"BAYES' THEOREM" FOR UTILITY

by

Leigh Tesfatsion

Discussion Paper No. 76-65, April 1976

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

ABSTRACT

An algorithm is proposed for updating an initial period objective (risk) function by means of transitional utility (loss) assessments, in a manner analogous to Bayes' theorem for probability. Specification of updated probability assessments is not required. The algorithm is shown to be robust with respect to a fully updated procedure, given certain empirically meaningful restrictions. Existence of the initial period objective function is axiomatized in the context of a model which adopts a symmetrical approach to utility and probability.

"BAYES' THEOREM" FOR UTILITY*

by

Leigh Tesfatsion

1. INTRODUCTION

The principal purpose of this paper is to propose a simple "utility algorithm" for updating an initial period objective (risk) function by means of transitional utility (loss) assessments, in a manner analogous to Bayes' theorem for probability.[1] Specification of updated probability assessments is not required. The algorithm is shown to be robust with respect to a fully updated procedure, given certain empirically meaningful restrictions.

The computation and information required for the numerical solution of Bayesian sequential decision procedures is often prohibitive (see DeGroot [2, Part Four]). Consequently, for sequential decision problems in which utility assessments are relatively straightforward and probability assessments are vague or of intractable form, a utility algorithm may be a desirable alternative to a Bayesian procedure.[2] Numerous interesting economic decision problems fit this description; for many important economic time series appear to be realizations for nonstationary stochastic processes, difficult to express analytically in any convincing manner, and intractable to work with even if so expressed. In contrast, utility assessments over outcomes can often be based on relatively noncontroversial revenue and cost considerations.[3]

---

*Research for this paper was supported by National Science Foundation Grant GS-31276X, University of Minnesota.

A seemingly essential prerequisite for devising general utility algorithms is that utility be interpreted as the proper twin of probability. Statistical studies have generally emphasized probability distributions, with utility (loss, cost...) assessments specified in abstract or routine (e.g. quadratic) form when they appear. Nevertheless, like probability, utility can be interpreted as a possibly conditioned distribution of "mass" over a collection of relevant events, however these are represented (e.g., sets, points, propositions...). Indeed, it has been cogently argued by de Finetti [1, 2.4] that in practice any distinction drawn between events to be represented as sets and events to be represented as points is purely a matter of convenience. Events, as distinct from sets, can always be subdivided by setting out their description in finer detail.[4]

The following concepts indicate the extent to which a meaningful analogy between utility and probability can be made.

utility density

$$u: \Omega \to R$$

where $R \equiv$ the real line and $u(\omega) \equiv$ the utility (loss, cost...) the decision maker associates with the realization of the event $\omega \in \Omega$.

conditioned utility densities

$$u(\cdot|\theta): \Omega \to R, \ \theta \in \Theta,$$

where $u(\omega|\theta) \equiv$ the utility (loss, cost...) the decision maker associates with the realization of event $\omega \in \Omega$, conditioned on the realization (choice, hypothesis, truth...) of θ.

nth period transitional utility density

$$u\,(\cdot|\omega_1,\ldots,\omega_{n-1}):\ \Omega_n \rightarrow R,$$

where the event space Ω_n relevant for period n is a function $\Omega(\omega_1,\ldots,\omega_{n-1})$ of the realized event stream $(\omega_1,\ldots,\omega_{n-1}) \in \Omega_1\times\ldots\times\Omega_{n-1}$, and $u(\omega_n|\omega_1,\ldots,\omega_{n-1}) \equiv$ the utility (loss, cost...) the decision maker associates with the realization of event $\omega_n \in \Omega_n$, conditioned on the realization of $(\omega_1,\ldots,\omega_{n-1})$.

Markov utility density of order 1

$$u(\omega_n|\omega_1,\ldots,\omega_{n-1}) = u(\omega_n|\omega_{n-1}):\ \Omega_n \rightarrow R;$$

i.e., the utility (loss, cost...) the decision maker associates with the realization of event ω_n in period n is independent of the events $(\omega_1,\ldots,\omega_{n-2})$ which he observed previous to period n-1. (One interpretation might be that habit formation is insignificant for the decision maker.) Generalization to other orders is obvious. In particular, a Markov utility density of order 0 might be termed a *stationary utility density*.

The utility algorithm proposed in this paper will require a period by period specification of conditioned transitional utility densities.

The paper is organized as follows. The algorithm is presented in section 4 and its properties are investigated in section 5. The definition of the algorithm and many of the properties established for the algorithm are independent of the exact form of the initial period objective function U. However, as shown in section 5, it is useful to assume that U has a general expected utility representation

$$U(\theta) = \int_{\Omega(\theta)} u(\omega|\theta)p(d\omega|\theta), \theta \varepsilon \Theta, \qquad (1.1)$$

where $\Theta = \{\theta, \ldots\}$ is a choice set consisting of "policies," and for each policy $\theta \varepsilon \Theta$, $u(\cdot|\theta)$ is a θ-conditioned utility density, $p(\cdot|\theta)$ is a θ-conditioned probability measure, and $\Omega(\theta)$ is a set of atomic events relevant for the initial period choice of θ. As shown in the preliminary sections 2 and 3, the representation (1.1) for U can be axiomatized in terms of primitive sets and relations.

The algorithm is derived recursively from a "Bayes' theorem for utility" of the form

$$U(\theta|\omega_1) \propto u(\omega_1|\theta) + U(\theta). \qquad (1.2)$$

Intuitively, (1.2) asserts that the "posterior" utility $U(\theta|\omega_1)$ which the decision maker associates with the choice of policy θ for period 2 after having observed event ω_1 in period 1 is proportional to the preobservation utility $u(\omega_1|\theta)$ he associated with the realization of event ω_1 in period 1, conditioned on the choice of policy θ for period 1, weighted by the "prior" (preobservation) utility $U(\theta)$ he associated with the choice of policy θ for period 1. For uniform (constant) priors U, (1.2) asserts that in period 2 the decision maker ought to choose the policy that would have given him maximum utility in period 1 for the realized ω_1 (a "maximum likelihood method" for utility - see section 5). As will be clarified in sections 4 and 5, for certain repetitive decision problems $U(\cdot|\omega_1)$ represents a reasonable choice of objective function for period 2.

For example, consider the following situation. A decision maker is about to introduce a perishable good into a market in the face of uncertain demand. Periodically he will have to order new stock θ

from a parent company. His realized utility $u(\omega_n|\theta_n)$ for period n is given by the difference $R_n(\min(\omega_n,\theta_n)) - C_n(\theta_n)$ between revenue $R_n(\min(\omega_n,\theta_n))$ corresponding to realized nth period sales $\min(\omega_n,\theta_n)$ and cost $C_n(\theta_n)$ corresponding to his nth period stock order θ_n. He has taken great pains to assess his first period objective function $\theta \mapsto U(\theta)$ incorporating his initial probability and utility assessments, but now he would appreciate having a reasonable, computationally feasible method for periodically updating this objective function as new events ω_n sequentially obtain.

