Economic Theory **22**, 447–455 (2003)

Economic Theory

Stochastic convexity in dynamic programming*

Alp E. Atakan

Department of Economics, Columbia University New York, NY 10027, USA
(e-mail: aea15@columbia.edu)

Received: February 23, 2001; revised version: May 19, 2002

Summary. This paper explores sufficient conditions for a continuous stationary Markov optimal policy and a concave value function in stochastic dynamic programming problems. Also, the paper addresses conditions needed for the differentiability of the value function. The paper uses conditions such as first order stochastic dominance, second order stochastic dominance and concave stochastic dominance that are widely applied in economics.

Keywords and Phrases: Dynamic programming, Stochastic dominance, Concave value function, Differentiable value function.

JEL Classification Numbers: C61, O41, D80, D90.

1 Introduction

Economists frequently look for continuous, concave and differentiable solutions to their models. Continuity and concavity are recurrent themes in economic theory. Differentiability enables the use of calculus to perform comparative statics exercises. The sufficient conditions needed for a continuous and differentiable solution are well established in static models as well as dynamic programming under certainty. As in many optimization problems monotonicity and convexity assumptions play an important role. Specifically concavity and monotonicity of the transition function is required to prove value function concavity in deterministic dynamic programming. Analogous sufficient conditions for stochastic dynamic programming

* I am deeply indebted to Prajit Dutta for patient assistance and advice. This paper has benefited from discussions with Tsz Cheong Lai, Tackseung Jun, Werner Stanzl and Satyajit Bose as well as feedback from an anonymous referee.

are less well articulated in the literature. Difficulties in the stochastic model arise since the definition of monotonicity and convexity for a transition probability is unclear.

The purpose of this paper is to investigate restrictions on the transition probability of a stochastic dynamic programming problem that ensure the existence of a continuous stationary Markov optimal policy function ("optimal policy") and a concave and differentiable value function. My approach is based on the concepts of first order ("FOSD"), second order ("SOSD") and concave ("CSD") stochastic dominance . I use these stochastic orders to define concavity and monotonicity for a transition probability and using these definitions provide the desired sufficient conditions.

Also, I analyze the implications of my concavity and monotonicity definitions in two common applications. The first application is the deterministic dynamic programming model where the state space is a subset of $\mathbb{R}$ – for example the one-sector model of optimal growth (see Stokey, Lucas and Prescott [12]). In this case I show my definition of concavity and monotonicity based on SOSD is exactly analogous to an increasing and concave transition function. The second application is when transitions have the form $h(s, a, r)$ where r is an exogenous shock. Many economic applications use this functional form. A prominent example is the stochastic growth model (see[12]). In this framework I prove that my definition is weaker than assuming that h is increasing and concave.

The economics literature on dynamic programming is extensive. Stokey, Lucas and Prescott [12] provide an excellent and exhaustive presentation of the earlier results. Amir [1] uses a condition related to SOSD to prove that the optimal policy is continuous and the value function is concave in a optimal growth problem. His results are a special case of mine. Dutta, Majumdar and Sundaram [5] examine conditions for parametric continuity in dynamic programming. In this paper they use FOSD monotonicity and concavity. As I show FOSD concavity is a restrictive assumption for deterministic transitions while the use of the weaker SOSD condition suffices. The results for the stochastic model presented in [12] are for a specific class of problems. They assume that the state space is partitioned into endogenous and exogenous stochastic components and the action is comprised of picking the next periods endogenous state variable. Their set-up allows them to avoid the need to define transition probability convexity and monotonicity.

The paper proceeds as follow: Section 2 provides the basic definitions and results for FOSD, SOSD and CSD, Section 3 discusses two applications; Section 4 outlines the dynamic programming problem and gives sufficient conditions for a continuous optimal policy and concave value function; finally Section 5 concludes.

2 Monotonicity and convexity of transition probabilities

In deterministic dynamic programming problems a transition function determines next period's state variable given the current state and choice of action. This transition function can be, for example, a production function. In many economic growth models the production function is assumed to be "neoclassical" – i.e. concave and monotonic. In this section I try to generalize the concept of monotonicity and concavity to the stochastic setting where the transition function is replaced by a transition probability.

