JOURNAL OF ECONOMIC THEORY **69**, 53–70 (1996)
ARTICLE NO. 0037

Discounting and Optimizing: Capital Accumulation Problems as Variational Minmax Problems

PIERRE-YVES GEOFFARD[1]

Département et laboratoire d'économie théorique et appliquée, CNRS, 48 boulevard Jourdan, 75014 Paris, France

Received July 14, 1992; revised May 30, 1995

The variational utility of a deterministic, infinite horizon consumption path is defined as the minimum value of an additive criterion, taken over all possible future values of the rate of time preference. Special cases of variational utility include additive and recursive models. A reduced-form capital accumulation problem under variational utility is studied here. The optimal capital path is characterized by a Hamiltonian dynamical system, together with two transversality conditions at infinity. The Hamiltonian structure is applied to standard additive, or recursive, capital accumulation problems. Under a separability assumption, a stability result is extended from the additive to the recursive case. *Journal of Economic Literature* Classification Numbers: C61, D90, E22.

1. INTRODUCTION

According to Magill [24], the concept of discounting was developed for practical purposes by financiers and accountants, long before it belonged to the standard toolbox of economists. Discounting was introduced into economic theory by Böhm–Bawerk [6]; Fisher [19] and Koopmans [22], among many others have also advocated its use, and have added important contributions to the formal definition of impatience.

At the same time, the practice of discounting was criticized, or even condemned. Ramsey wrote in his celebrated paper [26]: "we do not discount later enjoyments in comparison with earlier ones, a practice which is *ethically indefensible* and arises merely from the weakness of the imagination" (my emphasis). Others, like Rawls ([27], p. 293) with different concerns,

[1] This work was completed while I was visiting Stanford University. Conversations with Rose-Anne Dana, Darrell Duffie, Ivar Ekeland, Larry Epstein, José Scheinkman, Costis Skiadas, and seminar participants at the Universities of Chicago and Paris were fruitful. An anonymous referee and an associate editor suggested substantial improvements. Financial support from Caisse autonome de refinancement and CNRS is gratefully acknowledged. Any remaining errors are mine.

0022-0531/96 $18.00

even claim that "In the case of an individual the avoidance of pure time preference is a feature of being rational."

Nevertheless recent debates have most often focused on how to discount the future, rather than why the future should be discounted. In the theory of intertemporal allocation, the central concept is the rate of time preference. The simple and tractable structure of additive utility models relies on the assumption of a constant rate of time preference. Unfortunately, in multi-agent problems, this assumption leads to degenerate situations in which the most patient agent accumulates all the capital. Moreover, intertemporal substitution and risk aversion cannot be disentangled (Epstein and Zin [17]). A more general class of preferences has been introduced by Uzawa [35] and Epstein and Hynes [16], based on the assumption that the rate of time preference depends on consumption. Following Koopman's pioneering work [22], Lucas and Stokey [23] and Epstein [15] have introduced and developed the concept of recursive preferences.

The present paper introduces *variational utility functions*, a class of utility functions for which the dynamics of the discount factor obeys a minimum principle. Recursive utility functions (Epstein [15], Uzawa [35]), or standard additive utility functions, belong to this class. The variational utility of a given path is defined as the minimum value of the additive sum of all future *felicities*. At each date t in the future, felicity is a function of the current value of consumption, discount factor, and rate of time preference. The minimum is taken over all possible future values of the rate of time preference in some admissible set.

Optimal capital paths under variational utility are characterized by a stationary Hamiltonian dynamical system in dimension $n+1$, where n is the number of capital goods. The additional state variable is the discount factor. The dual to the capital stock is the discounted capital price, and the dual to the discount factor can be interpreted, in the recursive specification, as the recursive utility itself. Variational utility functions are the utility functions for which the optimal capital path is characterized by such a stationary Hamiltonian system.

Under some specification of the felicity function, variational utility coincides with standard additive utility models. This specification provides a method for embeddings familiar, modified Hamiltonian dynamical systems (associated with Cass, Shell, Brock, Scheinkman, and Rockafellar [8]) in a one-dimension-bigger space, and hence reveals a richer structure. This structure, interesting in itself, can be used to extend traditional results, which are known in the additive case, to recursive utility (or even to variational utility).

In a stochastic framework of finance theory, variational utilities and the associated backward stochastic differential equations have been used in pricing theory by El Karoui *et al.* [14] to establish a general duality principle between hedging and pricing problems.

Section 2 formally defines variational utility functions. When the felicity function is multiplicatively separable in discount, a characterization result proves that the variational utility function coincides with the recursive utility function.

Section 3 studies a reduced-form capital accumulation problem under variational utility. The optimal capital path is characterized by a Hamiltonian dynamical system, together with two transversality conditions at infinity.

Section 4 applies this Hamiltonian structure to the standard additive capital accumulation problem and to the recursive case. Under an assumption of separability, a stability result is extended from the additive to the recursive case.

2. The Minimum Principle and the Discount Function

This section presents a general class of utility functions, *variational utilities*, for which the evolution of the discounting function obeys a minimum principle. General recursive utility functions, as defined by Epstein [15], belong to this class.

Definition 1. Given a consumption path c, the *variational utility* associated with this path is equal to the minimum, taken over all future possible rates of time preference, of

$$U(c) \equiv \min_{r(\cdot) \in \mathscr{R}} \int_0^{\infty} f(c_t, B_t, r_t)\, dt, \tag{1}$$

where $\mathscr{R}$ is a set of admissible paths of rates of time preference, and f gives current felicity as a function of the current value of consumption c_t, discount factor B_t, and rate of time preference r_t.