2. THE PRIMITIVE POLICY MODEL[5]

Let $G = \{g,\ldots\}$ be a set of *candidate goals*, and for each $g \in G$ let $\Lambda_g = \{\lambda_g,\ldots\}$ be a set of *controls*. A *primitive policy model* is then characterized by the sets and relations

$$[(\Theta, \succcurlyeq), \{(\Omega(\theta), \succcurlyeq_\theta, \geq_\theta) \mid \theta \in \Theta\})]$$

where

$$\Theta = \{\theta,\ldots\} = \cup_{g\in G}\{(g,\lambda_g) \mid \lambda_g \in \Lambda_g\}$$

is the *policy choice set* consisting of candidate goal-control pairs (*policies*);

$\succcurlyeq$ (*policy preference order*) is a weak order[6] on Θ;

and for each policy $\theta \in \Theta$,

$\Omega(\theta) = \{\omega_\theta,\ldots\}$ is a nonempty set of *state flows* associated with the policy θ;

$\succcurlyeq_\theta$ (θ-*conditioned preference order*) is a weak order on $\Omega(\theta)$;

$\geq_\theta$ (θ-*conditioned probability order*) is a weak order on the algebra[7] $2^{\Omega(\theta)}$ comprising all subsets (*event flows*) $E_\theta \subseteq \Omega(\theta)$.

Remark: Axiomatic theories for the state-consequence-act models prevalent in economic and statistical decision theory generally assume the existence of only one primitive order, a preference order over acts, and hence obtain a simpler statement for their primitives. Nevertheless, the axioms used to justify an expected utility representation for the preference order then invariably impose strong nonnecessary restrictions on the primitives (e.g., L.J. Savage's reliance on constant acts in [5]). In contrast, assuming each state flow set $\Omega(\theta)$ is finite, necessary and sufficient conditions can be given which justify an expected utility representation for the policy preference order $\succcurlyeq$ (see section 3 below).

The controls may be interpreted as possibly conditioned sequences of actions (i.e., partial contingency plans) entirely under the decision maker's control at the time of his choice. The candidate goals $g \in G$ may be interpreted as potential objectives (e.g., production targets, intervals for quality control,...) whose realization the decision maker can attempt to achieve through appropriate choice of a control. The grouping of the controls into sets $\{\Lambda_g | g \in G\}$ reflects the possibility that different sets of controls may be relevant for different goals. A control $\lambda_g \in \Lambda_g$ may or may not provide for the communication of the goal g to other persons in the decision maker's problem environment.

The weak order $\succcurlyeq$ on Θ can be interpreted as a preference order as follows. For all $\theta, \theta' \varepsilon \Theta$,

$\theta \gtrsim \theta'$ <=> the choice of policy θ is at least as desirable to the decision maker at the choice of policy θ'.

It is assumed that the decision maker will choose a policy (candidate goal-control pair) $\theta^* \varepsilon \Theta$ which is optimal in the sense that $\theta^* \gtrsim \theta$ for all $\theta \varepsilon \Theta$. Throughout this paper "choose policy $\theta = (g, \lambda_g)$" and "implement control λ_g with g as the objective" are used interchangeably.

For each $\theta \varepsilon \Theta$, the set $\Omega(\theta)$ of state flows ω_θ can be interpreted as the decision maker's answer to the following question: "If I choose policy θ, what distinct situations (i.e., state flows ω_θ) might obtain?" The state flows may include references to past, present, and future happenings. In particular, they can subsume both the "states" and the "consequences" of traditional state-consequence-act models.

The θ-conditioned preference orders $\gtrsim_\theta$ can be interpreted as follows. For all $\omega, \omega' \varepsilon \Omega(\theta), \theta \varepsilon \Theta$,

$\omega \gtrsim_\theta \omega'$ <=> the realization of ω is at least as desirable to the decision maker as the realization of ω', given that he chooses θ.

Similarly, the θ-conditioned probability orders $\geq_\theta$ can be interpreted as follows. For all E, E' $\varepsilon\ 2^{\Omega(\theta)}$, $\theta \varepsilon \Theta$,

$E \geq_\theta E'$ <=> in the decision maker's judgment the realization of E is at least as likely as the realization of E', given that he chooses θ.

As usual, all preference and probability orders depend (here implicitly) on the decision maker's current information state.

A state flow ω may be relevant for the decision maker's problem under distinct potential policy choices; e.g., $\omega \in \Omega(\theta) \cap \Omega(\theta')$ for some $\theta, \theta' \in \Theta$. Given state flows $\omega, \omega' \in \Omega(\theta) \cap \Omega(\theta')$ for some $\theta, \theta' \in \Theta$, it may hold that $\omega \succcurlyeq_{\theta} \omega'$ whereas $\omega' \succcurlyeq_{\theta'} \omega$. Verbally, the relative desirability of the state flows ω and ω' may depend on which conditioning event the decision maker is considering, "I choose θ" or "I choose θ'." Similarly for the relative likelihood of event flows.

Examples illustrating these interpretations are given in Tesfatsion [6] and [8].

3. OBJECTIVE FUNCTION REPRESENTATION

In Tesfatsion [7] necessary and sufficient conditions are given which ensure that a primitive policy model with finite state flow sets $\Omega(\theta)$ has an expected utility (objective function) representation in the following sense. To each policy $\theta \in \Theta$ there corresponds a finitely additive probability measure $p(\cdot|\theta):2^{\Omega(\theta)} \rightarrow [0,1]$ satisfying

$$p(E|\theta) \geq p(E'|\theta) \iff E \geq_{\theta} E' , \qquad (3.1)$$

for all $E, E' \in 2^{\Omega(\theta)}$, and a utility function $u(\cdot|\theta):\Omega(\theta) \rightarrow R$ satisfying

$$u(\omega|\theta) \geq u(\omega'|\theta) \iff \omega \succcurlyeq_{\theta} \omega' , \qquad (3.2)$$

for all $\omega, \omega' \in \Omega(\theta)$, such that

$$\int_{\Omega(\theta)} u(\omega|\theta)\ p(d\omega|\theta) \geq \int_{\Omega(\theta')} u(\omega|\theta')\ p(d\omega|\theta') \iff \theta \succcurlyeq \theta' , \qquad (3.3)$$

for all $\theta, \theta' \in \Theta$.

The representation (3.1)-(3.3) for the preference order $\succcurlyeq$ can be interpreted as follows. To each state flow $\omega \in \Omega(\theta), \theta \in \Theta$, the decision maker assigns a number $u(\omega|\theta)$ representing the desirability of ω, given he chooses θ, and a number $p(\omega|\theta)$ representing the likelihood of ω, given he chooses θ. He then calculates the expected utility

$$\int_{\Omega(\theta)} u(\omega|\theta)\ p(d\omega|\theta)$$

corresponding to each choice of policy $\theta \varepsilon \Theta$, and chooses a policy which yields maximum expected utility.

Definition: A primitive policy model with representation as in (3.1)-(3.3) will be referred to as a *policy model* characterized by a vector

$$(\Theta, \{u(\cdot|\theta):\Omega(\theta)\to R|\theta\ \varepsilon\ \Theta\},\ \{p(\cdot|\theta):\ 2^{\Omega(\theta)}\to [0,1]|\theta\ \varepsilon\Theta\}),$$

with *objective function* $U:\Theta \to R$ given by

$$U(\theta) = \int_{\Omega(\theta)} u(\omega|\theta)\ p(d\omega|\theta),\quad \theta\ \varepsilon\Theta.$$

The principal distinctive feature of the policy model is the symmetrical treatment of utility and probability assessments. Both are conditioned on policy choices, and both are defined over "state flows" rather than having utility defined over a set of "consequences" and probability defined over a set of "states."

The conditioning of the utility functions $\{u(\cdot|\theta):\Omega(\theta)\to R|\theta\varepsilon\Theta\}$ and the probability measures $\{p(\cdot|\theta):2^{\Omega(\theta)} \to [0,1]|\theta\varepsilon\Theta\}$ on the policies θ is often essential; for, as illustrated by examples in Tesfatsion [6] and [8], a state flow may have a different utility and probability depending on which policy is chosen. By subsuming states and consequences into state flows and defining utility over state flows rather than consequences alone, the practical difficulty of specifying utility-free states can be avoided. More important, when probability is defined over state flows subsuming both states and consequences, it is no longer necessary for the decision maker to specify his available actions in the form of complete contingency plans (e.g., Savage acts, von Neumann strategies, Wald decision functions...) mapping states into consequences. For this reason the policy model appears to be more suitable than traditional state-consequence-act models for modeling certain limited information decision problems.