The transition probability, q, is a conditional probability measure on the state space S, given $S \times A$ where A denotes the action space.[1] The transition probability maps $S \times A$ into the set of probability measures on S. In this framework $q(\cdot|s,a)$ is a probability measure on S and $q(H|\cdot,\cdot)$ is a measurable function of s and a for any Borel subset H of S. A dynamic programming model is termed deterministic when $q(s'|s,a) = 1$ if $s' = h(s,a)$ and $q(s'|s,a) = 0$ otherwise. In this case $h : S \times A \to S$ is the transition function. In deterministic dynamic programming concavity and monotonicity of h is required to prove optimal policy continuity and value function concavity (see [12]). In the general stochastic model monotonicity and concavity are also required for the same result. However, it is not immediately obvious what monotonicity or concavity means for a transition probability. I use conditions based on FOSD, CSD or SOSD as partial orders to define concavity and monotonicity.

Definition 1 (Stochastic Dominance).

(i) A probability measure $q_1 \in \Delta(S)$ first order stochastically dominates $q_2 \in \Delta(S)$ (q_1 $FOSD$ q_2) if $\int_S W(\xi)q_1(d\xi) \geq \int_S W(\xi)q_2(d\xi)$ for all continuous, bounded and increasing functions W.

(ii) Also, q_1 second order stochastically dominates q_2 (q_1 $SOSD$ q_2) if $\int_S W(\xi)\ q_1\ (d\xi) \geq \int_S W(\xi)q_2(d\xi)$ for all continuous, bounded, increasing and concave functions W.

(iii) Finally, q_1 concave stochastically dominates q_2 (q_1 CSD q_2) if $\int_S W\ (\xi)\ q_1\ (d\xi) \geq \int_S W\ (\xi)\ q_2\ (d\xi)$ for all continuous, bounded and concave functions W.

From these definitions it is immediately obvious that if q_1 FOSD q_2 then q_1 SOSD q_2 also if q_1 CSD q_2 then q_1 SOSD q_2. This shows that SOSD is the weakest among the three stochastic orders.

Having defined the three orders for arbitrary probability measures I can use them to specify monotonicity and concavity for transition probabilities. The following statement defines concavity and monotonicity for the transition probability.

Definition 2. *(i) A transition probability $q : S \times A \to \Delta(S)$ is termed FOSD (or SOSD or CSD)* **increasing** *if $q(\cdot|s',a)$ FOSD (or SOSD or CSD) $q(\cdot|s,a)$ when $s' \geq s$.*

(ii) Given (s_1,a_1), (s_2,a_2) and their convex combination (s_3,a_3) a transition probability q is termed FOSD (or SOSD or CSD) **concave** *if $q(\cdot|s_3,a_3)$ FOSD (or SOSD or CSD) $\lambda \cdot q(\cdot|s_1,a_1) + (1-\lambda)\cdot\ q(\cdot|s_2,a_2)$.*[2]

This specification of concavity and monotonicity is beneficial since the implications of these notions of stochastic dominance for probability measures have sharp

[1] I take A and S as convex Borel subsets of $\mathbb{R}^n$ with the related Borel $\sigma-$algebras for a measurable structure.

[2] (s_3,a_3) will refer to $\lambda \cdot (s_1,a_1) + (1-\lambda)\cdot\ (s_2,a_2)$ for the remained of this paper.

characterizations in terms of distribution functions.[3] For example q FOSD increasing and concave implies that the function $F(x|s,a)$ is decreasing in s and convex in s and a where $F(\cdot|s,a)$ denotes the distribution function associated with the transition probability. For a detailed discussion see [11].

3 Two applications

It is worthwhile discussing in greater detail the deterministic versions of the model when the state space is a subset of $\mathbb{R}$ which is a widely used economic application. The following proposition shows the implication of Definition 2 for deterministic transition functions. It highlights the insight that FOSD and CSD may be fairly restrictive.