As noted by Simon [32], this "minimum principle" approach to the theory of choice under uncertainty was formally introduced by Arrow [1].

An interpretation of variational utility is given by the "worst-case scenario" analysis, often used in management or operational research, for example credit rating provided by independent agencies (Moody's, Standard and Poors, etc.). Typically, rating agencies have to issue a public statement about the value of a firm's credit; but they only have limited information about the firm. In particular, the future evolution of relevant variables (sales volumes, asset default rates or prepayment rates, etc.) is very important to the valuation of the firm's debt. A probability distribution on the future values of these variables may be difficult to define.

Instead, it may be more intuitive to assume that these variables will remain within some confidence interval, and to define the value of the debt as the value in the worst case, i.e., when the evolution of the relevant variables is systematically adverse. Another example is the alternative approach to stochastic models of the term structure of interest rates developed by Dermody and Rockafellar [9] and Bentami *et al.* [4]: they define the value of a given cash-flow as the worst discounted final value of the stock, assuming that the short rate is systematically unfavorable in the future (that it is high when borrowing, and low when lending).

Formally, a consumption path is a function $c\colon \mathfrak{R}_+ \to \Lambda$, where $\Lambda \subset \mathfrak{R}^m$ is the consumption set. The rate of time preference r_t belongs for every time t to some interval $[\delta, D]$, with $0<\delta<D$, and the viability set $\mathscr{R}$ is the set of all continuous functions such that this pointwise condition is met almost surely. It should be noted that since the discount factor, its derivative, and the rate of time preference are linked by the relation $\dot{B}_t = -r_t B_t$, then felicity can be expressed as a function of $(c_t, B_t, \dot{B}_t)$ instead. In this case the minimum in Eq. (1) is controlled by $\dot{B}$ instead of r. The pointwise condition $r_t \in [\delta, D]$ defines the viability set V for $(B, \dot{B})$ as $V = \{(x, y) \in \mathfrak{R}^2, x \in [0, 1], -Dx \leqslant y \leqslant -\delta x\}$.

Definition 2. The set of admissible discount functions, denoted $\mathscr{B}$, is the set of all absolutely continuous functions B from $\mathfrak{R}_+$ to $\mathfrak{R}$, such that:

- $B_0 = 1$,
- for almost any t, the viability condition $(B_t, \dot{B}_t) \in V$ holds.

If a finite measure μ on $\mathfrak{R}_+$ of density $\exp(-\delta t)\, dt$ is defined, and some $p>1$ is fixed, the Radon–Nykodym derivative $\dot{x}$ is defined, for any absolutely continuous function x, in the L^1_{loc} sense (almost everywhere, and locally integrable). An admissible discount function B satisfies, for almost every t, $|\dot{B}_t| \leqslant D$. This, in turn, implies that, as a function, $\dot{B}$ belongs to $L^p(\mu)$. This technical property is needed to prove that variational utility, as given in Definition 1, does exist.

A discount function is called *locally interior* if for almost every t, $(B_t, \dot{B}_t)$ belongs to the interior of V.

For felicity, as a function of $(c, B, \dot{B})$, the following assumption is made:

Assumption 1. For all c in Λ, the function $f(c, \cdot)$ is a convex, differentiable function from V to $\mathfrak{R}$, such that for all B, $f(c, B, \cdot)$ is non-decreasing on its domain, and bounded below by $-KB$, for some constant K, uniform in c.

In particular cases of interest (recursive utility), felicity can be shown to verify the convexity assumption, and Assumption 1 extends this property to the general case. The assumption of differentiability is made in order to rule out all corner solutions, and it will be used to characterize the solution of the minimization problem by first-order conditions. Consuming now, *ceteris paribus*, is preferred when the rate of time preference is high, i.e., when the agent is not eager to postpone consumption. Hence felicity is assumed to increase with r.

Values of f outside $\Lambda \times V$ are defined by $f = -\infty$ if c is not in Λ, and $f = +\infty$ if c is in Λ but $(B, \dot{B})$ is not in V. This extended function is now defined from $\mathfrak{R}^m \times \mathfrak{R}^2$, and takes its values in $\bar{\mathfrak{R}} = \mathfrak{R} \cup \{-\infty, +\infty\}$. The set of extended real numbers $\bar{\mathfrak{R}}$ is commonly used in convex analysis, e.g., to embed constraints in the integrand (see, e.g., Rockafellar [28]).

The utility of a consumption path is defined by the classical Lagrange variational criterion:

$$U(c) = \min_{B \in \mathscr{B}} \int_0^\infty f(c_t, B_t, \dot{B}_t)\, dt, \tag{2}$$

It should be noticed immediately that, as a direct consequence of Assumption 1, for any consumption path c and any admissible discount function B, the integrand in Eq. (2) is bounded below by $-Ke^{-\delta t}$. Hence, the integral is bounded below by $-K/\delta$, and cannot take negatively infinite values. Nevertheless, it can take positively infinite values. On the other hand, if some path c is not a consumption path (i.e., if on a non-null set of time, c_t is not in Λ), then the integral $U(c)$ is equal to $-\infty$ for all discount functions.