As will be discussed in sections 4 and 5, an additional benefit to be gained from the symmetrical treatment of utility and probability is the possibility of updating the initial objective function through transitional utility assessments in a manner analogous to Bayes' theorem for probability.

4. THE ALGORITHM

Let s_1 represent the initial information state of a decision maker facing a policy selection problem at some time t_1. Suppose his problem has the following policy model form (cf. section 3):

$$\Gamma_1 \equiv (\Theta_1, \{u_1(\cdot|\theta): \Omega_1(\theta) \to R \,|\, \theta \varepsilon \Theta_1\}, \{p_1(\cdot|\theta): 2^{\Omega_1(\theta)} \to [0,1] \,|\, \theta \varepsilon \Theta_1\})$$

with objective function

$$U_1(\theta) \equiv \int_{\Omega_1(\theta)} u_1(\omega|\theta) p_1(d\omega|\theta), \quad \theta \varepsilon \Theta_1,$$

where the policy choice set Θ_1 and state flow sets $\Omega_1(\theta)$ are functions of his information state s_1, and $u_1(\cdot|\theta)$, $p_1(\cdot|\theta)$, and $U_1(\theta)$ are abbreviations for $u(\cdot|s_1,\theta)$, $p(\cdot|s_1,\theta)$, and $U(\theta|s_1)$ respectively.[8]

Presumably, if Γ_1 is well-defined, the decision maker at some future time t_2 (uncertain as of time t_1) will have implemented the policy θ_1^* chosen at time t_1 and verified the realization of one (and only one) state flow $\omega_1^* \in \Omega_1(\theta_1^*)$. His new information state at time t_2 will then be $s_2 \equiv (s_1, \theta_1^*, \omega_1^*)$. A subsequent policy selection problem might then have the policy model form

$$\Gamma_2 \equiv (\Theta_2, \{u_2(\cdot|\theta):\Omega_2(\theta)\to R|\theta\in\Theta_2\}, \{p_2(\cdot|\theta):2^{\Omega_2(\theta)}\to[0,1]|\theta\in\Theta_2\})$$

with updated objective function

$$U_2(\theta) \equiv \int_{\Omega_2(\theta)} u_2(\omega|\theta)\, p_2(d\omega|\theta),\ \theta \in \Theta_2, \tag{4.1}$$

where the policy choice set Θ_2 and state flow sets $\Omega_2(\theta)$ are functions of his new information state s_2, and $u_2(\cdot|\theta)$, $p_2(\cdot|\theta)$, and $U_2(\theta)$ are abbreviations for $u(\cdot|s_2,\theta)$, $p(\cdot|s_2,\theta)$, and $U(\theta|s_2)$ respectively.

If the policy choice set Θ_2 or the state flow sets $\Omega_2(\theta)$ contain numerous alternatives, it may not be feasible or desirable for the decision maker to carry out all the updating required by Γ_2. If $\Theta_1 = \Theta_2 \equiv \Theta$ and one state flow set Ω is relevant for all policies in both periods, the decision maker could replace (4.1) by the updated objective function

$$U_1(\theta|\omega_1^*) \propto u_1(\omega_1^*|\theta) + U_1(\theta),\ \theta \in \Theta. \tag{4.2}$$

Intuitively, (4.2) asserts that the posterior utility $U_1(\theta|\omega_1^*)$ which the decision maker associates with the choice of policy θ for period 2 following the realization of ω_1^* in period 1 is proportional to the preobservation utility $u_1(\omega_1^*|\theta)$ he associated with the realization of ω_1^* in period 1, conditioned on the choice of policy θ for period 1, weighted by the prior (preobservation) utility $U_1(\theta)$ he associated with the choice of policy θ for period 1. Policies θ which score better with respect to the realized state flow ω_1^* than their anticipated "average" $U_1(\theta)$ have their relative utility value raised for period 2; conversely for those policies which score lower than anticipated. For repetitive decision problems in which the state flows are judged by the decision maker to be approximately stochastically independent of his policy choices, i.e., for all $\omega \varepsilon \Omega$ and $\theta \varepsilon \Theta$,

$$p_2(\omega|\theta) \approx p(\omega|s_1,\omega_1^*),$$

and the costs incurred by switching policies from period to period are negligible, $U_1(\cdot|\omega_1^*)$ may represent a reasonable choice of objective function for period 2.[9]

By taking exponents in (4.2), an exact analogy with Bayes' theorem for probability is obtained; namely,

$$V_1(\theta|\omega_1^*) \propto v_1(\omega_1^*|\theta)\ V_1(\theta), \tag{4.3}$$

where $V_1(\theta) \equiv \exp(U_1(\theta))$, $v_1(\omega_1^*|\theta) \equiv \exp(u_1(\omega_1^*|\theta))$, and $V_1(\theta|\omega_1^*) \equiv \exp\ (U_1(\theta|\omega_1^*))$. Since $v_1(\omega_1^*|\theta)\ V_1(\theta)$ and $\log(v_1(\omega_1^*|\theta)\ V_1(\theta)) \equiv u_1(\omega_1^*|\theta) + U_1(\theta)$ achieve their maxima (if any) over the same set of policies, (4.2) and (4.3) yield the same set of policy recommendations. The algorithm will be derived by recursion from the multiplicative relation (4.3) in preference to (4.2) in order to take direct advantage of results previously obtained by researchers investigating Bayes' theorem for probability.

The algorithm will now be formally presented.

4.1 Assumptions and Notation

At some time t_1 a decision maker in a certain information state s_1 faces a decision problem requiring periodic policy choices $\theta_1^*,\theta_2^*,\ldots$. A single policy choice set $\Theta=\Theta\ (s_1)$ is relevant for each period $[t_n,t_{n+1}]$, $n \geq 1$, and a single state flow set $\Omega=\Omega(s_1)$ is relevant for all policies in all periods. In addition, the cost of switching policies from period to period is negligible; i.e., letting ω_j^* denote the state flow which obtains in period j, the decision maker's policy-conditioned transitional utility densities $\{u_n(\cdot|\theta):\Omega\to R|\theta \in \Theta\}$ for period n satisfy

$$u_n(\omega|\theta) \equiv u(\omega|s_1,\theta_1^*,\omega_1^*,\ldots,\theta_{n-1}^*,\omega_{n-1}^*,\theta)$$

$$= u(\omega|s_1,\omega_1^*,\ldots,\omega_{n-1}^*,\theta),$$

for all $\omega \in \Omega$ and $\theta \in \Theta$. Finally, the decision maker is able to give initial period policy-conditioned probability assessments $\{p_1(\cdot|\theta): 2^{\Omega}\to[0,1]|\theta \in \Theta\}$ for the likelihood of the event flows $E \subseteq 2^{\Omega}$.

For normalization purposes it will be assumed that for some positive measure m over Θ all integrals

$$\int_{\Theta} f(\theta)m(d\theta)$$

appearing in the algorithm are well-defined and satisfy

$$\int_{\Theta}|f(\theta)|m(d\theta) < \infty.$$

For example, if Θ is finite, m could denote counting measure.

