THE STATISTICAL DYNAMICS OF SOCIO-ECONOMIC SYSTEMS

P.F. Lesse
Division of Building Research
Highett, Victoria, Australia

October 1983
CP-83-51

Contributions to the Metropolitan Study:3

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

Contributions to the Metropolitan Study:3

The Project "Nested Dynamics of Metropolitan Processes and Policies" was initiated by the Regional & Urban Development Group in 1982, and the work on this collaborative study started in 1983. The series of contributions to the study is a means of conveying information between the collaborators in the network of the project.

This paper by Paul Lesse deals with the relationship between dynamic economic changes at the microlevel and equilibrium descriptions applicable at the macrolevel. In this way the present contribution focuses on one of the basic theoretical issues in the Metropolitan Study: the possibility of relating effects observed on the aggregate level to the actions of people and economic agents whose individual decisions produce these effects. Thereby it also relates to the problem of multiple dynamic changes at the disaggregated level and the resulting static equilibrium at the global level.

The outcome of the approach is an aggregate equilibrium statistical representation in the form of an entropy maximizing probability distribution based on characteristics of the dynamic equations. The approach also makes it possible to relate entropy maximizing models to behavioral models based on cost minimizing or utility maximizing assumptions with reference to logit and probit types of models.

Börje Johansson
Acting Leader
Regional & Urban Development Group
IIASA
Laxenburg, September 1983

CONTENTS

THE STATISTICAL DYNAMICS OF SOCIO-ECONOMIC SYSTEMS

P.F. Lesse
Division of Building Research,
Highett, Victoria, Australia

Motto:

"One interested only in fruitful statics must study dynamics"

P.A. Samuelson: Foundations of Economic Analysis

SUMMARY

The paper deals with the following problems:

- relationship between the dynamic changes observed in an economic system at the disaggregated level and, equilibrium description applicable at the aggregated scale.
- relationship between entropy maximizing models and behavioral (cost minimizing, utility maximizing) models.
- the answers are provided in terms of an entropy maximizing probability distribution based on topological characteristics of the dynamic equations (Lyapunov functions).

1. INTRODUCTION

There exists a considerable body of literature dealing with socio-economic and urban planning models using entropy as a suitable concept and Jaynes' (1957a,b) principle as a handy tool for obtaining the values of variables at an aggregated level. Most of these papers can be traced back to Wilson (1966, 1970, 1974, and papers cited in the monographs) who has shown that the empirical gravity model can be derived by entropy maximization. Application of this approach led to a large number of planning models, many of which have been widely used in practice. The entropy method is thus theoretically sound and successful in practice; nevertheless, there appear to be a few aspects of it which are felt to be in need of further development.

One such aspect is the relationship between the effects observed on the aggregated level and the actions of people whose individual decisions produce these effects. Another is the problem of the relationship between the multiple dynamic changes at the disaggregated (micro) level and the resulting static equilibrium at the global level.

Both these problems have been identified by numerous writers and treated from different angles. The former problem can be seen as a problem of aggregation; the latter as a search for a relationship between descriptive and causative or behavioural models. In this sense, it can also be interpreted formally as the problem of relating the models based on maximization of entropy with those based on utility.

Utility is generally accepted as a behavioural concept well defined in terms of individual preferences over the set of socio-economic states (Papageorgiou, 1977). The problem of aggregation of the individual utility functions and the relation between the entropy and utility-based models is rather complicated and has been reviewed by Wilson (1975). An aggregated utility function was, in many cases, derived by the theory of random utility (for analysis, see Manski, 1977) whose relation to the entropy-based models was pointed out by McFadden (1974) and Cochrane (1975). The relationship between entropy maximizing models and random utility based models was explained by Anas (1982). Among those who have nontrivial reservations about the entropy maximizing technique, and who demand that a link connecting the behavioural and descriptive models should be established, belongs T. Smith (1978). T. Smith formulated an alternative theory based on the observation that the high cost trips are in general less probable than those with lower cost (cost efficiency principle). In a similar spirit entropy has been treated as a measure of accessibility by Erlander (1977, 1980). Boyce and Jansen (1980) see entropy as a concept related to spatial interaction among traffic flows and used as a smooting device (1981). The various aggregated forms of utility or consumer benefit were studied by Williams (1976, 1977). Wiliams and Senior (1977) have studied the appropriate measure of the consumer benefit in the field of locational analysis. The dual relations between models based on various utility-like functions and those maximizing entropy were explored by Wilson and Senior (1974) and Coelho and Wilson (1977). The common feature of these papers is the derivation of an aggregate measure of benefit (usually related to individual preferences) which attains its maximum at certain values of parameters describing the state of the system. Identical values are then shown to be obtained by maximizing entropy subject to certain constraints.

The derivation of the aggregate utilities is in general static, i.e. the dynamic nature of the decision making process is seldom tackled. Recently Bertuglia and Leonardi (1979) have published a dynamic model based on the theory of Markov processes. Wilson (1978) suggested using the theory of optimum control for this purpose and briefly outlined a way how such a result could be achieved.

In this paper I propose a theory capable of at least partly explaining the relationship between the utility maximizing or cost minimizing dynamics observed at the micro economic level, and the static entropy maximizing models describing the aggregate behaviour. The theory is based on a new approach to large dynamic models of socio-economic systems (e.g. Isard and Anselin, 1980; Kohno, Yoshida, Mitomo, 1981; Kohno, Higano, 1982).

The method is based on a very simple idea:
The equations describing the dynamics of a complicated economic systems contain two kinds of information: (i) 'ephemeral' information which affects the system's behaviour only for a short time (ii) 'essential' or important information which qualitatively determines the behaviour of the system.

The former is of little importance to the behaviour of the system as a whole and can be expected to be 'aggregated out' in a corresponding macro model. The latter determines the system's overall behaviour and is to be retained in the macro description. The aggregated model is thus a picture of reality which cancels out the ephemeral changes observed at the micro level and retains the results of the essential changes.

Mathematically, the ephemeral information in many cases includes the initial conditions; the essential comprises topological characteristics of solutions such as Lyapunov functions and constants of motion.