2.1. *Characterization of the Discount Function*

A theorem by Benveniste and Scheinkman [5] characterizes optimal solutions to Lagrange problems. In particular, for a given consumption path c, it can be used to characterize the discount factor B^*, solution to (2). To write this characterization theorem under a Hamiltonian form requires defining

$$H(t, B, u) \equiv \min_{\dot{B} \mid (B, \dot{B}) \in V} \{f(c_t, B, \dot{B}) + \dot{B}u\}.$$

PROPOSITION 2.1. *Under Assumption* 1, *an optimal solution* B^* *to Problem* (2) *exists.*

Assume B^* *is locally interior, and denote by* u^* *the dual path associated to* B^*. *These two paths are characterized by the Hamiltonian dynamics* $\dot{u}_t^* = -H_B$ *and* $\dot{B}_t^* = H_u$, *and the transversality condition* $\lim_{t\to\infty} B_t^* u_t^* = 0$.

Proof. Existence is proved by the traditional Weierstrass method (see, e.g., Ekeland and Temam [13], Geoffard [21], or Romer [29]), using the weak L^p topology on the derivative. When the criterium is concave, weak-L^p upper-semi-continuity and strong-L^p upper-semi-continuity are equivalent, and follow from Fatou's lemma. Moreover, since the maximization set is convex, weak-L^p compactness is equivalent to boundedness and strong-L^p closeness (by the Alaoglu theorem, which only holds in reflexive spaces, hence the condition $p > 1$).

The characterization result is a direct application of Theorem 3A in Benveniste and Scheinkman [5]. Under differentiability, characterization by the first-order conditions is a consequence of the Pontryagin principle, which is merely a different way of writing the Euler equation. Necessity of the transversality condition for a locally interior solution B^* is due to monotonicity of the integrand in $\dot{B}$, and positiveness of B: an alternative admissible discount function B can be constructed such that, for some positive constant α, and for all t large enough, $0 \leqslant B_t^* u_t^* \leqslant \alpha u_t^*(B_t - B_t^*)$, which tends to zero since B is admissible. Sufficiency of the transversality condition is due to convexity of the integrand. Q.E.D.

2.2. *Recursive Utility*

Recursive utility, as defined by Epstein [15], is a special case of variational utility. This case is obtained when felicity, as a function of consumption, discount, and rate of time preference, is multiplicatively separable in discount: $f(c, B, r) = BF(c, r)$. Function F gives the *undiscounted* felicity, as a function of current consumption and rate of time preference. Written in terms of undiscounted felicity F instead of felicity, Assumption 1 becomes:

Assumption 2. For all c, $F(c, \cdot)$ is a convex, continuously differentiable non-decreasing function, bounded below by some constant $-K$.

Let W be the Legendre transform of F w.r.t. r: $W(c, u) \equiv \min_r \{F(c, r) - ru\}$. Under Assumption 2, felicity f is convex w.r.t. $(B, \dot{B})$, since due to the conjugacy relation (see, e.g., Ekeland and Temam [13]), it is equal to:

$$\begin{aligned} f(c, B, \dot{B}) &= BF\left(c, -\frac{\dot{B}}{B}\right) \\ &= B \max_u \left\{ W(c, u) - \frac{\dot{B}}{B} u \right\} \\ &= \max_u \{ BW(c, u) - \dot{B}u \}, \end{aligned}$$

and the maximum of linear functions is a convex function.

The following proposition shows how variational utility coincides with recursive utility, when felicity is separable.

PROPOSITION 2.2. *Assume that* $f(c, B, r) = BF(c, r)$, *where* F *satisfies Assumption* 2, *and set* $W = F^*$, *the Legendre transform of* F. *Then*:

- for all t, the value of the dual variable u_t^* is the recursive utility of the remaining path defined by the aggregator W;
- the value of the problem is the recursive utility of the whole path c.

Proof. When $f = BF$, the Hamiltonian H in Proposition 2.1 satisfies $H = B \min_r \{F(c_t, r) - ru\} = BW(c_t, u)$. At the minimum, the rate of time preference r_t^* is equal to $-W_u(c_t, u_t^*)$, and the derivatives of the Hamiltonian are given by $H_B = W$ and $H_u = BW_u$. Hamiltonian equations of Proposition 2.1 are $\dot{u}_t^* = -W$ and $\dot{B}_t^* = B^* W_u$. Since B_0^* is equal to one, the second equation solves into $B_t^* = \exp\left(\int_0^t W_u(c_s, u_s^*)\, ds\right)$. The equation $\dot{u}_t^* = -W$ is the recursive definition of utility, given the aggregator W. The transversality condition $B^* u^* \to 0$ closes the definition and shows that the dual variable u^* equals the recursive utility.

For optimal control problems linear in the state, the value is equal to the scalar product of the state and the dual variable. Therefore, utility of the whole path, $U(c)$, is equal to $B_0^* u_0^*$. The conclusion of the proof follows from the fact that $B_0^* = 1$: the utility of the whole path c, $U(c)$, is equal to u_0^*, the recursive utility of c. Q.E.D.

As an important by-product, Proposition 2.2 gives a semi-explicit formula for the recursive utility U of a path defined by an aggregator W. With $F = W^*$, the Legendre transform of the aggregator W,

$$U(c) = \min_{B \in \mathscr{B}} \int_0^\infty B_t F(c_t, r_t)\, dt.$$

Proposition 2.2 also reveals the dual relation between the aggregator W and the felicity function F. In this light, the recursive equation $\dot{u} = -W(c, u)$ appears as the Euler equation characterizing the solution to a variational problem. It is analogous to the law of least action governing motion in physics.