4.2 The Algorithm

Step o: Calculate the initial period objective function $U_1:\Theta\to R$, given by

$$U_1(\theta) = \int_{\Omega} u_1(\omega|\theta)p_1(d\omega|\theta),\ \theta \in \Theta;$$

and select a policy $\theta_1^v \in \Theta$ for period 1 which satisfies

$$V_1(\theta_1^v) \geqq V_1(\theta),$$

for all $\theta \in \Theta$, where

$$V_1 \equiv \exp(U_1).$$

Step n ($n \geqq 1$): Verify a state flow $\omega_n^* \in \Omega$ in period n; calculate the updated objective function $V_{n+1}:\Theta\to R$ for period n+1, given by

$$V_{n+1}(\theta) \equiv \frac{v_n\ (\omega_n^*|\theta)\ V_n(\theta)}{v_n(\omega_n^*)} \quad , \ \theta \ \varepsilon\ \Theta\ , \tag{4.4}$$

where

$$v_n(\omega_n^*|\theta) \equiv \exp\ (u_n(\omega_n^*|\theta\ )),$$

$$v_n(\omega_n^*) \equiv \int_\Theta v_n(\omega_n^*(\theta))\ V_n(\theta)m(d\theta);$$

and select a policy $\theta_{n+1}^v\ \varepsilon\ \Theta$ for period n+1 which satisfies

$$V_{n+1}(\theta_{n+1}^v) \geq V_{n+1}(\theta),\ \text{for all}\ \theta\ \varepsilon\,\Theta\ .$$

Remark: All factors in (4.4) are strictly positive. The normalizing factors $v_n(\omega_n^*)$ guarantee that

$$\int_\Theta V_n(\theta)m(d\theta) = 1, \quad n \geq 2.$$

Since U_1 is only unique up to positive linear transformation under the axiomatization discussed in section 3, $V_1 \equiv \exp(U_1)$ can be similarly normalized.

5. ROBUSTNESS AND SMALL SAMPLE PROPERTIES

Certain "small sample" properties of the algorithm will first be established, followed by a number of robustness results. Although many of the properties are formally analogous to properties previously established for probability densities, their interpretation in terms of utility densities is equally natural and meaningful. The robustness results have no direct analogy in probability theory.

The assumptions, notation, and definitions of 4.1 and 4.2 will be maintained throughout this section, except where otherwise indicated. The existence of a particular sequence $\omega_1^*, \omega_2^* \ldots$ of state flows "realized" over periods 1, 2, ... will be implicitly assumed in the statement and proof of all theorems. The symbol K_n will denote a general nth period constant, i.e., a term independent of policy choices and state flows to be realized in periods $m \geq n$. The exact form of K_n may vary during a proof.

5.1 Small Sample Properties

The first theorem below guarantees that the algorithm is essentially invariant under similar positive linear transformation of the utility densities.

Theorem 5.1: For each $n \geq 1$, the set of maximizing policies for the objective function V_n is invariant under similar positive linear transformation of the utility densities $\{u_j(\cdot|\theta):\Omega \to R | 1 \leq j \leq n, \theta \varepsilon \Theta\}$.

Remark: The sets of maximizing policies for the V_n are not necessarily invariant under dissimilar positive linear transformations of the utility densities. The practical implication is that the utility densities u_n for each period n should be scored in relation to a fixed origin and unit determined in the initial period.

Proof: Let $n \geq 1$ be given. For each $j, 1 \leq j \leq n$, let $u_j^* \equiv au_j + b$ for some a, b ε R, a>0, and let V_n^* denote the nth period algorithm objective function calculated in terms of the densities u_j^*, $1 \leq j \leq n$. By assumption,

$$U_1(\theta) = \int_\Omega u_1(\omega|\theta)p_1(d\omega|\theta), \theta \varepsilon \Theta.$$

Hence

$$aU_1(\theta) + b = \int_\Omega u_1^*(\omega|\theta)p_1(d\omega|\theta) \equiv U_1^*(\theta), \theta \varepsilon \Theta,$$

and

$$\log V_1^* \equiv \log(\exp(U_1^*)) = a\log V_1 + b. \tag{5.1}$$

If $n \geq 2$, then for all $\theta \in \Theta$,

$$\begin{aligned}\log V_n^*(\theta) &= \log (v_{n-1}^*(\omega_{n-1}^*|\theta)V_{n-1}^*(\theta)) + K_n \\ &= \log(v_{n-1}^*(\omega_{n-1}^*|\theta)v_{n-2}^*(\omega_{n-2}^*|\theta)\ldots v_1^*(\omega_1^*|\theta)V_1^*(\theta))+K_n \\ &= \Sigma_{j=1}^{n-1}u_j^*(\omega_j^*|\theta) + U_1^*(\theta) + K_n \\ &= a(\Sigma_{j=1}^{n-1} u_j(\omega_j^*|\theta) + U_1(\theta)) + nb + K_n \\ &= a \log V_n(\theta) + K_n. \end{aligned} \tag{5.2}$$

It follows immediately from (5.1) and (5.2) that V_n^* and V_n have the same (possibly empty) set of maximizing policies.

Definitions and Notation: For each $n \geq 2$, let the *preference function* $D_n:\Theta \to R$ be defined by

$$D_n(\theta) \equiv \Pi_{j=1}^{n-1} v_j(\omega_j^*|\theta), \ \theta \in \Theta,$$

where $\omega_1^*,\ldots,\omega_{n-1}^*$ are the state flows realized over periods $1,\ldots n-1$. Thus

$$V_n(\theta) = K_n D_n(\theta) \ V_1(\theta), \ \theta \in \Theta.$$

The (possibly empty) set M_n^{mp} of policies $\theta_n^{mp} \in \Theta$ which maximize D_n will be called the *set of maximum preference policies for period n*. The (possibly empty) set of policies $\theta_n^V \in \Theta$ which maximize V_n will be denoted by M_n^V.

Remark: The preference function is clearly analogous to the likelihood function of probability theory, with maximum preference policies corresponding to maximum likelihood estimators.

For "smooth" initial period objective functions U_1 (equivalently, $V_1 \equiv \exp(U_1)$), the algorithm objective function $V_n = K_n D_n V_1$ can be replaced by the preference function D_n with little distortion. For uniform (constant) initial objective functions,[10]

$$M_n^v = M_n^{mp} \;, \; n \geq 2.$$

In addition, for various types of utility densities $\{u_n\}$ it can explicitly be shown that

$$\theta_n^v - \theta_n^{mp} \to 0 \;; \qquad (5.3)$$

i.e., the influence of the initial period objective function $V_1 \equiv \exp(U_1)$ on the selection of policies by means of V_n becomes negligible as n increases. The corollary to the following theorem establishes (5.3) for utility densities having a commonly used quadratic representation. A small sample normal distribution for V_n is simultaneously obtained.

Theorem 5.2: Suppose $\Omega \subseteq \Theta = R$, and for each n, $1 \leq n \leq n^*$, there exist constants k_n and d_n, $d_n > 0$, such that

$$u_n(\omega|\theta) = k_n - d_n[\omega-\theta]^2, \; \omega \in \Omega, \; \theta \in \Theta \;;$$

i.e., all utility densities u_n, $1 \leq n \leq n^*$, are quadratic. Then for each n, $1 \leq n \leq n^*$, the preference function $D_{n+1}: \Theta \to R$ satisfies

$$\exp(K_{n+1}) D_{n+1} \sim N(\bar{\omega}_n, \; 1/\sqrt{2\bar{d}_n}), \qquad (5.4)$$

with a unique maximum preference policy

$$\theta_{n+1}^{mp} = \bar{\omega}_n, \qquad (5.5)$$

where $\bar{d}_n \equiv \Sigma_{j=1}^n d_j$ and $\bar{\omega}_n \equiv \Sigma_{j=1}^n d_j w_j^* / \bar{d}_n \equiv$ the "weighted sample mean" of the realized state flows $\omega_1^*, \ldots, \omega_n^*$. In particular, for $d_1 = \ldots = d_n$,

$$\theta_{n+1}^{mp} = \Sigma_{j=1}^n \; \omega_j^* \; / \; n.$$

Remark: For each n, the nth period constants k_n and d_n may depend on the state flows $\omega_1^*, \ldots, \omega_n^*$ realized over periods 1,...,n. The utility densities may thus reflect fundamental changes in preferences over time.