We shall distinguish three basic descriptions of a dynamic system:

1) The ordinary dynamic description of the system's behaviour which is

obtained by solving the dynamic equations for a particular set of initial conditions and hence does not distinguish between the ephemeral and essential. This category also comprises the models of optimal growth using the theory of optimal control, differential games, etc.

2) The statistical description which uses the essential information for construction of a probability distribution. The probability distribution is defined in a space of the system's variables and regards the nonessential information as a random influence.

3) The macro (aggregated, or phenomenological) description in terms of quantities which are averages and higher momenta of the microvariables. The averaging process uses the probability distribution provided by the statistical description to suppress the ephemeral information and to highlight the essential. An outline of this approach has been given elsewhere (Lesse, 1982).

In the following I shall construct the statistical descriptions corresponding to several classes of dynamic models.

2. EQUILIBRIUM STATISTICAL REPRESENTATION OF LARGE DYNAMIC SYSTEMS

2.1 A few remarks on stability and boundedness of differential equations

The following elementary facts can be found e.g. in Hirsch and Smale (1979) or Varian (1981).

<u>Remark 2.1</u>. Consider a socio-economic system which can be modelled by a set of integrable differential equations

$$\dot{x} = f(x), \quad f(0) = 0, \quad x(0) = x_o \qquad (2.1)$$

where $x \in \Phi \subset R^n$, Φ is a compact subspace.

This system is stable at the origin $x = 0$ if there exists a positive definite function $V(\ )$ (Lyapunov function): $\Phi \to R$ which is nonincreasing along the paths generated by (1), i.e. which has the property

$$\dot{V} = \nabla V(x)\, f(x) \leq 0 \qquad (2.2)$$

If $-\dot{V}$ is negative definite the stability of the origin is asymptotic.

By removing the condition $f(0) = 0$ in (2.1) we obtain a more general dynamic system with the equilibrium point not necessarily at the origin. In such a system the existence of a Lyapunov function implies that the solutions are uniformly bounded, i.e. for any $\alpha > 0$ there exists a constant $\beta > 0$ such that $|x_o| \leq \alpha$ implies that the solutions of (2.1) are bounded: $|x(t, t_o, x_o)| < \beta$ for all $t > t_o$.

Remark 2.2

If the system (2.1) is linear the Lyapunov function is a positive definite quadratic form $V = x^T Hx$ with nxn matrix H which is a solution of a matrix equation

$$A^T H + HA = -G \qquad (2.3)$$

where G is any nxn symmetric positive definite matrix and A is the operator defining the linear system:

$$\dot{x} = Ax, \ x(0) = x_0 \ . \qquad (2.4)$$

The existence of the positive definite matrix H satisfying (2.3) is both sufficient and necessary for the asymptotic stability of (2.4) (Siljak, 1978).

Remark 2.3

Φ being compact, V: $\Phi \to R$ is bounded, i.e. there is a constant M such that $V(x) \leqq M, \ \forall x \in \Phi$

2.2 The probability distributions associated with dynamic systems

Definition 2.1:

The probability that a differential neighbourhood of an arbitrary point $x_a \in \Phi$, i.e. the interval $[x_a, x_a + dx]$, contains a solution of the system (1) irrespective of initial conditions is

$$P(x) = \exp\,[\Omega - \lambda\, V(x)] \qquad (2.5)$$

where $\Omega = -\log \int_{x \in \Phi} \exp\,[-\lambda\, V(x)]\, dx \qquad (2.6)$

and λ satisfies

$$\int_{x \in \Phi} P(x)\, V(x)\, dx \leqq M \qquad (2.7)$$

We observe that the probability distribution (2.5) maximizes Shannon entropy subject to the usual normalization condition and to (2.7). The virtue of Shannon entropy as a probability estimator rests on the well known arguments of Jaynes (1957a,b).

Definition 2.1 has the following significance: If the system (2.1) is too large and/or if the initial conditions are not known and therefore if the solution cannot be obtained, Definition 2.1 paves the way to a statistical description of the dynamic system. This description is time indepedent, therefore it may be called the equilibrium statistical representation (ESR) of the dynamic system. The notion of equilibrium whose meaning has often been discussed in economic literature in the past (e.g. Samuelson, 1948; Hicks, 1939), and at present (Andersson and Persson, 1980; Erlander, 1982) is thus given a new, and I believe a more precise meaning: An equilibrium is that statistical description of a dynamic (micro) economic system which is based on the knowledge of Lyapunov function(s) only. This definition generalizes that used in physics (Katz, 1967). In addition the definition resolves the old and vexing question whether an economic system at a particular time is, or can be, at equilibrium. The answer is that no system can 'be at equilibrium' in this sense as equilibrium is not a property of the system but refers to a given description which in turn depends on the state of our knowledge. However, it is legitimate to ask when the equilibrium description of a system is adequate for a given purpose.

This question in effect tests the reliability of the ESR and therefore it is important both from the philosophical and practical point of view.

2.3 Reliability of ESR

ESR is based on the maximization of entropy and hence any investigation of reliability should start with the appraisal of this method for assigning probabilities. However, so much has been written on this subject in recent years that only a quotation should be sufficient: "The Principle of Maximum Entropy, like Ockham, tells us to refrain from inventing Urn Models when we have no Urn" (Jaynes, 1978). The reader who has reservations concerning the method is referred to the excellent exposition from which the quotation has been taken. Alternatively, those with allergies towards physical sciences can consult the monographs by Wilson (1970), (1974) or Webber (1980) on this subject.

If the maximization of entropy is accepted as a valid method for estimation or probabilities, we can seek answers to the two questions which determine the reliability of ESR:

1) what is the significance of ESR, i.e. what kind of information does ESR provide?
2) how can this information be verified by data?

I shall deal with these questions in more detail now.

Information provided by ESR

ESR is a probabilistic representation of the system obtained by disregarding the nonessential features of the dynamics. As a consequence it models the system from a somewhat detached point of view, making it possible to omit the micro level details of the dynamic changes which do not affect the overall picture of the system. This overall picture is obtained by using the expected (average) values of the variables considered as interesting or important.

A simple example can serve as an illustration.