Also note that Proposition 2.2 uses a limit condition at infinity to define the recursive utility of a path. Epstein [15], focusing on bounded utilities, defines the recursive utility of a path as the only initial value u_0, such that the solution to $\dot{u} = -W(c, u)$ exists and is bounded for all t. Duffie *et al.* [11] state explicitly the hidden transversality condition: the recursive utility is the only initial value u_0 such that $\exp(-\delta t)\, u_t$ goes to zero at infinity. This transversality condition is equivalent to the condition $\lim_{t \to \infty} \exp\left(\int_0^t W_u(c_s, u_s)\, ds\right) u_t = 0$ of Proposition 2.2.

The separability assumption has also simple implications for the value function. For a given path c, the function V is defined by

$$V(\beta, t) \equiv \min_{B \in \mathscr{B};\, B_t = \beta} \int_t^\infty B_s F(c_s, r_s)\, ds.$$

V gives the value of the minimization problem as a function of the initial value of the discount. The recursive utility $U(c)$ of a path c is equal to $V(1, 0)$, and the recursive utility of its tail, $U({}_tc)$, is equal to $V(1, t)$. Given linearity of the integrand in B, β can be factorized, and

$$V(\beta, t) = \beta \min_{B \in \mathscr{B};\, B_t = 1} \int_t^\infty B_s F(c_s, r_s)\, ds = \beta V(1, t) = \beta U({}_tc).$$

The value function is differentiable, and is homogeneous of degree one, and for all β, $V_\beta(\beta, t) = U({}_tc)$. The derivative of the value function w.r.t. the state (the discount factor) at time t is equal to the dual variable at that time. This is a traditional duality result (see Benveniste and Scheinkman [5]) that holds in more general situations.

2.3. *Uzawa or Additive Utility*

The previous construction still applies to the special case where the generator W is linear w.r.t. u. Two cases are of interest: additive utility and Uzawa utility. Uzawa [35] writes the utility of a path as

$$U(c) = \int_0^\infty \exp\left(\int_0^t - R(c_s)\, ds\right) v(c_t)\, dt.$$

This utility is a straightforward extension of the additive utility, making the rate of time preference R depend on current consumption. The explicit form of the utility and its flexibility make it of great use in many applications (e.g., problems in international trade by Findlay [18] or Obstfeld [25]).

This utility corresponds to the special case when the aggregator W is linear in u and given by $W(c, u) = v(c) - R(c)\, u$. This specification is obtained for a rate of time preference that depends only on current consumption. When the rate of time preference is constant, we obtain the traditional additive utility model.

The felicity function associated to a linear aggregator may take infinite values, and is defined by

$$\begin{cases} F(c, r) = v(c) & \text{if} \quad r = R(c) \\ F(c, r) = +\infty & \text{if} \quad r \neq R(c). \end{cases} \tag{3}$$

Taking the minimum of the integral over all possible future rates ensures that the "worst" rate at time t will almost always be equal to $R(c_t)$ (to a constant ρ in the additive case): this is the only way the integral can be finite.

3. Capital Accumulation

This section studies an optimal growth model under variational utility. Formally, the problem is a maxmin problem: the maximum is taken over all feasible capital paths, and the minimum over all discount functions. Under standard convexity assumptions, the optimal capital path is characterized by its Hamiltonian dynamics in dimension n, the number of capital goods. Together with the evolution of the discount factor and its dual variable, the solution to the maxmin problem is fully characterized by a Hamiltonian system in dimension $n+1$, autonomous in time. The construction applies to all specific cases of variational utility, such as additive utility, for which the construction shows that the familiar modified Hamiltonian dynamics (Cass and Shell [7, 8], Scheinkman [30]) in dimension n can be embedded in a larger space (in dimension $n+1$) and exhibit a true Hamiltonian structure.

3.1. *Capital Paths*

Definition 3. Given an initial capital stock ξ, a *feasible capital path* is an absolutely continuous functions $k\colon \mathfrak{R}_+ \to \mathfrak{R}^n$ such that

- $k_0 = \xi$;
- $\dot{k} \in L^p(\mu)$;
- the technology condition $(k_t, \dot{k}_t) \in T$ holds for a.e. t.

The *feasible set* $X(\xi)$ is the set of all feasible capital paths.

Since μ is a finite measure, all L^p spaces form a decreasing sequence with p. Therefore, it is true that the condition $\dot{k} \in L^1(\mu)$ would be weaker, and could possibly extend the applicability of the characterization result below. However, finding a counter-example, i.e., an economic model where investment flows are in $L^1(\mu)$ without being in $L^p(\mu)$ for any $p > 1$, is not easy.

The technology set T satisfies the following assumption:

Assumption 3. The technology set T is a closed, convex subset of $\mathfrak{R}^n_+ \times \Theta$, with non-empty interior, where Θ is a bounded subset of $\mathfrak{R}^n$.

3.2. *The Reduced Form Model*

The consumption path associated with a capital path is given by the *production function* $\pi: T \to \Lambda$. Every good produced that is not invested is consumed: $c_t = \pi(k_t, \dot{k}_t)$.

An extra technical assumption is made concerning the feasible set:

Assumption 4. If k is a locally interior, feasible capital path, then when τ is sufficiently large, there exists another feasible path, dominated on $[\tau, +\infty]$ by $\frac{1}{2}k$.

This assumption states that if technology T, with proper investment, can asymptotically sustain some capital path k, then it can also sustain half of k; it is required to prove the necessity of the transversality condition.

The production function π satisfies the following assumption:

Assumption 5. The production function π is a quasi-concave, continuously differentiable function of $(k, \dot{k})$, non-increasing in $\dot{k}$.