Proof: Let $\theta \in \Theta$ and n, $1 \leq n \leq n^*$, be given. Then

$$\Sigma_{j=1}^{n} u_j(\omega_j^* | \theta) = K_{n+1} - \Sigma_{j=1}^{n} d_j [\omega_j^* - \theta]^2$$

$$= K_{n+1} - \Sigma d_j [\omega_j^{*2} - 2\omega_j^* \theta + \theta^2]$$

$$= K_{n+1} + 2\theta \Sigma d_j \omega_j^* - \bar{d}_n \theta^2$$

$$= K_{n+1} + \bar{d}_n [2\theta\ \bar{\omega}_n - \theta^2]$$

$$= K_{n+1} - \bar{d}_n [\bar{\omega}_n - \theta]^2 .$$

Thus

$$\exp(K_{n+1})\ D_{n+1}(\cdot) = \exp(K_{n+1})\ \exp(\Sigma_{j=1}^{n} u_j(\omega_j^* | \cdot))$$

$$\sim N(\omega_n, 1/\sqrt{2\bar{d}_n}),$$

with

$$\theta_{n+1}^{mp} = \bar{\omega}_n .$$

Corollary 5.3: Suppose in addition to the assumptions of Theorem 5.2 that Ω has finite cardinality N, and the initial objective function U_1 is given by

$$U_1(\theta) = \Sigma_{\Omega}\ u_1(\omega|\theta)\ p_1(\omega),\ \theta \in \Theta. \quad (5.6)$$

Then for each n, $1 \leq n \leq n^*$, the algorithm objective function $V_{n+1}: \Theta \to R$ for period n+1 satisfies

$$\exp(K_{n+1})\ V_{n+1} \sim N(m_n, \sigma_n),$$

with unique maximizing policy

$$\theta_{n+1}^{v} = m_n ,$$

where

$$m_n \equiv [\bar{d}_n\bar{\omega}_n + d_1\bar{\omega}]/[\bar{d}_n+d_1],$$

$$\sigma_n^{-2} \equiv 2[\bar{d}_n+d_1],$$

and

$$\bar{\omega} \equiv \Sigma_\Omega \omega p_1(\omega),$$

the "prior mean" of $\omega \varepsilon \Omega$. Moreover,

$$\theta_n^v - \theta_n^{mp} \to 0 \text{ as } n \to \infty.$$

Remark: Condition (5.6) implies that the decision maker is initially certain the realization of a state flow ω in period 1 is stochastically independent of his policy choice; i.e., $p_1(\omega|\theta) = p_1(\omega)$ for all $\omega\varepsilon\Omega, \theta\varepsilon\Theta$.

Proof: Let $\Omega \equiv \{\omega^1,\ldots,\omega^N\}$. As in the proof for Theorem 5.2, it is easily calculated that

$$\begin{aligned} U_1(\theta) &= \Sigma_{s=1}^N u_1(\omega^s|\theta)p_1(\omega^s) \\ &= K_1 - d_1\Sigma[\omega^s-\theta]^2 \; p_1(\omega^s) \\ &= K_1 - \Sigma d_s'[\omega^s-\theta]^2 \\ &= K_1 - d'[\bar{\omega}-\theta]^2, \end{aligned}$$

where $d_s' \equiv d_1 p_1(\omega^s)$, $1\leq s\leq N$, $d' \equiv \Sigma_{s=1}^N d_s' = \Sigma_{s=1}^N d_1 p_1(\omega^s) = d_1$, and $\bar{\omega} \equiv \Sigma_{s=1}^N d_s'\omega^s/d' = \Sigma_{s=1}^N \omega^s p_1(\omega^s)$. Thus

$$\begin{aligned} \exp(K_1)V_1 &\equiv \exp(K_1)\exp(U_1) \\ &\sim N(\bar{\omega}, 1/\sqrt{2d_1}). \end{aligned} \qquad (5.7)$$

Combining (5.7) and (5.4),

$$\exp(K_{n+1})\, V_{n+1}(\theta) \equiv \exp(K_{n+1})\, D_{n+1}(\theta)\, V_1(\theta)$$

$$= \exp(K_{n+1})\, \exp(-\bar{d}_n[\bar{\omega}_n-\theta]^2 - d_1[\bar{\omega}-\theta]^2)$$

$$= \exp(K_{n+1})\, \exp(-[2\sigma_n{}^2]^{-1}[m_n-\theta]^2)$$

$$\sim N(m_n,\sigma_n),$$

where

$$m_n = [\bar{d}_n\bar{\omega}_n + d_1\bar{\omega}] \;/\; [\bar{d}_n + d_1];$$

$$\sigma_n^{-2} = 2[\bar{d}_n + d_1].$$

The unique maximizing policy θ^v_{n+1} for V_{n+1} is therefore

$$\theta^v_{n+1} = m_n. \tag{5.8}$$

Combining (5.8) and (5.5),

$$\theta^v_{n+1} - \theta^{mp}_{n+1} = m_n - \bar{\omega}_n$$

$$= \bar{d}_n\ \bar{\omega}_n/[\bar{d}_n+d_1] - \bar{\omega}_n + d_1\bar{\omega}/[\bar{d}_n+d_1]$$

$$\to 0 \text{ as } n \to \infty.$$

Definitions: A function $t_n:\Pi_n\Omega\to S$, S an arbitrary set, will be called a *sufficient statistic for the utility densities* $\{u_n(\cdot|\theta): \Omega\to R|\theta\varepsilon\,\Theta\}$ if there exist functions f_n and g_n such that

$$u_n(\omega|\theta) = f_n(t_n(\overset{*}{\omega}_1,\ldots,\overset{*}{\omega}_{n-1},\omega),\theta) + g_n(\overset{*}{\omega}_1,\ldots,\overset{*}{\omega}_{n-1},\omega)$$

for all $\omega \in \Omega$ and $\theta \in \Theta$. A function $T_n:\Pi_n\Omega\to S$, S an arbitrary set, will be called a *sufficient statistic for* V_{n+1} if there exists a function H_n such that

$$V_{n+1}(\theta) = H_n(T_n(\overset{*}{\omega}_1,\ldots,\overset{*}{\omega}_n),\theta),\ \theta \in \Theta.$$

Theorem 5.4: Suppose for each n, $1 \le n \le n^*$, there exists a sufficient statistic t_n for the utility densities $\{u_n(\cdot|\theta):\Omega \to R | \theta \varepsilon \Theta\}$. Then for each n, $1 \le n \le n^*$,

$$V_{n+1}(\theta) = H_n(t_1,\ldots,t_n,\theta), \theta \varepsilon \Theta;$$

i.e., the algorithm objective function V_{n+1} for period n+1 depends on the realized state flows $\omega_1^*,\ldots,\omega_n^*$ only through the sufficient statistic $T_n \equiv (t_1,\ldots,t_n)$.