2.4 An Example

Consider a Samuelson (1947) dynamic model of competitive market given by the equations

$$\frac{dp_i}{dt} = k_i\ f_i(p_1, \ldots p_n) \qquad i=1, \ldots n, \qquad (2.8)$$
$$p_i(0) = p_{io}$$

where p_i is the price, $f_i\ (p_1, \ldots p_n)$ the excess demand and k_i the price adjustment coefficient of the i-th commodity. Following the usual linearization procedure (Siljak, 1978) (2.8) can be transformed into

$$\frac{d\hat{p}}{dt} = A\ \hat{p}\ ,\ \hat{p}\ \ \varepsilon\ E^n, \qquad (2.9)$$

where $\hat{p} = p - p^*$ is the excess of the price vector p over the equilibrium price p^*, and A is a stable Metzler nxn matrix.

It can be seen that (2.9) is a special case of (2.4). If p(0) is not known the classical analysis does offer very little beyond the determination of the equilibrium price p^* towards which the system tends.

In contrast, the statistical approach makes it possible to determine the positive definite matrix H from (2.3) and also the corresponding ESR which has the form:

$$P(\hat{p}) = \exp\ [\Omega - \hat{p}^T\ \lambda\ H\ \hat{p}]. \qquad (2.10)$$

The multiplier λ can be determined using (2.7) if there is a suitable Φ and if $\sup\limits_{\hat{p}\varepsilon\Phi}\ \hat{p}^T\ H\ \hat{p} = M$ can be found.

ESR makes it possible to answer the following type of questions:

1) If prices are confined to a subspace Φ, what is the expected average deviation of price from the equilibrium price, $<\hat{p}>$?

The answer is, of course

$$<\hat{p}> = \int_{\hat{p}\varepsilon\Phi} P(\hat{p})\ \hat{p}\ d\hat{p} \qquad (2.11)$$

2) How much does $\hat{p}$ deviate from $<\hat{p}>$? The deviation is measured by the variance

$$\text{var } \{\hat{p}\} = <\hat{p}\ \hat{p}^T> - <\hat{p}><\hat{p}>^T, \qquad (2.12)$$

where $<\hat{p}\ \hat{p}^T> = \int_{\hat{p}\varepsilon\Phi} P(\hat{p})\ \hat{p}\ \hat{p}^T\ d\hat{p}$ (2.13)

We can make the following observations:

1) ESR of a stable linear dynamic system is a normal (Gaussian) probability distribution with zero mean and a variance matrix Q determined by the Lyapunov function H and by the Lagrange multiplier λ ($Q^{-1} = \lambda H$). Conversely, equilibrium probit models can be viewed as ESR's generated by stable dynamic linear systems describing the changes at the micro level.

2) Using equation (2.3) it is possible to find a number of Lyapunov matrices H corresponding to a given dynamic operator A and generated by various G. Obviously, those H leading to sharper probability distributions are to be preferred. It is possible to formulate a problem of finding the optimum Lyapunov function which yields a probability distribution (2.10) with variance in some

sense minimal.

3) For a given H the definition of the feasible domain Φ determines the upper bound M and, finally λ. It can be seen that the smaller M the sharper the probability distribution will be. This is in agreement with common sense, of course, as M in general decreases with decreasing size of the feasible domain which in turn reflects a more constrained (i.e. more predictable) behaviour of the system.

4) The amount of information provided by ESR can be measured by inserting the probability distribution into the definition of Shannon entropy.

The result is an estimate

$$S \leq \lambda M - \Omega$$

with $\Omega = - \log \int_{\hat{p} \epsilon \Phi} \exp [- \hat{p}^T \lambda H \hat{p}] d\hat{p}$.

The bilinear form on the r.h.s. can be seen as a basis for the macroscopic (phenomenological) description of the system (Lesse, 1982). The larger the S the less information is inherent in the corresponding ESR.

2.5 Verification of ESR

In the foregoing section it was shown that ESR can provide some information about the behaviour of the system even under conditions which make the use of other modelling methods impracticable, for example, when the state of the system at present is unknown. However, the question remains how this information, e.g. the values $\langle\hat{p}\rangle$, $\langle\hat{p}\,\hat{p}^T\rangle$ provided by (2.11), (2.13), could be verified by the data or observations.

Let us assume that there are some quantities depending on the state of the socio-economic system whose values can be observed. Referring to the example 2.4 one such quantity can be 'the price of a standard basket of goods' defined by $\pi = \sum_i q_i\, p_i$, where q_i is an arbitrarily determined quantity of commodity i which forms a part of the basket. The equilibrium price p* being known, it is possible to calculate $\pi - \sum_i q_i\, p_i^* = \pi_\Delta$, i.e. the difference between the price of a standard basket at a given time and that at equilibrium. Repeated observations of π_Δ can be considered as a stochastic process $\pi_\Delta(t)$, i.e. as a family of random variables π_Δ indexed by t. The index t is to be interpreted as the time when the observation was made.

Let us introduce a few auxiliary definitions

Definition 2.2: The function $R(\lambda)$ defined by

$$R(\lambda) = \langle [\pi_\Delta\ (t+\lambda) - \langle\ \pi_\Delta\ (t+\lambda)\ \rangle][\pi_\Delta(\lambda) - \langle\ \pi_\Delta\ (\lambda)\rangle]\ \rangle$$

is called covariance of the stochastic process $\pi_\Delta(t)$.

The sign < > denotes, as before, the expected value (c.f. (11)). We shall occasionally refer to this expected value as to the ensemble average.

Definition 2.3: The random variable $\bar{\pi}_\Delta(T)$ defined by $\bar{\pi}_\Delta\ (T) = \frac{1}{T}\int_0^T \pi_\Delta\ (t)\ dt$, T >

is called the time average of the stochastic process $\pi_\Delta(t)$. T is the averaging time.

The integrals of stochastic processes are to be understood as limits in the mean here.

Definition 2.4: A stochastic process $\pi_\Delta(t)$ satisfying

$$\lim_{T\to\infty} \frac{1}{T^2} \int_{-T}^{T} (T - |\lambda|)\, R(\lambda)\, d\lambda = 0$$

is called strongly ergodic in the mean.

