.

Cornes, R., and T. Sandler, The Theory of Externalities, Public Goods, and Club Goods, Cambridge University Press, 1986.

Dacunha-Castelle, D., "A Survey on Exchangeable Random Variables In Normed Spaces," in G. Koch and F. Spizzichino (eds.), Exchangeability in Probability and Statistics, North-Holland, 1982: 47-60.

Day, R.H., "Disequilibrium Economic dynamics: A Post-Schumpeterian Contribution," J. Econ. Behav. and Org., 5 (1984), 57-76.

de Finetti, B., "Bayesianism: Its Unifying Role for Both the Foundations and Applications of Statistics," Int. Stat. Rev., 42 (1974), 117-30.

__________, Theory of Probability, New York: Wiley, 1974

__________, Probability, Induction and Statistics, New York: Wiley, 1972.

Denker, M., Asymptotic Distribution Theory in Nonparametric Statistics, Vieweg, 1985.

Diaconis, P., "Applications of the Method of Moments in Probability and Statistics," in H.J. Landau (ed.), Moments in Mathematics, Amer. Math Soc., 1987.

__________ and D. Freedman, "On the Statistics of Vision: The Julesz Conjecture," J. Math. Psychol., 24 (1981), 112-38.

__________ and __________, "Finite Exchangeable Sequences," Ann. Probability, 8 (1980a), 745-64.

__________ and __________, "de Finetti's Theorem for Markov Chains," Ann. Probability, 8 (1980b), 115-30.

Doyle, P.G., and J.L. Snell, Random Walks and Electric Networks, Math. Association of America, 1984.

Dubins, L.E., and L.J. Savage, How to Gamble If You Must, McGraw-Hill, 1965.

Dynkin, E.B., "Sufficient Statistics and Extreme Points," Ann. Prob., 6 (1978), 705-30.

Dynkin, E.B., and A.A. Yushkevich, Markov Processes, Plenum, 1969.

Eagleson, G.K., "Weak Limit Theorems for Exchangeable Random Variables," in G. Koch and F. Spizzichino (eds.), Exchangeability in Probability and Statistics, North Holland, 1982.

Feller, W., An Introduction to Probability Theory and Its Applications, I and II, Wiley and Sons, 1968 and 1971.

French, M.J., Invention and Evolution: Design in Nature and Engineering, Cambridge University Press, 1988.

Georgii, H.O., Canonical Gibbs Measures, Springer-Verlag, 1979.

Greene, D.H., and D.E. Knuth, Mathematics and the Analysis of Algorithms, 2nd ed., Birkhäuser, 1986.

Greenwood, P.E., and A.N. Shriyayev, Contingulty and the Statistical Invariance Principle, Gordon and Breach, 1985.

Grötschel, M., L. Lovasz and A. Schrijver, Geometric Algorithms and Combinatorial Optimization, Springer-Verlag, 1988.

Hall, P., and C.C. Heyde, Martingale Limit Theory and Its Application, New York: Academic Press, 1980.

Halmos, P.R., A Hilbert Space Problem Book, 2nd ed., Springer-Verlag, 1982.

Hardy, G.H.J., E. Littlewood, and G. Polya, Inequalities, Cambridge University Press, 1934.

Heckman, J.J., "Simple Statistical Models for Discrete Panel Data Developed and Applied to Test the Hypothesis of True State Dependence Against the Hypothesis of Spurious State Dependence," Annales de l'INSEE, 30/31 (1978), 227-69.

__________ and B. Singer, "Econometric Duration Analysis," J. Econometrics, 24 (1984), 63-132.

Hewitt, E., and L.J. Savage, "Symmetric Measures on Cartesian Products," Trans. Amer. Math. Soc., 80 (1955), 470-501.

Hicks, J., A Theory of Economic History, Oxford University Press, 1969.

Johnstone, I., "Admissible Estimation, Dirichlet Principles and Recurrence of Birth-Death Chains on Z_+^P," Prob. Th. Rel. Fields, 71 (1986), 231-69.

Kailath, T., "Signal Processing Applications of Some Moment Problems, in H.J. Landau (ed), Moments in Mathematics, Amer. Math. Soc., 1987: 71-109.

Kallenberg, O., "Spreading and Predictable Sampling in Exchangeable Sequences and Processes," Ann. of Probab., 16 (1988), 508-34.

__________, Random Measures, 4thed., Academic Press, 1986.

Karlin, S., and H. Taylor, A Second Course in Stochastic Processes, Academic Press, 1981.

Kelly, F.P., Reversibility and Stochastic Networks, Wiley, 1979.

Kemperman, J.H.B., "Review of Marshall and Olkin," Bulletin of the American Mathematical Society, 1981.

Kendall, D.G., "On Finite and Infinite Sequences of Exchangeable Events," Studia Sci. Math. Hungar., 2 (1967), 319-27.

Kingman, J.F.C., "Uses of Exchangeability," Ann. Probab., 6 (1978), 183-97.

Klein Haneveld, W.K., Duality in Stochastic Linear and Dynamic Programming, Springer-Verlag, 1986.

Lauritzen, S.L., Statistical Models and Extremal Families, Aalborg University Press, 1982.

Leamer, E.E., "Model Choice and Specification Analysis," in Z. Griliches and M.D. Intriligator (eds.), Handbook of Econometrics I, North-Holland, 1983: 285-300.

Liggett, T.M., Interacting Particle Systems, Springer-Verlag, 1985.

Loève, M., Probability Theory II (4th ed.), Springer-Verlag, 1978.

Lucas, R.E., Jr., "Adaptive Behavior and Economic Theory," Journal of Business, 59 (1986), S401-26.