Quasi-concavity of the production function (for any vector C, the set of (x, y) such that $\pi(x, y) \geqslant C$ is a convex subset of T) precludes increasing returns. As has been show by Romer [29] in the additive case, convexity with respect to the highest derivative is sufficient to prove the existence of an optimal solution. In the recursive case a similar existence result has been proved [20] assuming that the aggregator is uniformly continuous. However, Assumption 5 simplifies the characterization result below.

The trade-off between investment and consumption is formally expressed by the assumption that π is non-increasing in $\dot{k}$.

As a function of consumption, the felicity function satisfies the following assumption:

Assumption 6. Felicity f is a concave, continuously differentiable, and non-decreasing function of c.

Variational utility of a capital path k is defined via the Lagrangian L:

$$L(k, \dot{k}, B, \dot{B}) \equiv f(\pi(k, \dot{k}), B, \dot{B}). \tag{4}$$

It should be noted that the Lagrangian L is autonomous in time; first order optimality conditions lead to true Hamiltonian dynamics. Defining the integral as $\mathscr{L}(k, B) = \int_0^\infty L(k_t, \dot{k}_t, B_t, \dot{B}_t)\, dt$, the capital accumulation problem can be stated as:

$$\max_{k \in X} U(k) = \max_{k \in X} \min_{B \in \mathscr{B}} \mathscr{L}(k, B). \tag{5}$$

An *optimal path* is a feasible path k^*, such that $U(k^*) \geqslant U(k)$ for any other feasible path k.

3.3. *Conditions for Optimality*

The following theorem characterizes the solution (k^*, B^*) to the variational problem $\max_k \min_B \mathscr{L}(k, B)$. A combination of traditional existence results for minmax convex–concave problems, together with recent existence results for recursive utility maximization problems (either the topology used in Balder [2] or that used in Geoffard [20] can be used), could be developed to prove existence of a solution to variational utility maximization problems. Interiority could be guaranteed, in principle, by some "Inada" conditions (see Duffie *et al.* [12] for an example of such conditions under recursive utility). Assuming that a locally interior optimal solution exists, the following result may be stated:

THEOREM 3.1. *A locally interior* (k^*, B^*) *is a solution to the variational problem* (5) *if and only if the two Euler equations and the two transversality conditions hold*:

- *Euler equation* (*discounting*):

$$\frac{d}{dt} L_{\dot{B}} = L_B$$

- *Transversality condition* (*discounting*):

$$\lim_{t \to \infty} L_{\dot{B}} B = 0$$

- *Euler equation* (*optimizing*):

$$\frac{d}{dt} L_{\dot{k}} = L_k$$

- *Transversality condition* (*optimizing*):

$$\lim_{t \to \infty} \langle L_{\dot{k}} \,|\, k \rangle = 0.$$

Proof. The first two conditions are straightforward consequences of the characterization theorem of Benveniste and Scheinkman [5], applied to the definition of the utility U, which must hold for any capital path k. To prove that the last two conditions hold, we prove that the maxmin problem (5) is equivalent to the corresponding minmax problem.

Required assumptions on L as a function of $(k, \dot{k})$ follow from assumptions (5) and (6) on π and f. Since π is quasi-concave, and since f is concave and non-decreasing in c, L is quasi-concave in $(k, \dot{k})$. Since f is non-decreasing in c and π is non-increasing in $\dot{k}$, L is non-increasing in $\dot{k}$. With respect to $(B, \dot{B})$, L inherits the convexity property of f.

The integral $\mathscr{L}$, is therefore quasi-concave in $(k, \dot{k})$ and convex in $(B, \dot{B})$. If the maximum is attained at some (k^*, B^*), then the following saddle-point property holds:

$$\mathscr{L}(k^*, B^*) = \min_B \mathscr{L}(k^*, B); \tag{6}$$

$$\mathscr{L}(k^*, B^*) = \max_k \mathscr{L}(k, B^*). \tag{7}$$

The maxmin is a supinf, and the order can be reversed. The optimal solution (k^*, B^*) is characterized by the set of two Euler equations, each with a corresponding transversality condition to infinity: given B^*, since k^* satisfies (7) and $\mathscr{L}$ is quasi-concave in k, the characterization result in Benveniste and Scheinkman [5] proves that the Euler equation $(d/dt)\, L_{\dot{k}} = L_k$ must hold, together with the transversality condition $\lim_{t \to \infty} \langle L_{\dot{k}} | k \rangle = 0$. Similarly, Eq. (6) proves that the Euler equation $(d/dt)\, L_{\dot{B}} = L_B$ and the transversality condition $\lim_{t \to \infty} L_{\dot{B}} B = 0$ hold.

The first two conditions prove that $\mathscr{L}(k^*, B^*)$ is equal to $U(k^*)$. The last two conditions prove that $\mathscr{L}(k^*, B^*)$ also equals $\max_k \mathscr{L}(k, B^*)$. Therefore,

$$U(k^*) = \max_k \mathscr{L}(k, B^*) \geqslant \max_k \min_B \mathscr{L}(k, B) = \mathscr{L}(k^*, B^*),$$

which proves that k^* is an optimal capital path, and concludes the proof. Q.E.D.

3.4. *Hamiltonian Formulation*

In order to write the Hamiltonian dynamics of an optimal interior solution (k^*, B^*), the dual variables associated with the state variables k and B are denoted by q and u, respectively. The Hamiltonian is defined as

$$H(k, q, B, u) \equiv \max_{\dot{k}} \min_{\dot{B}} \{ L(k, \dot{k}, B, \dot{B}) + u\dot{B} + \langle q | \dot{k} \rangle \}. \tag{8}$$

The whole set of necessary and sufficient conditions for optimality is

$$\begin{cases} \dot{k} = H_q & k(0) = k_0 \\ \dot{q} = -H_k & \lim\limits_{t \to \infty} \langle q_t | k_t \rangle = 0 \\ \dot{B} = H_u & B(0) = 1 \\ \dot{u} = -H_B & \lim\limits_{t \to \infty} B_t u_t = 0. \end{cases} \tag{9}$$

Notice that the quantity $E_t = H(k_t, q_t, B_t, u_t)$ remains constant along a solution (a fundamental property of autonomous Hamiltonian systems).