Proposition 1. *If the model is deterministic and $S \subset \mathbb{R}$ then*

(i) the only transition function that is FOSD increasing and concave is a constant function with $h(s,a) = c \; \forall \; s,\, a$,
(ii) the transition function is SOSD increasing and concave if and only if it is concave in s and a and increasing in s,
(iii) the transition function is CSD concave if and only if it is linear.

Proof. In the deterministic model $\int_S W(\zeta)\, q(d\zeta|s,a) = W(h(s,a))$. (i) requires that $W(h(s,a))$ be increasing in s and concave in s and a for all increasing W. Since W can be chosen concave or convex it is possible to find an increasing function that will make the composition non-concave for any function h that is not constant. Assume that $h(s_3,a_3) < h(s_1,a_1)$ and without loss of generality $h(s_2,a_2) \le h(s_1,a_1)$. Define continuous, increasing and bounded W as follows:

$$W(x) = (x - h(s_3,a_3)) \cdot 1_{[h(s_3,a_3),h(s_1,a_1))}(x) + (h(s_1,a_1) - h(s_3,a_3)) \cdot 1_{[h(s_1,a_1),\infty)}(x) \tag{1}$$

In this case FOSD concavity is violated since $W(h(s_2,a_2)) \ge W(h(s_3,a_3))$ and $W(h(s_1,a_1)) > W(h(s_3,a_3))$. Consequently

$$h(s_3,a_3) \ge \max\{h(s_1,a_1), h(s_2,a_2)\} \; \forall s, a. \tag{2}$$

However Eq. (2) can not hold strictly in the interior of $S \times A$ hence Eq. (2) holds with equality. This shows that h is constant on $S \times A$. (ii) If h is concave then so is an increasing concave transformation. For the other direction assume that there are point such that $h(s_3,a_3) < \lambda \cdot h(s_1,a_1) + (1-\lambda) \cdot h(s_2,a_2)$. Now if we choose W to be increasing and linear then SOSD concavity is not satisfied. (iii) requires that $W(h(s,a))$ be concave in s and a for all concave W. However since W can be chosen increasing or decreasing the composition can be made non-concave for non-linear h. For the other direction see the proof of Proposition 2 (iii) .[4] □

---

[3] F and G are distribution functions and $s \in \mathbb{R}$. F FOSD G if and only if $F(x) \le G(x)\; \forall x$. F SOSD G if and only if $\int_{-\infty}^{z} F(x)\, dx \le \int_{-\infty}^{z} F(x)\, dx \; \forall z$ and $\mathbb{E}_F[s] \ge \mathbb{E}_G[s]$. F CSD G if and only if $\int_{-\infty}^{z} F(x)\, dx \le \int_{-\infty}^{z} F(x)\, dx \; \forall z$ and $\mathbb{E}_F[s] = \mathbb{E}_G[s]$.

[4] Full proofs are available from the author on request.

In many economic applications the next period's state variable is taken to be a function of the current state s, the action a and an exogenous shock r with distribution function G i.e. $s' = h(s, a, r)$.[5] Concavity and monotonicity assumptions are then made on the function h. A prominent example is the stochastic optimal growth where h is a production function and r can be an additive or a multiplicative shock (see [12]). The following proposition relates assumptions on h to assumptions on the associated transition probability.

Proposition 2. *Assume that $s' = h(s, a, r)$ where h is a Borel function and r is a random variable.*

(i) If h is increasing in s then the associated transition probability is FOSD increasing;
(ii) if h is concave in s and a then the associated transition probability is SOSD concave;
(iii) if h is linear in s and a then the associated transition probability is CSD concave.