Now we can state the following proposition:

Proposition 2.1:

If a stochastic process $\pi_\Delta(t)$ is strongly ergodic in the mean then $\bar{\pi}_\Delta(T) = \langle \hat{p} \rangle$ with probability 1,

i.e. (i) $\langle \bar{\pi}_\Delta (T) \rangle = \langle \hat{p} \rangle$ (i)

and

(ii) $\lim_{T\to\infty} \langle \bar{\pi}_\Delta (T) - \langle \hat{p} \rangle \rangle = 0$. (ii)

The proof of Proposition 2.1 is elementary and can be found in textbooks on stochastic processes, e.g. Melsa et al. (1973).

The assertion contained in the proposition is worth examining in detail.

The most important outcome is the connection between the estimate $\widehat{\langle p \rangle}$ provided by ESR (cf. (2.11)) and the observed values $\pi_{\Delta}(t)$. According to (i), (ii) the estimates $\langle p \rangle$ are unbiased estimators of the time averages of observed values $\pi\ (t)$. This makes it possible to verify the validity of ESR in any given case.

The premise of the proposition requires that the observed quantities should behave as a strongly ergodic process. Using the Definition 2.4 it is possible to interpret this requirement in practical terms as a prohibition to use ESR for predicting the values of those system's characteristics whose measurements 'do not forget past history'. This is a sensible limitation: If we are interested in those properties of the system which are determined by history and this dependence on history does not fade away with passing time then the knowledge of the initial state of the system is an essential desideratum and ESR is clearly inadequate. On the other hand, if the behaviour of the system can be described for a given purpose in terms of quantities which ultimately become independent of the past, the equilibrium representation can be useful.

It is perhaps not necessary to remark that while the use of ESR has been demonstrated using the Example on a particularly simple (linear) system, the method is applicable to dynamic systems of considerably wider generality. The following section should justify this assertion.

3. SOME OTHER DYNAMIC SYSTEMS AND THEIR EQUILIBRIUM STATISTICAL REPRESENTATIONS

3.1 Differential games

A few results from the field of optimum control and differential games are needed. The standard monographs such as Isaacs (1965), Friedman (1971), Intriligator (1971) should be consulted for more detailed information. More advanced reviews include Varaiya (1971) and Gupta (1981).

Let there be a set of differential equations

$$\frac{dx}{dt} = f(x,u_1,\dots,u_N,t) \tag{3.1}$$

with initial considerations $x(0) = x_0$; with the state vector $x \in R^K$ and control vectors $u_1(t),u_2(t),\dots,u_N(t)$ being measurable functions of time with values $u_1 \in \Omega_1$, $u_2 \in \Omega_2,\dots,$ $u_N \in \Omega_N$ where the control sets Ω_k ($k=1,\dots,N$) are compact subsets of some Euclidean space with appropriate dimensions. The functions $f(\ )$ are assumed to satisfy the standard conditions of integratility for any $u_k \in \Omega_k$.

Further, let there be a list of N functions $h_k(x_1,u_1,\dots,u_N,t)$ continuous in all arguments in the sets defined above and another list of functions $g_k(t_f,x(t_f))$ which are bounded for $t_f,x(t_f)$ from a closed terminal set $F:[0,T] \times R^K$. The cost function of the i-th player $J_k(u_1,\dots,u_N)$ is defined as

$$J_k(u_1,\dots,u_N) = g_k(t_f,x(t_f)) + \int_0^{t_f} h_k(x,u_1,\dots,u_N,t)\,dt \tag{3.2}$$

$(k=1,\dots,N)$.

The equations (3.1), (3.2) together with a criterion of optimality define a differential game.

If the equations (3.1), (3.2) do not depend on time explicitly the game is called autonomous.

The most frequently met criteria of optimality are

(a) Pareto optimality

(b) security pay off

(c) Nash equilibrium

The Pareto optimal strategies $\check{u}_k$ satisfy $J_k(\check{u}_1, \dots \check{u}_N) \leq J_k\ (u_1, \dots u_N)$ $\forall\ u_k\ \varepsilon\ \Omega_k$ $k=1, \dots N$

The Pareto optimal strategies correspond to a situation when all participants cooperate to achieve a minimum cost solution.

The security pay off strategies $\bar{u}_k$ satisfy

$$J_k\ (\bar{u}_1,\ \bar{u}_2, \dots \bar{u}_k, \dots \bar{u}_N) = \max_{u_1 \varepsilon \Omega_1} \dots \min_{u_k \varepsilon \Omega_k} \dots \max_{u_N \varepsilon \Omega_N} J_k\ (u_1, \dots u_k \dots u_N)$$

These strategies minimize the damage to the k-th player caused by a concerted action of the rest of the players.

Nash equilibrium strategies $\hat{u}_k$ are defined by

$$J_k\ \hat{u}_1 \dots \hat{u}_{k-1},\ u_k,\ \hat{u}_{k+1} \dots \hat{u}_N) \geq J_k\ (\hat{u}_1 \dots,\ \hat{u}_k, \dots\ \hat{u}_N).$$

The Nash equilibrium strategies characterize a state which is stable in the sense that any player deviating from the equilibrium strategy is penalized by an increased cost.

The solutions of these games can be obtained in the following way:

A differential game optimal in the Pareto sense can be reduced to solving a simple dynamic optimization problem with the objective

$$J = \sum_k \mu_k J_k \qquad \sum_k \mu_k = 1, \; \mu_k \geq 0$$

(Friedman, 1971).

A security pay off differential game can be solved as a set of N zero sum games each played by a single player against an aggregate opponent formed by all the remaining participants.

Nash equilibrium solution can be obtained by using the following theorem (Friedman, 1971):

Theorem 3.1.