4. Some Applications

4.1. *Additive Utility: A Hamiltonian Characterization*

In additive utility maximization problems of the kind:

$$\max_{k \in X(k_0)} \int_0^\infty e^{-\rho t} v(k_t, \dot{k}_t)\, dt, \tag{10}$$

the stock is the state variable, and the dual variable is the capital price. Since the discount factor is written explicitly as $e^{-\rho t}$, the dual variable can be either the current price p or the discounted price q, and two Hamiltonians may be written. (For all details, see Cass and Shell [8].)

In terms of the current price, the Hamiltonian H^1 is

$$H^1(k, p) = \max_{\dot{k}} \{v(k, \dot{k}) + \langle p \mid \dot{k} \rangle\}. \tag{11}$$

The Hamiltonian H^1 is autonomous in time, but the optimal solution is characterized by a *Modified Hamiltonian Dynamical System*, in the terminology of Scheinkman [30]:

$$\begin{cases} \dot{k} = H^1_p \\ \dot{p} = \rho p - H^1_k. \end{cases}$$

The term ρp breaks the Hamiltonian structure of the problem.

However, a true Hamiltonian H^2 can be written, in terms of the discounted price:

$$H^2(t, k, q) = \max_{\dot{k}} \{e^{-\rho t} v(k, \dot{k}) + \langle q \mid \dot{k} \rangle\}. \tag{12}$$

This is the standard Hamiltonian associated with the Lagrangian of problem (10). The dynamics in terms of H^2 are truly Hamiltonian; but H^2 is not autonomous in time.

When a problem is not autonomous in time, the natural question to be asked is whether the problem could be embedded in a bigger space and reveal a richer structure. The answer is often affirmative; it is always affirmative in some weak sense. The additional variable z, defined by $\dot{z} = 1$, and $z(0) = 0$, turns any non-autonomous problem in x into an autonomous problem in (x, z). But of course, this artefact does not reveal any structure. However, in the particular case of additive utility growth models, the variational approach and the fully Hamiltonian dynamics (9) totally characterize the optimal solution. The additional state variable is the discount factor, and its dual variable is the utility of the capital path to come.

As stated in Section 2.3, additive utility corresponds to the following Lagrangian:

$$\begin{cases} L(k, \dot{k}, B, \dot{B}) = Bv(k, \dot{k}) & \text{if } \dot{B} = -\rho B \\ L(k, \dot{k}, B, \dot{B}) = +\infty & \text{if } \dot{B} \neq -\rho B. \end{cases}$$

The Hamiltonian H is equal to

$$H(k, q, B, u) = \max_{\dot{k}} \{ Bv(k, \dot{k}) - \rho Bu + \langle q \mid \dot{k} \rangle \}.$$

With previous notations, $H(k, q, B, u) = BH^1(k, B^{-1}q) - \rho Bu$. Once written in terms of H^1, the dynamical system may be written as

$$\begin{cases} \dot{k} = H_q = -BH^1_k \\ \dot{q} = -H_k = H^1_p \\ \dot{B} = H_u = -\rho B \\ \dot{u} = -H_B = -H^1 + B^{-1} \langle H^1_p \mid q \rangle + \rho u. \end{cases}$$

The last equation (the dynamics of u) can be simplified into the recursive equation: $\dot{u} = -v(k, \dot{k}) + \rho u$. The variation in utility through time ($-\dot{u}$) is equal to the utility given by current consumption, less the "carrying costs," or the costs of postponing satisfaction. Moreover, the utility of the remaining path at time t can be written explicitly as:

$$u_t = B_t^{-1} \int_t^{\infty} B_s v(k_s, \dot{k}_s)\, ds.$$

The initial conditions (on the state variables) are $k(0) = k_0$, the initial stock, and $B_0 = 1$. The (traditional) transversality condition is $\lim_{t \to \infty} \langle q_t \mid k_t \rangle = 0$. The additional transversality condition is $\lim B_t u_t = 0$. Since $B_t u_t = \int_t^{\infty} B_s v(k_s, \dot{k}_s)\, ds$, this condition ensures that the integral $\int_0^{\infty} B_s v(k_s, \dot{k}_s)\, ds$ is defined.

4.2. *Recursive Utility*

In the recursive case, the same Hamiltonian equations can be written. Since the aggregator W is the conjugate (w.r.t. state B) of the Lagrangian L, the Hamiltonian H defined by (8) may be written as

$$\begin{aligned} H(k, q, B, u) &= \max_{\dot{k}} \min_{\dot{B}} \{ L(k, \dot{k}, B, \dot{B}) + u\dot{B} + \langle q \mid \dot{k} \rangle \} \\ &= \max_{\dot{k}} \{ BW(k, \dot{k}, u) + \langle q \mid \dot{k} \rangle \}. \end{aligned}$$

This is the Hamiltonian introduced by Epstein [15] in a multi-agent context. Hamiltonian dynamics, together with appropriate limit conditions, characterize the optimal solution to a recursive utility maximization problem.