Proof: By assumption, for $1 \le n \le n^*$,

$$D_{n+1}(\theta) = \exp(\Sigma_{j=1}^n f_j(t_j(\omega_1^*,\ldots,\omega_j^*),\theta) + \Sigma_{j=1}^n g_j(\omega_1^*,\ldots\omega_j^*))$$

$$\equiv F_n(t_1(\omega_1^*),\ldots,t_n(\omega_1^*,\ldots,\omega_n^*),\theta)G_n(\omega_1^*,\ldots,\omega_n^*).$$

Thus

$$V_{n+1}(\theta) \equiv v_n(\omega_n^*|\theta) \; V_n(\theta) \; / \; v_n(\omega_n^*)$$

$$= D_{n+1}(\theta) \; V_1(\theta) \; / \int_\Theta D_{n+1}(\theta) \; V_1(\theta)m(d\theta)$$

$$= F_n(t_1,\ldots,t_n,\theta) \; V_1(\theta) \; / \int_\Theta F_n(t_1,\ldots,t_n,\theta) \; V_1(\theta)m(d\theta)$$

$$\equiv H_n(t_1,\ldots,t_n,\theta).$$

Theorem 5.5: Suppose for each n, $1 \le n \le n^*$, there exist functions $h_{sn}:\Omega \to R$, $y_n:\Omega \to R$, $b_n:\Theta \to R$, and $c_s:\Theta \to R$, $1 \le s \le r$, such that

$$u_n(\omega|\theta) = \Sigma_{s=1}^r \; c_s(\theta)h_{sn}(\omega) + y_n(\omega) + b_n(\theta). \qquad (5.9)$$

Then for each $\theta \varepsilon \Theta$ and each n, $1 \le n \le n^*$,

$$V_{n+1}(\theta) = H_n(T_n(\omega_1^*,\ldots,\omega_n^*),\theta),$$

where

$$T_n(\omega_1^*,\ldots,\omega_n^*) \equiv (\Sigma_{j=1}^n h_{1j}(\omega_j^*),\ldots,\Sigma_{j=1}^n h_{rj}(\omega_j^*)),$$

a sufficient statistic of fixed dimension r.

Remark: Quadratic utility $u_n(\omega|\theta) = k_n - d_n[\omega-\theta]^2$ (cf. Theorem 5.2) is a special case of (5.9) with r=2. Take $(c_1(\theta), c_2(\theta)) = (2\theta, 1)$, $(h_{1n}(\omega), h_{2n}(\omega)) = (d_n\omega, -d_n\omega^2)$, $y_n(\omega) = k_n$, and $b_n(\theta) = -d_n\theta^2$.

Proof: By assumption, $D_{n+1}(\theta) = \exp(\Sigma_{j=1}^n u_j(\omega_j^*|\theta)) = \exp(\Sigma_{j=1}^n y_j(\omega_j^*))F(T_n(\omega_1^*,\ldots,\omega_n^*),\theta)$, where $F(T_n,\theta) \equiv \exp(\Sigma_{j=1}^n b_j(\theta))\exp(\Sigma_{s=1}^r[\Sigma_{j=1}^n h_{sj}(\omega_j^*)]c_s(\theta))$; And $v_n(\omega_n^*) \equiv \int_\Theta D_{n+1}(\theta)V_1(\theta)m(d\theta) = \exp(\Sigma_{j=1}^n y_j(\omega_j^*))\int_\Theta F(T_n,\theta)V_1(\theta)m(d\theta)$.

Hence

$$\begin{aligned} V_{n+1}(\theta) &= D_{n+1}(\theta)\ V_1(\theta)/v_n(\omega_n^*) \\ &= F(T_n,\theta)\ V_1(\theta)/\int_\Theta F(T_n,\theta)\ V_1(\theta)m(d\theta) \\ &\equiv H_n(T_n,\theta). \end{aligned}$$

5.2 Robustness

The algorithm objective functions $\{V_n|n\geq 1\}$ require only the prior U_1 and the transitional utility densities $\{u_n\}$. Consequently they require a lower information state than the "fully updated" policy model objective functions $\{U_n|n\geq 1\}$ given by

$$U_n(\theta) = \int_\Omega u_n(\omega|\theta)p_n(d\omega|\theta),\ \theta \in \Theta,$$

where the functions $\{u_n\}$ are transitional utility densities as in the definition of the algorithm objective functions $\{V_n\}$, and the functions $\{p_n\}$ are transitional probability densities (see section 4).

Generally the state flows $\omega_1^*, \omega_2^* \ldots$ realized under the algorithm would differ from the state flows $\omega_1^u, \omega_2^u \ldots$ realized under the objective functions $\{U_n\}$. However, if the state flows ω are stochastically independent of present and past policy choices, i.e., for all $\omega\varepsilon\Omega, \theta\varepsilon\Theta, n\geq 1$,

$$p_n(\omega|\theta) = p(\omega|s_1,\omega_1^u,\ldots,\omega_{n-1}^u),$$

then presumably it would hold that

$$(\omega_1^*,\omega_2^*,\ldots) = (\omega_1^u,\omega_2^u,\ldots).$$

In this case maximizing policy sets $M_n^v = \{\theta_n^v,\ldots\}$ and $M_n^u = \{\theta_n^u,\ldots\}$ for V_n and U_n can be simultaneously derived; and the relative robustness of the algorithm can be measured by the differences

$$U_n(\theta_n^u) - U_n(\theta_n^v), \; n\geq 1, \tag{5.10}$$

between "true" maximum expected utility for period n and "true" expected utility corresponding to the policy choice θ_n^v selected to maximize the algorithm objective function $V_n, n\geq 1$.

In Theorem 5.7 below it will be shown that for a given sequence $\omega_1^*,\omega_2^*\ldots$ of "realized" state flows the differences in (5.10) are asymptotically negligible if the transitional utility assessments $\{u_n\}$ eventually stabilize and the transitional probability assessments $\{p_n\}$ asymptotically approximate empirical frequencies. In Theorem 5.10 below it will be shown that V_n asymptotically selects out the maximizing policy for U_n under essentially the same conditions. Theorems 5.7 and 5.10 therefore establish empirically meaningful conditions for the robustness of the algorithm relative to the fully updated policy model objective functions $\{U_n\}$, assuming state flows are stochastically independent of policy choices.

Notation: Let $\omega_1^*,\omega_2^*\ldots$ be a sequence of state flows realized over periods 1,2... For any $\omega\varepsilon\Omega$ and $n\geq 1$, define

$$n(\omega) \equiv \Sigma_{j=1}^{n} 1_{\omega}(\omega_j^*),$$

where

$$1_{\omega}(\omega_j^*) = \begin{cases} 1 & \text{if } \omega=\omega_j^* \ ; \\ 0 & \text{if } \omega\neq\omega_j^* \ . \end{cases}$$

Thus $n(\omega)/n$ denotes the frequency with which ω appears in the first n realizations $\omega_1^*,\ldots,\omega_n^*$.

The short-hand notation

$$f_n(A_n) \sim g_n(B_n)$$

will be used for

$$\lim_{n\to\infty} [\sup\{|f_n(a)-g_n(b)|:a\varepsilon A_n, b\varepsilon B_n\}] = 0.$$

For each $n\geq 1$, the (possibly empty) set of policies $\theta_n^u \ \varepsilon\ \Theta$ which maximize U_n will be denoted by M_n^u.