If the functions $f(\ )$, $g_k(\ )$, $h_k(\ )$ are continuously differentiable in all their arguments and if the equilibrium strategies $\hat{u}_k$ are Lipchitz continuous then there exists a vector valued function $W_k(t,x)$ satisfying on the boundary of the terminal set $W_k = g_k$, and in the set $[0,T] \times R^K \times \Omega_1 \times \ldots \Omega_N$ the equations

$$\frac{\partial W_k}{\partial t} + \min_{u_k \varepsilon \Omega_k} \Big(f(x,\ t,\ u_1,\ \ldots\ u_k,\ \ldots\ u_N) \frac{\partial W_k}{\partial x} + \\ + h_k\ (x,\ u_1, \ldots\ u_k,\ \ldots u_N,\ t) \Big) = 0 \qquad (3.3)$$

The function $f(x, t, u_1, \ldots u_N)\, p_k + h_k(x, u_1 \ldots u_n, t) = H_k(x, t, p_k, u_1 \ldots u_N)$ is called the hamiltonian of the k-th player.

The new variables $p_k = \dfrac{\partial W_k}{\partial x}$

are the shadow prices corresponding to the state variables x and associated with the k-th player.

Differential games are a natural generalization of the optimum growth models (e.g. Intriligator, 1971; Isard et al., 1979) and hence their ESR can be of considerable practical interest.
The functions W_k can be related to Lyapunov functions:

Theorem 3.2 (Stalford, Leitman, Skowronski)
Consider a two person zero sum game, i.e. N=2 with player no.1 minimizing, and player no.2 maximizing J. Let there be a continuously differentiable Lyapunov function

$V(x,t) : \Phi \to R$ with the following properties:

a) $a(\| x \|) \leq V(x,t) \leq b(\| x \|)$

b) $\dfrac{\partial V}{\partial t} + \dfrac{\partial V}{\partial x} f(x,t,u_1^*, u_2) \leq - h(x,t,u_1^*, u_2)$
for all $u_2 \in \Omega_2$, $x \in \Phi$, $t \geq 0$

c) $\dfrac{\partial V}{\partial t} + \dfrac{\partial V}{\partial x} f(x,t,u_1,u_2^*) \geq - h(x,t,u_1,u_2^*)$
for all $u_1 \in \Omega_1$, $u_\eta \in \Omega_2$ $x \in \Phi$, $t \geq 0$

where $u_1^* \in \Omega_1$, $u_2^* \in \Omega_2$, and $a(\| x \|)$, $b(\| x \|)$ are positive, continuous and increasing functions.

If the conditions a), b) and c) are satisfied then $u_1^* = \hat{u}_1$, $u_2^* = \hat{u}$, i.e.

$$J(u_1^*, u_2) \leq J(u_1^*, u_2^*) \leq J(u_1, u_2^*)$$

Proof is easy and follows the ideas of Stalford and Leitman (1971), and Skowronski (1977). A generalization to N-person differential games is available (Stalford and Leitman, 1973).

We observe that according to b) and c)

$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(x, t, u_1^*, u_2^*) + h(x, t, u_1^*, u_2^*) = 0 \qquad (3.4)$$

The following proposition follows immediately from (3.4) and from Remark 2.1:

Proposition 3.1

Let there be a function $V(x,t)$ satisfying the conditions. of Theorem 3.1 and let $h(x,t,u_1^*,u_2^*) \geq 0$, $x \in \Phi$, $t \geq 0$. Then the optimal trajectories of the differential system (3.1) are bounded.

3.2 Equilibrium statistical representation of some games

The ESR corresponding to a set of autonomous differential equations with bounded trajectories has been introduced in Section 2. Proposition 3.1 makes it possible to treat certain differential games in the same way.

Definition 3.1

The equilibrium statistical representation of socio-economic system whose dynamics is described by an autonomous optimum growth model or by an autonomous two person zero sum game and which satisfies Proposition

3.1 is given by the following equations

$$P(x) = \exp\,[\Omega - \lambda\, V(x)] \qquad (3.5)$$

$$\int_{x \varepsilon \Phi} P(x)\, V(x)\, dx \leq M$$

$$\max_{x \varepsilon \Phi} V(x) = M$$

Interpretation

a) Lyapunov function

The interpretation of Lyapunov function in the context of differential games is of special interest.

We observe that the solution of the equation (3.4) corresponding to an autonomous system in general has the form $V(x,t) = Et + V(x)$, where E is an arbitrary constant whose magnitude does not change the character of the dynamic equations. The function $V(x)$ thus coincides with $V(x,t)$ either if $E=0$, or if $t=0$, and hence $V(x,t)$ and $V(x)$ can be identified by an appropriate choice of the time scale. The function $V(x,t)$ is to be interpreted as the present value of cost corresponding to the initial state, and accrued along the optimum path:

$$V(x,t) = \int_t^{t_f} h(x,t,u_1^*,u_2^*)\, dt;$$

we observe that $\frac{dV}{dt} = -\, h(x,t,u_1^*,u_2^*)$ in agreement with (3.4)

The Lyapunov function $V(x)$ is then the present value of minimum/optimum cost (i.e. of the cost accumulated along the optimum path) evaluated at $t=0$.

b) Probability distribution

It can be seen that the ESR given by Definition 3.1 assigns highest probability to the path with minimum cost. This is certainly reasonable from the common sense point of view and perhaps it can be interpreted as cost effiency in the Smith (1978) sense. This cost efficiency is contingent upon the nonnegativity of the Lagrange multiplier λ. However, Lagrange multipliers corresponding to the minimization of real-valued convex functionals on convex subsets subject to inequality constraints $G(x) \leq 0$ are nonnegative provided that $G(x)$ is convex (e.g. Luenberger, 1969, p.213ff). These requirements are not difficult to satisfy as the negative entropy functional is convex and the condition a) in Theorem 3.2 makes $V(x,t)$ bounded by two convex functions $a(\| x \|)$, $b(\| x \|)$.

3.3 An Example

Let there be a large number of firms each attempting to follow a planned path. For simplicity I shall assume that the planned growth of the i-th firm can be described in terms of a single economic indicator $\eta_i(t)$. The actual growth is given by $\xi_i(t)$.

The growth fluctuation is $x_i(t) = \dfrac{\xi_i(t) - \eta_i(t)}{\eta_i(t)}$.