Proof. Observe that $\int_S W(s')\, q(s'|s,a) = \int_{\mathbb{R}} W(h(s,a,r))\, dG$. (i) since h is increasing in s for all r so is $W \circ h$ for increasing W. (ii) If h is concave in s and a for all r then so is $W \circ h$ for increasing and concave W. (iii) If h is linear in s and a and W is concave then $W \circ h$ is also concave in s and a. □

Interestingly the converse to Proposition 2 (ii) is not true even when $S \subset \mathbb{R}$. As demonstrated in Example 1, SOSD concavity does not imply that h is concave. This shows that the condition used in this paper is weaker than the assumption which is usual in the economics literature. In many applications h is taken as concave and increasing. For example in [12] the production function in the stochastic growth model is assumed to be concave and increasing to prove value function concavity. As I show, the weaker condition of SOSD concavity and monotonicity actually suffices. The following provides a counter example to the converse of Proposition 2 (ii).

Example 1. Assume that h only depends on s. The state space is taken as $[0, 2]$ and r is a Bernoulli random variable. The function $h(s, r)$ is defined as follows:

$$h(s,r) = \begin{cases} \min\left\{s^2+1, 2\right\} \text{ with probability } 1-p \\ \sqrt{s+0.1} \text{ with probability } p. \end{cases} \tag{3}$$

In this example the transition probability is SOSD increasing and concave for large p however the function h is not concave in s.[6]

Proof. SOSD increasing follows since h is increasing. Let $F(x|s)$ denote the conditional distribution function for s'.

[5] For a Borel subset B of S let $A(s,a) = \{r : h(s,a,r) \in B\}$. Now define $q(B|s,a) = \int_{-\infty}^{+\infty} 1_{A(s,a)}(r)\, dG(r)$ which is the associated transition probability.

[6] This example is based on an example in [1].

$$F(x|s,a) = \begin{cases} 0 \text{ if } x < \sqrt{s+0.1} \\ p \text{ if } \sqrt{s+0.1} \leq x < \min\left\{s^2+1, 2\right\} \\ 1 \text{ if } \min\left\{s^2+1, 2\right\} \leq x \end{cases} \tag{4}$$

Showing that F is SOSD concave is equivalent to showing that $\int_0^z F(x|s)\,dx$ is a convex function of s for all z for large p (see [11]). Which follows since

$$\int_0^z F(x|s)\,dx = \begin{cases} 0 \text{ if } z < \sqrt{s+0.1} \\ p\left(z - \sqrt{s+0.1}\right) \text{ if } \sqrt{s+0.1} \leq z < \min\left\{s^2+1, 2\right\} \\ z - (1-p)\min\left\{s^2+1, 2\right\} - p\sqrt{s+0.1} \text{ otherwise} \end{cases} \tag{5}$$

is a convex function when p is large. Observe that the function is convex for all p except when $z \geq \min\left\{s^2+1, 2\right\}$ and $s \in [0, 1)$. However in this range the second derivative is always positive for $p > 0.91$. $\square$

4 Application to dynamic programming

4.1 Dynamic programming preliminaries

The dynamic programming problem is a collection $\langle S, A, \pi, \phi, T, q\rangle$ where q is the transition probability, S is the state space, A is the action space, $\pi : S \times A \to \mathbb{R}$ is the immediate reward function, $\phi : S \to \wp(A)$ is the feasible action correspondence, $\delta \in (0, 1)$ is the discount rate and $T = \infty$ is the time horizon. At each point in time the decision-maker chooses an action from the set of permitted actions, receives a payoff $\pi(s, a)$ and the system moves to a new state according to the transition probability.

A history at time t, denoted h_t, is a sequence $\{s_0, a_0, ..., s_t\}$. H_t is the set of all possible histories at time t. A policy, $\sigma \in \Sigma$, is defined as a sequence $\{f_t\}$ where, $f_t : H_t \to A$ is a measurable function and Σ is the set of all policies. A policy is Markov if $f_t : S \to A$ for all t. The sequence is termed stationary Markov if, in addition, $f_t = f$ for all t. For each policy σ, there is an associated function $I_\sigma : S \to \mathbb{R}$. This function gives the total expected discounted reward from employing policy σ given the initial state s_0 i.e., $I_\sigma(s_0) = E_\sigma[\sum_{t=0}^{\infty} \delta^t \pi(s_t, a_t)|s_0]$. The decision-maker aims to choose the policy $\sigma \in \Sigma(s_0)$ that maximizes $I_\sigma(s_0)$ where $\Sigma(s_0)$ denotes the set of feasible policies given the initial state. The value function, $V : S \to \mathbb{R}$, is defined as the supremum of the discounted reward functions, $V(s_0) = \sup_{\sigma\in\Sigma(s_0)} I_\sigma(s_0)$.