Other attempts have been made to find the dual properties, or the Hamiltonian structure, of recursive problems. Major contributions have been made by Becker and Boyd [3] and Sorger [33]. They only deal with a recursive criterion *á la* Uzawa, where the aggregator is $W(c, u) = v(c) - R(c)\, u$. It is linear in utility, and not concave.

Becker and Boyd give conditions under which the value function is smooth, and provide a more complete duality theory in the Uzawa case. The limit condition in utility-and-discount is stated as an assumption, and Sorger conjectures that it is a transversality condition. Theorem 3.1 in the present paper proves this conjecture.

However, linearity of the aggregator leads in a wrong direction: all attempts to write the problem as a "max–max" variational problem fail. Extra state and dual variables are introduced, but the variational problem is written under a concavity assumption. As Sorger [33] notes, this assumption is not met in the general case: the Hamiltonian is the maximum over all possible controls of a "pre-Hamiltonian," which is linear in the discount variable. Therefore it is convex, and corresponds to a minimization variational problem. The concavity assumption is only met in the edge case of a linear aggregator (i.e., both concave and convex), that is, precisely for Uzawa utility.

Indeed, it is only by investigating the general recursive problem that the present study was able to exhibit the maxmin structure, and the minimum principle governing the evolution of the discount.

4.3. *The Value Function in the Recursive Case*

This section derives simple properties of the value function in the recursive case, assuming differentiability. In the Uzawa case, sufficient conditions for the value function to be differentiable are given by Becker and Boyd [3]. We did not investigate the conditions under which the same holds true for the general recursive case, and simply assumed differentiability.

The value function V to the maxmin problem is defined by

$$V(\beta, x, t) \equiv \max_{k \in X,\, k_t = x} \ \min_{B \in \mathscr{B};\, B_t = \beta} \int_t^{\infty} B_s F(k_s, \dot{k}_s, r_s)\, ds.$$

PROPOSITION 4.1. *In the recursive case, the value function is separable and independent of time*: $V(\beta, x, t) = \beta \tilde{V}(x)$. *Moreover, if the value function* *is differentiable in stock, then the dual variable* q_t *is equal to* $B_t \tilde{V}'(k_t)$ *for all* t.

Proof. Stationarity implies that V does not depend on t, and linearity of the integrand w.r.t. B implies that V is homogeneous of degree one in β: for all (β, x, t), $V(\beta, x, t) = \beta V(1, x, 0)$: that is with $\tilde{V}(x) \equiv V(1, x, 0)$, we obtain $V(\beta, x, t) = \beta \tilde{V}(x)$. Function $\tilde{V}$ is the *undiscounted* value function for the capital accumulation problem, and it only depends on the current capital stock.

Under the differentiability assumption, the derivative of the value function w.r.t. capital, calculated along the optimal capital path, is equal to the dual variable (the price): $V_k(B_t, k_t, t) = q_t$. This completes the proof: $q_t = B_t \tilde{V}'(k_t)$.

4.4. *Separability and Stability*

The Hamiltonian characterization can be useful in many applications. This section presents hints of a stability result under a separability assumption. It is merely an extension, to the recursive case, of a result given by Scheinkman [31] in the additive case. This example shows how the Hamiltonian characterization can be a simple and powerful tool to extend results known in the additive case to the recursive case.

Concavity of the integrand w.r.t $(k, \dot{k})$ and convexity of the technology set imply that the value function is concave.

We assume the value function is differentiable. Parallel to the separability property $u(k, \dot{k}) = u^1(\dot{k}) + u^2(k)$ assumed by Scheinkman [31], it is assumed that the aggregator is separable: $W(k, \dot{k}, u) = W^1(\dot{k}) + W^2(k, u)$. Then the Hamiltonian is separable, $H(k, q, B, u) = BH^1(B^{-1}q) + H^2(k, B, u)$.

The function $\phi(t) \equiv H^1(B_t^{-1} q_t)$, defined along an optimal path, is a Lyapunov function. Recalling that $q_t = B_t \tilde{V}'(k_t)$, we obtain

$$\begin{aligned}
\dot{\phi}(t) &= H_p^1(B_t^{-1} q_t) \cdot \left[-\frac{\dot{B}_t}{B_t^2} q_t + \frac{1}{B_t} \dot{q}_t \right] \\
&= H_p^1(B_t^{-1} q_t) \cdot \left[-\frac{\dot{B}_t}{B_t^2} q_t + \frac{1}{B_t} (\dot{B}_t V'(k_t) + B_t V''(k_t) \dot{k}_t) \right] \\
&= H_p^1(B_t^{-1} q_t) \cdot [V''(k_t) \dot{k}_t] \\
&= \dot{k}_t V''(k_t) \dot{k}_t.
\end{aligned}$$

This quantity is non-positive given concavity of the value function, and therefore $\phi(t)$ decreases along an optimal path. This proves that the steady-state is globally stable.

5. Conclusion

The variational approach to recursive utility has identified the recursive equation $\dot{u}_t = -W(c_t, u_t)$ as the Euler equation to a minimization problem. This Lagrange problem defines the recursive utility of a path by $U(c) = \min_r \int_0^\infty B_t F(c_t, r_t)\, dt$, where r_t is the rate of time preference and B_t the associated discount factor. The evolution of the recursive discount rate has been shown to obey some "law of least-action."

The capital accumulation model, where investment is controlled to maximize a recursive utility objective for consumption, was characterized by a true Hamiltonian system. Standard, additive utility maximization problems were embedded in a bigger space, and revealed a richer structure which was previously hidden in modified Hamiltonian systems.