The fluctuations have a dynamics

$$\frac{dx_i}{dt} = F(x_i, u_i, w_i),$$

where u_i is the control variable available to the firm for stabilization of fluctuations, and w_i is the lumped effect of exogenous destabilizing factors. Assuming that the deviations from the planned path are small we can linearize and obtain

$$\frac{dx_i}{dt} = a\, x_i + b\, u_i + w_i \tag{3.6}$$

where the coefficients a, b are assumed to be identical for all firms. The cost of stabilization is assumed to be

$$J_i = \frac{1}{2}\int_0^{t_f} q\, x_i^2 + r\, u_i^2 - s\, w_i^2\, dt \tag{3.7}$$

where t_f, q, r, s > 0, $r < s\, b^2$. The firm seeks a policy u_i minimizing the cost irrespective of the influence of the exogenous factors. The problem can be treated as a zero sum game played by the firm against nature.

The hamiltonian is

$$H_i = \frac{dV}{dx_i}(a\, x_i + b\, u_i + w_i) + \frac{1}{2}(q\, x_i^2 + r\, u_i^2 - s\, w_i^2).$$

The optimum controls are

$$u_i^* = -\frac{\partial V}{\partial x_i}\frac{b}{r}\ ;\ w_i^* = \frac{\partial V}{\partial x_i}\cdot\frac{1}{s}$$

The optimum hamiltonian becomes

$$H_i^{opt} = \frac{1}{2}\left(\frac{1}{s} - \frac{b^2}{r}\right)\left(\frac{dV}{dx_i}\right)^2 + a\, x\, \frac{dV}{dx_i} + \frac{1}{2}\, q\, x_i^2$$

To obtain the optimal path we shall not follow the usual procedure of solving the canonical equations, instead, we shall seek the Lyapunov function in the form

$$V = \frac{1}{2} K x^2, \text{ where } K > 0 \text{ is to be determined.} \qquad (3.8)$$

The equation (3.4) is now

$$\frac{1}{2} \left(\frac{r-sb^2}{rs}\right) K^2 x_i^2 + a K x_i^2 + \frac{1}{2} q x_i^2 = 0$$

which can be solved for K.

The positive root is

$$K = \frac{rs}{sb^2-r} [a + \sqrt{D}] \qquad (3.9)$$

$$D = a^2 + \frac{(sb^2-r)q}{rs}$$

We observe that

$$\frac{dV}{dt} = - \sqrt{D} \, . \, K x_i^2$$

and hence V qualifies as Lyapunov function

The optimum strategies

$$u_i^* = - \frac{b}{r} K x_i \ , \ w_i^* = \frac{1}{s} K x_i$$

yield the dynamics

$$\frac{dx_i}{dt} = - \sqrt{D} \, x_i$$

with trajectories

$$x_i = x_{io} \exp(-\sqrt{D}\, t) \qquad (3.10)$$

where x_{io} are the initial fluctuations.

If the number of firms is very large the cost of collecting the data on x_{io} may be excessive. Therefore, unless we use ESR, the only practical conclusions we can draw from (3.10) are

(i) the fluctuations tend to zero,

(ii) the halftime of stabilization is $t_{1/2} = \frac{\ln 2}{\sqrt{D}}$

(halftime of stabilization is a time required for reducing the initial fluctuation to half of its size), and hence after about $10.t_{1/2}$ the system will be practically stabilized.

To demonstrate the connection between dynamics and ESR we shall consider the initial fluctuations as random and equation (3.10) as defining a stochastic process

$$x(t) = x_o \exp(-\sqrt{D}\, t) \qquad (3.11)$$

where x_o is a random variable uniformly distributed in some interval Φ.

It is possible to show that this process is strongly ergodic in the mean in the sense of Definition 2.4. Proposition 2.1 indicates that for such a process the time average of x(t) (cf. Definition 2.3) equals the ensemble average $\langle x \rangle$ with probability one. Therefore, the expected value of x corresponding to all possible realizations of the system irrespective of initial conditions is an unbiased estimator of the time average. The ensemble average can be obtained, as in Section 2, using the probability distribution

$$P(x) = \exp [\Omega - \lambda V(x)], \qquad (3.12)$$

where $V(x) = \frac{1}{2} K x^2$,

λ is to be determined from

$$\int_{x \varepsilon \Phi} \exp [\Omega - \frac{\lambda}{2} K x^2] \frac{1}{2} K x^2 \, dx \leq M$$

$$\max_{x \varepsilon \Phi} \frac{1}{2} K x^2 = M,$$

and K is given by (3.9).

The example should illustrate the following points:

In large systems the variables may be aggregated by considering them as stochastic processes. For example, the equations (3.10) have been replaced by the equation (3.11). The time averages of these processes sometimes (cf. Proposition 2.1) can be replaced by ensemble averages, i.e. by averages taken over all possible realisations of the system. The corresponding probability distribution is obtained from the essential features of the microdynamics as represented by Lyapunov functions, and from some pre-estimates of the system's behaviour (M, Φ). The Lyapunov function is interpreted as the minimum or optimum cost.

The cost minimizing microdynamics thus forms a natural basis for the descriptive statistical representation.

4. Discussion

The patient reader will have realized that the underlying theme of this paper is simple to the point of triviality: Take a set of dynamic equations, extract whatever information on the behaviour of the solution can be obtained from the structure of the equations, and use it for estimating the probability that the neighbourhood of a point in a suitably defined space is a part of some trajectory. If the information describing the system's behaviour is time independent, the probability distribution is stationary and the corresponding statistical model corresponds to equilibrium.

This recipe makes it necessary to use certain ingredients: the behaviour of a dynamic system is qualitatively analyzed in terms of Lyapunov functions. The probability estimation procedure must be sufficiently general to admit Lyapunov functions preferably without any further ad hoc considerations. This points to the Principle of Maximum Entropy. Indeed, the feature of entropy maximization which is sometimes criticized as its weakness, namely, its lack of behavioural meaning, is turned here into a strength: the behavioural characteristics of the statistical model are supplied by the person formulating the behavioural microdynamics, not by the statistician.