In this paper I assume that S and A are convex subsets of $\mathbb{R}^n$, π is continuous and bounded, $\phi : S \to \wp(A)$ is continuous and compact valued and $q(\cdot|s, a)$ has the Feller property, i.e. if $(s_n, a_n) \to (s, a)$ then

$$\int_S W(\xi)q(d\xi|s_n, a_n) \to \int_S W(\xi)q(d\xi|s, a) \tag{6}$$

for all continuous and bounded W. For a function W on S define the operators L and T as follows:

$$L_a W(s) = \pi(s,a) + \delta \int W(\xi) q(d\xi | s,a) \tag{7}$$

$$TW(s) = \sup_{a \in \phi(s)} \{L_a W(s)\} \tag{8}$$

Under the assumptions mentioned above it is shown in [3], [7] and [8] that an optimal policy exists and the continuous, bounded value function is the unique fixed point of the operator T with $V = TV$. Also, the set of maximizers, $F(s) = \arg\max_{a \in \phi(s)} \{L_a V(s)\}$, is an upper-hemi-continuous correspondence and any measurable selection f from F is a stationary Markov optimal policy.

4.2 Main results

In deterministic dynamic programming if $\pi(s,a)$ is strictly concave in s and a and increasing in s, the feasible action correspondence is increasing and convex graph and the transition function is continuous, bounded and concave then the value function is continuous, bounded, concave and increasing and the optimal policy is continuous (see [12]).[7] The main idea when trying to generalize the result to the stochastic case is to ensure that the operator T, as defined in Eq. (8), is monotone and concave. The following proposition generalizes the result using Definition 2.

Theorem 1. *Assume that π is strictly concave in s and a and increasing in s, and ϕ is increasing and convex graph. If q is SOSD increasing and concave then the value function is a continuous, bounded, increasing and strictly concave function of s and the optimal policy is continuous.*

Proof. Let $C(S)$ denote the set of continuous, bounded, concave and increasing functions. From previous results we know that if W is continuous and bounded then so is TW. Assume $W \in C(S)$ and let $TW(s_1) = L_{a_1}W(s_1)$, $TW(s_2) = L_{a_2}W(s_2)$. Since π is concave and q SOSD concave $TW(s_3) \geq L_{a_3}W(s_3) > \lambda \cdot L_{a_1}W(s_1) + (1-\lambda) \cdot L_{a_2}W(s_2)$ which shows that the operator T preserves concavity. Also $TW(s_1) \geq L_{a_2}W(s_1) \geq L_{a_2}W(s_2) = TW(s_2)$ if $s_1 \geq s_2$ since π is increasing and q is SOSD increasing which shows that the operator T preserves monotonicity. $C(S)$ is a Banach space, $T : C(S) \to C(S)$ is a contraction and the value function satisfies $TV = V$. Hence the value function is the unique fixed point of T and an element of $C(S)$. However since $\pi$is strictly concave so is V. Since $L_a V \in C(S)$ and strictly concave then by Berge's Maximum Theorem $F(s) = \arg\max_{a \in \phi(s)} L_a W(s)$ is a continuous function. $\square$

Corollary 1. *Under the assumptions of Theorem 1 if q is FOSD increasing and concave then the value function is continuous, bounded, increasing and strictly concave.*

[7] ϕ is increasing if $s' \geq s$ implies $\phi(s') \sqsupseteq \phi(s)$ and ϕ is convex graph if $graph(\phi) = \{(s,a) \in S \times A | a \in \phi(s)\}$ is a convex set.