Recursive utility functions belong to the more general class of variational utility functions, where the minimum principle still holds, and for which therefore optimal capital paths are characterized by the same Hamiltonian dynamics. It may be asked what set of axioms concerning preferences would lead to such utility functions. These axioms would have to be weaker, as the model is more general than the recursive one. Intuitively, it is the time-consistency axiom set in the "recursive literature" (e.g., [22, 15]) which would be weakened; this suggests a new direction for research in addressing the so-called myopia problem [34].

References

1. K. Arrow, Alternative approaches to the theory of choice in risk-taking situations, *Econometrica* **19** (1951), 404–437.
2. E. J. Balder, Existence of optimal solutions for control and variational problems with recursive objectives, *J. Math. Anal. Appl.* **178** (1993), 418–437.
3. R. A. Becker and J. H. Boyd, III, Recursive utility and optimal capital accumulation. II. Sensitivity and duality theory, *Econ. Theory* **2** (1992), 547–563.
4. S. Bentami, B. Hallak, and H. Kallal, Interest rates risk immunization by linear programming, *in* "Proc. Coll. Internat. AFIR, 1990," pp. 223–239.
5. L. M. Benveniste and J. A. Scheinkman, Duality theory for dynamical optimal models of economy: The continuous time case, *J. Econ. Theory* **27** (1982), 1–19.
6. E. von Böhm-Bawerk, "Positive Theory of Capital," South Holland, Illinois, 1959.
7. D. Cass and K. Shell, The structure and stability of competitive dynamical systems, *J. Econ. Theory* **12** (1976), 31–70.
8. D. Cass and K. Shell, "Hamiltonian Dynamics in Economics," Academic Press, 1976.
9. J. Dermody and T. Rockafellar, A new approach to term structure for markets with shorting costs, *Math. Finance* **1** (1989), 67–80.
10. D. Duffie and L. Epstein, Stochastic differential utility, *Econometrica* **60** (1992), 353–394.
11. D. Duffie, L. Epstein, and C. Skiadas, "Infinite Horizon Stochastic Differential Utility," Working Paper, Graduate School of Business, Stanford University, June 1991.

12. D. Duffie, P.-Y. Geoffard, and C. Skiadas, Efficient and equilibria allocations with stochastic differential utility, *J. Math. Econ.* **23** (1994), 133–146.
13. I. Ekeland and R. Temam, "Analyse Convexe et Problèmes Variationnels," Dunod, Paris, 1974.
14. N. El Karoui, S. Peng, and M.-C. Quenez, "Backward Stochastic Differential Equations in Finance and Optimization," Working Paper, Université Paris IV, January 1994.
15. L. Epstein, The global stability of efficient intertemporal allocations, *Econometrica* **55** (1987), 329–355.
16. L. Epstein and A. Hynes, The rate of time preference and dynamic economic analysis, *J. Polit. Econ.* **91** (1983), 611–635.
17. L. Epstein and S. Zin, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* **57** (1989), 937–969.
18. R. Findlay, An "Austrian" model of international trade and interest rate equalization, *J. Polit. Econ.* **86** (1978), 989–1007.
19. I. Fischer, "The Theory of Interest," Macmillan, New York, 1930.
20. P.-Y. Geoffard, "Equilibres à Taches Solaires, Utilité Récursive: Deux Modèles Economiques Intertemporels," Thèse de doctorat, Université Paris IX, July 1991.
21. P.-Y. Geoffard, Utilité récursive: Existence d'un chemin optimal de consommation, *C.R. Acad. Sci.* **312** (1991), 657–660.
22. T. J. Koopmans, Stationary ordinal utility and impatience, *Econometrica* **28** (1960), 287–309.
23. R. Lucas, Jr., and N. Stokey, Optimal growth with many consumers, *J. Econ. Theory* **32** (1984), 139–171.
24. M. J. P. Magill, "On A General Economic Theory of Motion," Springer-Verlag, Berlin/New York, 1970.
25. M. Obstfeld, Macroeconomic policy, exchange-rate dynamics, and optimal asset accumulation, *J. Polit. Econ.* **89** (1981), 1142–1161.
26. F. Ramsey, A mathematical theory of savings, *Econ. J.* (1928), 543–559.
27. J. Rawls, "A Theory of Justice," Belknap, Cambridge, MA, 1971.
28. R. T. Rockafellar, "Conjugate Duality and Optimization," *SIAM Reg. Conf. Series in Appl. Math.* Vol. 16, SIAM, Philadelphia, 1974.
29. P. Romer, Cake eating, chattering, and jumps: Existence results for variational problems, *Econometrica* **54** (1986), 897–908.
30. J. A. Scheinkman, On optimal steady states of n-sector growth models when utility is discounted, *J. Econ. Theory* **12** (1976), 11–30.
31. J. A. Scheinkman, Stability of separable Hamiltonians and investment theory, *Rev. Econ. Stud.* **45** (1978), 559–570.
32. H. Simon, A behavioral model of rational choice, *Quart. J. Econ.* **69** (1954), 99–118.
33. G. Sorger, "A Hamiltonian Approach to Uzawa Utility Maximization," Working Paper, IEORST, University of Technology, Vienna, 1991.
34. R. H. Strotz, Myopia and inconsistency in dynamic utility maximization, *Rev. Econ. Stud.* **23** (1957), 165–180.
35. H. Uzawa, Time preference, the consumption function, and optimum asset holdings, *in* "Value, Capital, and Growth: Papers in Honor of Sir John Hicks" (J. N. Wolfe, Ed.), Edinburgh University, Edinburgh, 1968.