Remark 5.6: For any strictly positive real-valued function f, log(f) and f have the same (possibly empty) set of maximizing elements. Thus for every $n\geq 2$,

$$\begin{aligned} M_n^V &\equiv \{\theta' \ \varepsilon\ \Theta \mid V_n(\theta') = \max_{\theta\varepsilon\Theta} V_n(\theta)\} \\ &= \{\theta' \ \varepsilon\Theta \mid \tfrac{1}{n-1}\log V_n(\theta') = \max_{\theta\varepsilon\Theta} \tfrac{1}{n-1}\log V_n(\theta)\}; \\ M_n^{mp} &\equiv \{\theta' \ \varepsilon\Theta \mid D_n(\theta') = \max_{\theta\varepsilon\Theta} D_n(\theta)\} \\ &= \{\theta' \ \varepsilon\Theta \mid \tfrac{1}{n-1}\log D_n(\theta') = \max_{\theta\varepsilon\Theta} \tfrac{1}{n-1} \log D_n(\theta)\}. \end{aligned}$$

Theorem 5.7: Assume (a) Ω has finite cardinality N; (b) for some constant K, $|u_n(\omega|\theta)| \leq K$ for all $\theta\ \varepsilon\ \Theta$, $\omega\ \varepsilon\ \Omega$, $n\geq 1$; (c) there exists h: $\Omega\times\Theta\to R$ such that for each $\omega\varepsilon\Omega$, $u_n(\omega|\theta) \sim h(\omega|\theta)$ uniformly in θ;

(d) for each $\omega\epsilon\Omega$, $p_{n+1}(\omega|\theta) \sim n(\omega)/n$ uniformly in θ; and (e) for all sufficiently large n, the maximizing policy sets M_n^{mp}, M_n^u, and M_n^v for D_n, U_n, and V_n are nonempty. Then

$$U_n(M_n^{mp}) \sim U_n(M_n^u) \sim U_n(M_n^v).$$

Remark: Condition (c) is satisfied if the utility densities $\{u_n\}$ are stationary; i.e., if $u_n(\omega|\theta) = u_1(\omega|\theta)$ for all $\omega\epsilon\Omega$, $\theta \epsilon \Theta$, $n \geq 1$. Condition (e) is satisfied if Θ is compact and the utility and probability densities are continuous in θ.

Proof: By (a) and (b), uniformly in θ,

$$U_1(\theta)/n \equiv \Sigma_\Omega[u_1(\omega|\theta)/n]p_1(\omega|\theta) \to 0.$$

By (a), (b), and (c), uniformly in θ,

$$\Sigma_{j=1}^n u_j(\omega_j^*|\theta)/n \sim \Sigma_{j=1}^n h(\omega_j^*|\theta)/n \equiv \Sigma_\Omega h(\omega|\theta)n(\omega)/n;$$

and

$$\Sigma_\Omega h(\omega|\theta)n(\omega)/n \sim \Sigma_\Omega u_n(\omega|\theta)n(\omega)/n.$$

Finally, by (a), (b), and (d), uniformly in Θ,

$$\Sigma_\Omega u_n(\omega|\theta)n(\omega)/n \sim \Sigma_\Omega u_n(\omega|\theta)p_n(\omega|\theta).$$

Thus, uniformly in θ,

$$\frac{1}{n-1} \log V_n(\theta) \equiv \frac{1}{n-1} \Sigma_{j=1}^{n-1} u_j(\omega_j^*|\theta) + \frac{1}{n-1} U_1(\theta) - K_n \tag{5.11}$$

$$\sim \frac{1}{n-1} \Sigma_{j=1}^{n-1} u_j(\omega_j^*|\theta) - K_n$$

$$\equiv \frac{1}{n-1} \log D_n(\theta) - K_n \tag{5.12}$$

$$\sim \frac{1}{n-1} \Sigma_{j=1}^{n-1} h(\omega_j^*|\theta) - K_n$$

$$\equiv \Sigma_\Omega h(\omega|\theta)(n-1)(\omega)/(n-1) - K_n$$

$$\sim \Sigma_\Omega u_n(\omega|\theta)(n-1)(\omega)/(n-1) - K_n$$

$$\sim \Sigma_\Omega u_n(\omega|\theta)p_n(\omega|\theta) - K_n$$

$$\equiv U_n(\theta) - K_n. \tag{5.13}$$

By definition, the set of maximizing policies for (5.13) is M_n^u; and by Remark 5.6, the set of maximizing policies for (5.11) is M_n^v and for (5.12) it is M_n^{mp}.

Given $\varepsilon>0$, it follows from the above that for sufficiently large n,

$$\left| U_n(\theta) - \frac{1}{n-1} \Sigma_{j=1}^{n-1} u_j(\omega_j^*|\theta) \right| < \varepsilon/2$$

for all $\theta \in \Theta$. In particular, given any $\theta_n^u \in M_n^u$ and $\theta_n^{mp} \in M_n^{mp}$,

$$U_n(\theta_n^u) \equiv \Sigma_\Omega u_n(\omega|\theta_n^u)p_n(\omega|\theta_n^u)$$

$$< \frac{1}{n-1} \Sigma_{j=1}^{n-1} u_j(\omega_j^*|\theta_n^u) + \varepsilon/2$$

$$\leq \frac{1}{n-1} \Sigma_{j=1}^{n-1} u_j(\omega_j^*|\theta_n^{mp}) + \varepsilon/2$$

$$< U_n(\theta_n^{mp}) + \varepsilon.$$

Since $U_n(\theta_n^{mp}) \leq U_n(\theta_n^u)$ by definition of M_n^u, it follows that

$$\left| U_n(\theta_n^u) - U_n(\theta_n^{mp}) \right| < \varepsilon.$$

Thus

$$U_n(M_n^{mp}) \sim U_n(M_n^u).$$

The proof for $U_n(M_n^u) \sim U_n(M_n^v)$ is similar.

Corollary 5.8: Assume (a) Ω and Θ are finite; (b) $u_n(\omega|\theta) \sim h(\omega|\theta)$ for all $\omega \in \Omega$ and $\theta \in \Theta$; and (c) $p_{n+1}(\omega|\theta) \sim n(\omega)/n$ for all $\omega \in \Omega$ and $\theta \in \Theta$. Then

$$U_n(M_n^{mp}) \sim U_n(M_n^{u}) \sim U_n(M_n^{v}).$$

Proof: It is easily shown that conditions (a)-(c) in Corollary 5.8 imply conditions (a)-(e) in theorem 5.7.

Lemma 5.9: Suppose the policy choice set Θ is finite, and there exists a policy $\theta^* \varepsilon \Theta$ such that for each $\theta \neq \theta^*$,

$$\lim_{n\to\infty} \sup \left[\frac{1}{n} \Sigma_{j=1}^n u_j(\omega_j^*|\theta) - \frac{1}{n} \Sigma_{j=1}^n u_j(\omega_j^*|\theta^*)\right] < 0. \qquad (5.14)$$

Then

$$\lim_{n\to\infty} V_n(\theta) = \begin{cases} 0 \text{ if } \theta \neq \theta^*; \\ 1 \text{ if } \theta = \theta^*. \end{cases}$$

Proof: In terms of $v_j \equiv \exp(u_j)$, $1 \leq j \leq n$, (5.14) implies that for $\theta \neq \theta^*$,

$$\lim_{n\to\infty} \sup \frac{1}{n} \log \left(\Pi_{j=1}^n v_j(\omega_j^*|\theta)/\Pi_{j=1}^n v_j(\omega_j^*|\theta^*)\right) < 0,$$

hence

$$\lim_{n\to\infty} \sup \left[\Pi_{j=1}^n v_j(\omega_j^*|\theta)/\Pi_{j=1}^n v_j(\omega_j^*|\theta^*) \right] \leq 0.$$

By positivity of the v_j, it follows that

$$\lim_{n\to\infty} \left[\Pi_{j=1}^n v_j(\omega_j^*|\theta)/\Pi_{j=1}^n v_j(\omega_j^*|\theta^*)\right] = 0.$$