__________, Models of Business Cycles, Basil Blackwell, 1987.

Manne, A., "Linear Programming and Sequential Decisions," Management Science, 6 (1960), 259-67.

Marshall, A.W., and I. Olkin, Inequalities: Theory and Majorization and Its Applications, 1979.

Masani, P., "Dilations as Propagators of Hibertian Varieties," SIAM Math. Anal., 9 (1978).

McCall, J.J., "Economic Processes and Exchangeability," Hoover Institution, 1988a.

__________, "A Symmetric Approach to Economics: Dilations and the Dilation Operator," Hoover Institution, 1988b.

__________, "Exchangeability and the Structure of the Economy," in H. Brink (ed), Business Cycles, forthcoming, Macmillan.

McCall, J.J., and K. Velupillai, "Dynamic Economics: 'Stopping Times, Conditioning and Ergodicity'," forthcoming.

__________ and __________, Dynamical Systems and Exchangeability: Economic Applications, Basil Blackwell, forthcoming.

[illegible]

Olshen, R.A., "A Note on Exchangeable Processes," Zeitschrift fur Wahrscheinlichkeitstheorie verw. Gebiete, 28 (1973), 317-21.

Padmanabahn, K., and A.N. Netravali, "Dilated Networks for Phonetic Switching," IEEE Trans. Com-35, 12 (1987), 1357-65.

Pfeiffer, P.E., Conditional Independence in Applied Probability, Newton, MA: UMAP, Educational Development Center, 1979.

Phelps, R.R., Lectures on Choquet's Theorem, van Nostrand, 1966.

Pigou, A.C., Industrial Fluctuations, Macmillan, 1926.

Pitman, J., "An Extension of de Finetti's Theorem," Adv. Appl. Prob., 10 (1978), 268-79.

Poirer, D.J., "Frequentist and Subjectivist Perspectives in the Problems of Model Building in Economics," Econ. Perspect., 2 (1988), 121-70.

Protter, P., "Three Books on Stochastic Integration," Review in Ann. Prob., 14 (1986), 343-46.

Ramakrishma, M.V., "Computing the Probability of Hash Table/Urn Overflow," Commun. Statist. Th. Meth., 16 (1987), 3343-53.

Ressel, P., "de Finetti-Type Theorems: An Analytic Approach," Ann. Prob., 13 (1985), 898.

Riordan, J., An Introduction to Combinatorial Analysis, Wiley, 1978.

Rockafellar, R.T., Network Flows and Monotropic Optimization, John Wiley, 1984.

Rota, G.C., "Theory of Möbius Functions," Z. Wahrsch., 2 (1964), 340-68.

Rustem, B., and K. Velupillai, "Objective Functions and the Complexity of Policy Design," J. Econ. Dyn. Cont., 11 (1987).

Samuelson, P.A., "Spatial Price Equilibrium and Linear Programming," American Economic Review, (1952), 283-303.

Sarason, D., "Moment Problems and Operators in Hilbert Space," in H.J. Landau (ed), Moments in Mathematics, American Math. Soc., 1987.

Serfling, R.J., Approximation Theorems of Mathematical Statistics, Wiley, 1980.

Serra, J., Image Analysis and Mathematical Morphology, Academic Press, 1982.

Shiryayev, A.N., "Martingales: Recent Development Results and Applications," Inst. Statist. Rev., 49 (1981), 199-233.

Simon, H.A., "On the Behavioral and Rational Foundations of Economic Dynamics," J. Ec. Beh. Oreg., 5 (1984), 35-55.

Sims, C.A., "Martingale-Like Behavior of Prices and Interest Rates," Center for Econ. Resch., University of Minnesota, 1984.

Sinden, F.W., "A Geometric Representation for Pairs of Dual Quadratic or Linear Programs," J. Math. Anal. and Appl., 5 (1962), 378-402.

Smale, S., "Dynamics in General Equilibrium Theory," American Economic Review, 66 (1976), 288-94; reprinted in The Mathematics of Time, Springer-Verlag, 1980.

Solow, R.M., "Economic History and Economics," American Economic Review, 75 (1985), 328-31.

Speed, T.P., "Cumulants and Partition Lattices II: Generalized k-Statistics," J. Austral. Math. Soc., A40 (1986), 34-53.

Spitzer, F., "A Combinatorial Lemma and Its Application to Probability Theory," Trans. Amer. Math. Soc., 82 (1956), 323-39.

Stein, E.M., "Harmonic Analysis on R^n," in J.M. Ash (ed.), Studies in Harmonic Analysis, Math. Assoc. of Amer., 1976: 97-135.

Stewart, J., "Positive Definite Functions and Generalizations, An Historical Survey," Rocky Mtn. J. Math., 6 (1976), 409-33.

Strang, G., "A Framework For Equilibrium Equations," SIAM Rev., ;30 (1988), 283-97.

Strassen, V., "The Existence of Probability Measures with Given Marginals," Ann. Math. Statist., 36 (1965), 423-39.

Telser, L.G., "A Theory of Self-Enforcing Agreements," Journal of Business, 53 (1980), 27-44.

Velupillai, K., The History and Theory of Business Cycles, 1988.

Whitt, W., "Bivariate Distributions with Given Marginals," Ann. Statist., 4 (1976), 1280-89.

Winter, S.G., "Comments on Arrow and Lucas," Journal of Business, 59 (1986), S427-34.

Zaman, A., "A Finite Form of de Finetti's Theorem for Stationary Markov Exchangeability," Ann. Prob., 14 (1986), 1418-27.

Zeckhauser, R., Comments: Behavioral versus Rational Economics: What You See Is What You Conquer," Journal of Business, 59 (1986), S435-49.