The flavour of the product is, of course, to a large degree determined by the ingredients. However, the formal mathematical considerations guiding the construction of the theory lead to results with an interesting economic interpretation. This was illustrated by the example in Section 3.3, where a probability distribution was derived which depended on the cost function in a sensible manner. Of course, this is due to the fact that some Lyapunov functions have reasonable economic interpretations, e.g. as valuation functions, cost functions, etc. As a consequence, the cost function, or, in a complementary formulation, utility, enters the probability distribution neither by design nor by accident: it is there by mathematical necessity, because it is an entity defining the character of the optimal trajectories at the micro level.

We have seen that the Lyapunov functions associated with the dynamics lead to reasonable results. However, it may be that there are some other characteristics which can also be used for estimating probabilities. In this context it is instructive to explore the possibility of relaxing the requirement that ESR should be based on Lyapunov functions and replace the cost function by the total utility flow/total expenditure flow, i.e. by the hamiltonian (cf. equation (3.3)). This possibility is rather seductive for several reasons: Firstly, the hamiltonian (or the total expenditure flow) H unifies both aspects of the optimization process i.e. the optimum allocation of the state vector x and the optimum valuation as represented by the shadow prices p. As a consequence, it is an important characteristics of the dynamics. "We find the geometry of the hamiltonian function to be the fundamental determinant of the long-run behaviour of a competitive dynamical system" (Cass and Shell, 1976). The solution of Pareto optimal games, of two person zero sum games and, by implication, the security pay-off solutions of N-person games can all be obtained (under certain plausible assumptions) by solving canonical equations

$$\frac{dx}{dt} = \frac{\partial H_{opt}}{\partial p}; \quad -\frac{dp}{dt} = \frac{\partial H_{opt}}{\partial x} \qquad (4.1)$$

Further, if the game is autonomous, the hamiltonian is a constant of motion,

i.e.

$$\frac{dH_{opt}}{dt} = \frac{\partial H_{opt}}{\partial t} + \frac{\partial H_{opt}}{\partial x}$$

$$+ \frac{dx}{dt} + \frac{\partial H_{opt}}{\partial p}\frac{dp}{dt}$$

and, using (4.1), if $\frac{\partial H_{opt}}{\partial t} = 0$ then

$$\frac{dH_{opt}}{dt} = 0 \text{ , or } H_{opt} = \text{const.}$$

As a consequence, a probability distribution $P(x,p) = \exp [\Omega - \lambda H(x,p)]$ defined over the phase space $[x,p]$, with λ determined by $\int\int P(x,p)\, H(x,p)\, dxdp = \text{const.}$ appears to make sense. Indeed, this is the type of probability distribution which one meets in statistical mechanics.

However, there is an important difference between physics and economics. The physical hamiltonians can be positive definite whereas the economic ones are in general indefinite (Samuelson, 1972; Rockafellar, 1973; Cass and Shell, 1976). Generally speaking this is due to the fact that the physical analogy of shadow prices (linear momentum) is brought into the mechanical hamiltonian via a positive definite quadratic form (kinetic energy). In contrast the dependence of the (optimum) economic hamiltonian on shadow prices is the result of interaction between the dynamic constraints and optimal policies. This means that the simplest quadratic hamiltonian in physics has the form $H_{phys} = \frac{1}{2}(x^2 + p^2)$ and in economics $H_{ec}^{opt} = \frac{1}{2}(x^2 - p^2)$. As a consequence the integral $\int\int \exp(\Omega - \lambda H(x,p))\, H(x,p)\, dxdp$ converges in physics but not necessarily in economics. It appears that the hamiltonian is useful for probability estimation only so long as it retains the Lyapunov property and hence its use in an economic statistical theory is of limited value.

The construction of ESR for a given dynamic system thus depends on us finding a suitable Lyapunov function. According to Remark 2.2 for a linear dynamic system the existence of a Lyapunov function is equivalent to asymptotic stability. Therefore, an ESR of an autonomous linear asymptotically stable dynamic system always exists. Can we make a similar statement for a more general dynamic system?

The answer is provided by the Converse Theorems on Stability and Boundedness (Yoshizawa, 1975). Broadly speaking the existence of a Lyapunov function is guaranteed if the dynamic system is locally Lipshitzian and if it is uniformly asymptotically stable.

We thus arrive at a conclusion:
Any model of a socio-economic system (ordinary dynamic model, welfare optimizing model, differential game model, etc.) whose dynamic equations are autonomous, locally Lipshitzian, and which is uniformly asymptotically stable has an equilibrium statistical representation.

5. CONCLUSIONS

The equilibrium statistical representation of a dynamic system (ESR) is the probability distribution which maximizes entropy subject to an upper bound on Lyapunov function associated with an autonomous dynamic system. ESR has the following properties.

1. ESR is interpreted as probability that the neighbourhood of a given point in the state (configuration) space belongs to some trajectory of the dynamic system.

.2 The time average of an arbitrary observation made on a dynamic system can be treated as a stochastic process. If this process is strongly ergodic in the mean, i.e. vaguely speaking if the system forgets its past, then the ESR produces an expected value of the observation which is an unbiased estimator of the time average.

3. If the dynamic system is generated by a dynamic optimization model or a zero sum game against nature, the Lyapunov function (if it exists) can be interpreted as present cost/utility. The resulting ESR then satisfies the criterion of cost efficiency postulated by Smith. The ESR probability distribution thus bridges the gap between behavioural and descriptive models.

The existence of a Lyapunov function associated with a dynamic system implies that the system is stable and bounded. The Lyapunov function certainly exists if the dynamic equations satisfy the Converse Theorems of Stability (Yoshizawa, 1975) i.e. if the solutions are uniformly asymptatically stable and the dynamic equations locally Lipshitzian.

It follows that the microdynamic equations must have solutions which are at least stable (or bounded) if the equilibrium statistical representation should exist. Conversely, if a socio-economic system has a meaningful ESR, the stability/boundedness of the microdynamic solutions is implied. As a consequence, the theory establishes a link between the existence of a (macro-) equilibrium description and stability/boundedness at the micro level.

REFERENCES

Anas, A. (1982) Residential Location Markets and Urban Transportation Economic Theory, Econometrics, and Policy Analysis with Discrete Choice Models. Academic Press, New York.