Proof. q FOSD increasing and concave implies that q is SOSD increasing and concave. □

As I stated earlier SOSD increasing and concave is equivalent to an increasing and concave transition function in the deterministic model. Consequently Theorem 1 can be viewed as an exact analog to the deterministic result. The corollary shows that concavity and monotonicity based on the FOSD used in [5] implies the much weaker SOSD condition.

Theorem 2. *Assume that π is strictly concave in s and a, and ϕ is convex graph. If q is CSD concave then the value function is a continuous, bounded and strictly concave function of s and the optimal policy is continuous.*

Proof. Let $B(S)$ denote the set of continuous, bounded and concave functions. Assume $W \in B(S)$ and let $TW(s_1) = L_{a_1}W(s_1), TW(s_2) = L_{a_2}W(s_2)$. Since π is concave and q CSD concave $TW(s_3) \geq L_{a_3}W(s_3) > \lambda \cdot L_{a_1}W(s_1) + (1 - \lambda) \cdot L_{a_2}W(s_2)$ which shows that the operator T preserves concavity. The rest is identical to the proof of Theorem 1. □

Theorem 2 is an improvement on Theorem 1 since we can dispense with the monotonicity assumptions. However this comes at a cost: we have to impose a more stringent (CSD) concavity requirement on q. As I showed earlier the CSD concavity assumption is essentially a linearity requirement on the transition function.

Value function concavity and differentiability are closely linked. Results concerning the differentiability of a concave value function were first presented in [2]. Under the assumptions in Theorem 1 (a) or (b) it is straight forward to give sufficient conditions for a differentiable value function. For example if both A and S are compact subsets of $\mathbb{R}^n$ and ϕ is a constant correspondence with $\phi(s, a) = A$ for all s and a then the result in [2] implies that the value function is continuously differentiable on S.

5 Conclusion

In this paper I defined monotonicity and concavity for a transition probability using some well-known stochastic orders and applied these definitions to dynamic programming. I proved the existence of a continuous optimal policy and a concave and differentiable value function using these definitions. I also outlined the implication of my definitions of concavity in the deterministic univariate case and when the transition function takes the form $h(s, a, r)$. I proved that SOSD concavity and monotonicity is weaker than assuming that h is increasing and concave and exactly analogous to an increasing and concave deterministic transition function.

References

1. Amir, R.: A new look at optimal growth under uncertainty. Journal of Economic Dynamics and Control **22**, 67–86 (1997)

2. Benveniste, L. M., Scheinkman, J. A.: On the differentiability of the value function in dynamic models of economics. Econometrica **47**, 727–732 (1979)
3. Blackwell, D.: Discounted dynamic programming. Annals of Mathematical Statistics **36**, 226–235 (1965)
4. Burdett, K., Mortensen, D.: Search, layoffs and labor market equilibrium. Journal of Political Economy **88**, 652–672 (1980)
5. Dutta, P., Majumdar, M., Sundaram, R.: Parametric continuity in dynamic programming problems. Journal of Economic Dynamics and Control **18**, 1069–1092 (1994)
6. Folland, G.: Real analysis: modern techniques and their applications. New York: Wiley 1999
7. Furukawa, N.: Markovian decision processes with compact action spaces. Annals of Mathematical Statistics **43**, 1612–1622 (1972)
8. Maitra, A.: Dynamic programming for compact metricspaces. Sankhya Series A **30**, 211–216 (1968)
9. Micheal Rothschild and Joseph Stiglitz. Increasing risk: I. a definition. Journal of Economic Theory **2**, 225–243 (1970)
10. Rothschild, M., Stiglitz, J.: Increasing risk: Ii. its economic consequences. Journal of Economic Theory **3**, 66–84 (1971)
11. Shaked, M., Shanthikumar, G.: Stochastic orders and their applications. New York: Academic Press 1994
12. Stokey, N., Lucas, R., Prescott, E.: Recursive methods in economic dynamics. Cambridge, MA: Harvard University Press 1989