Thus for any policy $\theta' \varepsilon \Theta$,

$$V_{n+1}(\theta') \equiv v_n(\omega_n^*|\theta')\ V_n(\theta')/v_n(\omega_n^*)$$

$$= \frac{\left[\Pi_{j=1}^n v_j(\omega_j^*|\theta')/\Pi_{j=1}^n v_j(\omega_j^*|\theta^*) \right] V_1(\theta')}{\Sigma_{\theta\varepsilon\Theta}\left[\Pi_{j=1}^n v_j(\omega_j^*|\theta)/\Pi_{j=1}^n v_j(\omega_j^*|\theta^*)\right] V_1(\theta)}$$

$$\to \begin{cases} 0 \text{ if } \theta' \neq \theta^*; \\ 1 \text{ if } \theta' = \theta^*. \end{cases}$$

Theorem 5.10: Assume (a) Ω and Θ are finite; (b) there exists $h:\Omega\times\Theta\to R$ such that $u_n(\omega|\theta) \sim h(\omega|\theta)$ for each $\omega\varepsilon\Omega$ and $\theta \varepsilon \Theta$; (c) $p_{n+1}(\omega|\theta) \sim n(\omega)/n$ for each $\omega\varepsilon\Omega$ and $\theta \varepsilon \Theta$; and (d) there exists a policy $\theta^* \varepsilon \Theta$ such that for each policy $\theta \neq \theta^*$,

$$\lim_{n\to\infty} \sup\ [U_n(\theta) - U_n(\theta^*)] < 0.$$

Then

$$\lim_{n\to\infty} V_n(\theta) = \begin{cases} 0 \text{ if } \theta \neq \theta^* ; \\ 1 \text{ if } \theta = \theta^* . \end{cases}$$

Proof: Under conditions (a)-(c) it is easily shown (cf. theorem 5.7, (5.11)-(5.13)) that uniformly in θ,

$$\frac{1}{n} \Sigma_{j=1}^{n} u_j(\omega_j^*|\theta) \sim U_{n+1}(\theta). \tag{5.15}$$

Together with condition (d), (5.15) implies

$$\lim_{n\to\infty} \sup\ [\frac{1}{n} \Sigma_{j=1}^{n} u_j(\omega_j^*|\theta) - \frac{1}{n} \Sigma_{j=1}^{n} u_j(\omega_j^*|\theta^*)] < 0$$

for $\theta \neq \theta^*$. The claim then follows from Lemma 5.9.

FOOTNOTES

[1]Bayes' Theorem for the probability of events E and H states that

$$\text{Prob }(E|H) = \text{Prob }(H|E)\text{ Prob }(E)/\text{Prob }(H),$$

provided that Prob(H) $\neq 0$. Bayes' theorem for probability has been justified by fundamental coherency arguments (see de Finetti [1, Chapter 4]).

[2]Similar arguments have been given to justify the minimax and maximax decision rules which are also based on utility assessments. However, under these rules prior beliefs play no formal role, and new observations only become relevant for the decision maker when they change the worst or best possible outcomes he associates with each of his available actions.

[3]This may have been a motivating factor for the Marschak-Radner team model (see Marschak and Radner [4]) in which all team members are assumed to have identical preferences over outcomes. However, the team members are also assumed to be in complete agreement concerning the basic probability distribution over states.

[4]Unlike the situation for probability (see de Finetti [1]), no general coherency arguments exist which necessitate the finite additivity of a decision maker's utility assessments; i.e., the utility U(E) a decision maker associates with the realization of an event with finite set representation E is not generally expressable by

$$U(E) = \Sigma_{\omega\varepsilon E} U(\omega).$$

On the other hand, it is the content of a number of expected utility theories (e.g., see Jeffrey [3]) that a decision maker's utility assessments are finitely additive with respect to some (probability) measure P, appropriately conditioned; i.e.,

$$U(E) = \Sigma_{\omega \varepsilon E} U(\omega)P(\omega)/P(E).$$

[5]The "policy model" presented in sections 2 and 3 is discussed in greater detail in Tesfatsion [6] and [7]. As shown in Tesfatsion [6], the expected utility model of Savage, the Marschak-Radner team model, the Bayesian statistical decision model, and the standard optimal control model can be viewed as special cases of the policy model. The policy model is extended to a policy game in Tesfatsion [8] and shown to be a generalization of the standard n-person game in normal form.

[6]A binary relation $\geq$ on a set D is a *weak order* if for all a, b, $c \ \varepsilon \ D$

(1) $a \geq b$ or $b \geq a$ (i.e., $\geq$ is connected);

(2) $a \geq b$ and $b \geq c$ implies $a \geq c$ (i.e., $\geq$ is transitive).

[7]A collection F of subsets of a nonempty set X is said to be an *algebra* in X if F has the following three properties:

(1) $X \ \varepsilon \ F$;

(2) If $A \ \varepsilon \ F$, then $A^c \ \varepsilon \ F$, where A^c is the complement of A relative to X;

(3) If $A, B \ \varepsilon \ F$, then $A \cup B \ \varepsilon \ F$.

[8]It will always be assumed that a decision maker's utility assessments specified in an information state s known with certainty coincide with the utility assessments he would make conditioned on that information state; symbolically, $u_s(\omega) = u(\omega|s)$. Similarly for a decision maker's probability assessments. (The latter assumption is implicit in nearly all Bayesian arguments.)

[9]If the state flows ω are not stochastically independent of policy choices θ, then the reasoning implied by (4.2) can lead to undamped vacillation between policies as the decision maker vainly attempts to achieve certain state flow-policy matches. Costs incurred by switching policies from period 1 to period 2 could be handled by the inclusion in the definition of $U_1(\theta|\omega_1^*)$ of an additive cost term $c(\theta|\theta_1^*)$, where θ_1^* is the decision maker's first period policy choice. Such costs will not be considered in this paper.

[10]In probability theory the use of uniform probability densities to represent initial ignorance has often been criticized. For example, if the decision maker is ignorant of the true value of $\theta \in (0,\infty)$, then the same must be said for the true value of $1/\theta$. Yet the simultaneous specification of uniform probability densities over $(0,\infty)$ for both θ and $1/\theta$ leads to contradiction. Moreover, uniform probability densities are often "improper;" i.e., their associated distribution functions assign infinite probability to the set of possible values.

In contrast, a uniform initial period objective function U_1 unambiguously represents the initial indifference of the decision maker between available policy choices. Since utility assessments need not be additive, no contradictions arise. In particular, it is meaningless to ask whether U_1 is "proper" or "improper."

REFERENCES

[1] de Finetti, B., Theory of Probability (Volume 1), New York: John Wiley & Sons, Ltd., 1974.

[2] DeGroot, M. H., Optimal Statistical Decisions, New York: McGraw-Hill, Inc., 1970.

[3] Jeffrey, R. C., The Logic of Decision, New York: McGraw Hill, Inc., 1965.

[4] Marschak, J. and Radner, R., Economic Theory of Teams, New Haven: Yale University Press, 1972.

[5] Savage, L. J., The Foundations of Statistics, second revised edition, New York: Dover Publications, Inc., 1972.

[6] Tesfatsion, L., "An Expected Utility Model with Endogenously Determined Goals," Discussion Paper No. 75-58, Center for Economic Research, University of Minnesota, September 1975.

[7] Tesfatsion, L., "Axiomatization for an Expected Utility Model With Endogenously Determined Goals," Discussion Paper No. 75-59, Center for Economic Research, University of Minnesota, September 1975.

[8] Tesfatsion, L., "Games, Goals, and Bounded Rationality," Discussion Paper No. 76-63, Center for Economic Research, University of Minnesota, January 1976.