Andersson, Å. and Persson, H. (1980) Integration of Transportation and Location Analysis: A General Equilibrium Approach, IIASA RR 80-40.

Bertuglia, C.S. and Leonardi, G. (1979) Dynamic Models for Spatial Interaction. Sistemi Urbani, Vol.2, pp.3-25.

Boyce, D.E. and Janson, B.N. (1980) A Discrete Transportation Network Design Problem with Combined Distribution and Assignment, Transp. Res. Vol.14B, pp.147-154.

Boyce, D.E., Le Blanc, L.J., Chon, K.S., Lee, Y.J. and Lin, K.T. (1981) Combined Models of Local, Destination, Mode and Route Choice: A Unified Approach Using Nested Entropy Constraints. Transportation Planning Group, University of Illinois, Urbana, Illinois.

Cass, D. and Shell, K. (1976) The Structure and Stability of Competitive Dynamical Systems. J. Econ. Theory Vol.12, pp.31-70.

Cochrane, R.A. (1975) A Possible Economic Basis for the Gravity Model. J. of Transp. Economics and Policy, Vol.9, No.1, pp.34-49.

Coelho, J.D. and Wilson, A.G. (1977) Some Equivalence Theorems to

Integrate Entropy Maximizing Submodels Within Overall Mathematical Programming Frameworks, Geographical Analysis, Vol.9, pp.160-173.

Erlander, S. (1977) Accessibility, Entropy and the Distribution and Assignment of Traffic. Transp. Res. Vol.11, pp.149-153.

Erlander, S. (1980) Optimal Spatial Interaction and the Gravity Model, Lecture Notes in Economics and Mathematical Systems, No.173, Springer, Berlin.

Erlander, S. (1982) On the Classical Problem of Equilibrium in Statistical Mechanics, Linköping University, Department of Mathematics, LiTH-MAT-R-82-50.

Friedman, A. (1971) Differential Games. Wiley Interscience, New York.

Gupta, N.K. (1981) An Overview of Differential Games. In: Control and Dynamic Systems, C.T. Leondes (Ed.). Academic Press, New York.

Hicks, J.R. (1939) Value and Capital, Oxford Clarendon Press.

Isaacs, R. (1965) Differential Games, Wiley , New York.

Isard, W., Liossatos, P., Kanemoto, Y. and Kaniss, P.C. (1979) Spatial Dynamics and Optimal Space-Time Development. North Holland, New York.

Isard, W. and Anselin, L. (1980) Multi-Region Programming Models: Linkages to and from. Journal of Regional Science. Symposium on Multiregional Forecasting and Policy Simulation Models, Vol.20, pp.129-

206.

Jaynes, E.T. (1957) Information Theory and Statistical Mechanics. Phys. Rev. Vol.106, pp.620-630.

Jaynes, E.T. (1957) Information Theory and Statistical Mechanics II. Phys. Rev. Vol.108, pp.171-190.

Jaynes, E.T. (1978) Where Do We Stand on Maximum Entropy? In: The Maximum Entropy Formalism, A Conference held at MIT, 1978, R.D. Levine, M. Tribus (Eds.), MIT Press, Cambridge, Mass., pp.15-118.

Katz, A. (1967) Principle of Statistical Mechanics (The Information Theory Approach). W.H. Freeman & Co., San Francisco, London.

Kohno, H., Yoshida, M. and Mitomo, H. (1981) Simultaneous Optimal Allocation of Public Investments to the Interurban Comprehensive Transport Systems and Regional Living-Environmental Facilities. Proc. 28th North American Meeting RSA November, Montreal.

Kohno, H. and Higano, Y. (1982) Improvement of the Quality of Life in Tokyo through the Total Reorganization and Strengthening of Economic Functions of the Japanese National Capital Region. Proc. 22 European Congress RSA, Groningen.

Lesse, P.F. (1982) A Phenomenological Theory of Socio-economic Systems with Spatial Interactions. Environment and Planning A, Vol.14, pp.869-888.

Melsa, J.L. and Sage, A.P. (1973) An Introduction to Probability and Stochastic Processes. Prentice Hall, Englewood Cliffs, N.J.

Papageorgiou, G.J. (1977) Fundamental Problems of Theoretical Planning. Environment and Planning A., Vol.9, 1329-1356.

Samuelson, P.A. (1948) Foundations of Economic Analysis. Cambridge, Harvard University Press.

Siljak, D. (1978) Large Scale Dynamic Systems. North Holland, New York, Amsterdam.

Skowronski, J.M. (1977) Lyapunov Type Design of Lumped Systems in Conflict with Environment. Department of Mathematics, University of Queensland, Control Theory Rep, pp.77-1.

Smith, T.E. (1978) A Cost Efficiency Principle of Spatial Interaction Behaviour, Reg. Sci and Urban Economics, Vol.8, pp.313-337.

Stalford, H. and Leitman, G. (1973) Sufficiency conditions for Nash Equilibria in N-person Differential Games in: Topics in Differential Games, A. Blaquiere (Ed.). North Holland, Amsterdam.

Stalford, H. and Leitman, G. (1971) Sufficient Conditions for optimality in two person zero sum differential games. J. Math. Anal. and Appl. 33, pp.650-654.

Varaiya, P. (1971) Differential Games with Dynamical Systems. In: Differential Games and Related Topics. H.W. Kuhn, G.P. Szegö (Eds.).

North Holland, Amsterdam.

Varian, H.E. (1981) Dynamical Systems with Applications to Economics. In: J.K. Arrow, and M.D. Intriligator (Eds.) Handbook of Mathematical Economics, Vol.I, North Holland, Amsterdam, pp.93-110.

Webber, M.J. (1980) Information Theory and Urban Spatial Structure. Croom Helm, London.

Williams, H.C.W. and Senior, M.L. (1977) Accessibility, Spatial Interaction and the Evaluation of Land-Use Transport Plans. Proc. International Research Conference on Spatial Interaction Theory and Models, Bastad, Sweden.

Wilson, A.G. (1970) Entropy in Urban and Regional Modelling. Pion, Ltd., London.

Wilson, A.G. (1974) Urban and Regional Models in Geography and Planning. Wiley and Sons, London.