Necessity of the Transversality Condition for Stochastic Models with Bounded or CRRA Utility

Takashi Kamihigashi[*]

RIEB, Kobe University
Rokkodai, Nada, Kobe 657-8501, Japan

October 8, 2004

Abstract

This paper shows that the standard transversality condition (STVC) is necessary for optimality in stochastic models with bounded or constant-relative-risk-aversion (CRRA) utility under fairly general conditions. We consider an infinite-horizon stochastic maximization problem that takes a general form of a multi-sector growth model with a single consumption good. We show that the STVC is necessary if utility is bounded or logarithmic. We also show that the STVC is necessary in the case of non-logarithmic CRRA utility as long as lifetime utility is finite at the optimum. These results apply to various stochastic growth models, including real business cycle models with endogenous labor supply. Since unbounded utility functions that do not belong to the CRRA class are rather rare in applications, our results provide a fairly complete set of solutions regarding necessity of the STVC in practice.

Keywords: Transversality condition, stochastic optimization, stochastic growth, CRRA, real business cycle.

JEL Classification Numbers: C61, D90, G12

[*]Tel./fax: +81-78-803-7015, Email address: tkamihig@rieb.kobe-u.ac.jp

1 Introduction

Since Shell (1969) and Halkin (1974), necessity of transversality conditions (TVCs) has been an uneasy matter to economic theorists who use infinite-horizon optimization problems. The most standard TVC, which we call the standard TVC (STVC), is the condition that the value of optimal stocks at infinity must be zero. Necessity of the STVC becomes an issue particularly for models with unbounded utility, such as constant-relative-risk-aversion (CRRA) utility.

Models with CRRA utility are widely used in macroeconomics, for example, in the endogenous growth literature (e.g., Barro and Sala-i-Martin, 1995), the indeterminacy literature (e.g., Benhabib and Farmer, 1999), and the real business cycle (RBC) literature (e.g., King and Rebelo, 1999). Although some results on necessity of the STVC for models with CRRA utility are available in the literature, they only deal with deterministic cases (Ekeland and Scheinkman, 1986; Alvarez and Stokey, 1998; Kamihigashi, 2000, 2001).

This paper shows that the STVC is necessary in stochastic models with bounded or CRRA utility under fairly general conditions. In particular, we consider an infinite-horizon stochastic maximization problem that takes a general form of a multi-sector growth model with a single consumption good. Our model encompasses various stochastic growth models (e.g., Brock and Mirman, 1972), including RBC models with endogenous labor supply (e.g., King and Rebelo, 1999). Our assumptions on the technology side of the model are quite general. They require only a weak concavity property and some technical conditions that are rarely violated in applications.

We establish three results in this paper. Our first result is that the STVC is necessary if utility is bounded. This of course means that the STVC may fail to be necessary only if utility is unbounded. Hence we next turn to the case of unbounded utility. We restrict attention to the CRRA class since unbounded utility functions that do not belong to this class are rather rare in applications. Our second result is that the STVC is necessary in the case of logarithmic utility. This is a strong result since it does not require any additional condition (other than the technical assumptions on the technology side of the model). Our last result is that the STVC is necessary in the case of non-logarithmic CRRA utility as long as lifetime utility is finite *at the optimum*. It is important to emphasize that we do *not* assume the finiteness of lifetime utility for all feasible paths, for such an assumption is often violated

in models with unbounded utility.

Since these results are established for a general stochastic growth model, they generalize similar results shown for deterministic models (e.g., Kamihigashi, 2000, 2001). Though our results are based on a general theorem shown in Kamihigashi (2003), they are surprisingly clean and powerful. Due to their general applicability, they provide a fairly complete set of solutions regarding necessity of the STVC in practice since utility is usually assumed to be bounded or of the CRRA class in applications. It is useful to know that the STVC is necessary in these cases, not only because it helps characterize optimal paths, but also becasue it has important implications for the stability properties of balanced growth paths, as recently shown by Martínez-García (2003).

The rest of the paper is organized as follows. Section 2 shows how the Bellman equation is related to the STVC, examines two examples in which the STVC is not necessary, and discusses related results in the literature. Section 3 presents our general model and discusses our assumptions. Our main results are stated in Section 4. The proofs appear in Appendix B. Section 5 explains how one can apply our results to models with endogenous labor supply such as RBC models. Section 6 discusses how one can apply our results to other models with control variables. Section 7 concludes the paper. Appendix A contains preliminary lemmas.

2   Informal Discussions

This section clarifies some of the issues regarding necessity of the STVC.[1] In particular, we show how the Bellman equation is related to the STVC, and discuss when the STVC possibly fails to be necessary by examining two simple examples. Related results in the literature are also discussed. Though this paper is concerned with stochastic models, most of the known issues on the STVC are deterministic in nature. Hence for the most part we consider only deterministic models in this section.

---

[1]This section owes much to Editor Woulter J. den Haan's suggestions.

2.1 The Squeezing Argument

Consider the following maximization problem:

$$\max_{\{x_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t v(x_t, x_{t+1}) \quad \text{s.t. } x_0 = \overline{x}_0, \quad \forall t \in \mathbb{N}, x_{t+1} \geq 0, \tag{1}$$

where $\beta \in (0,1)$ is the discount factor, v is assumed to be nonincreasing in the second argument and concave, and $\overline{x}_0$ is given exogenously. There may be other constraints, but we assume that they are not binding at the optimum.

For the moment we assume that the objective function, or lifetime utility, is finite for all feasible paths. Then we may assume without loss of generality that

$$v(0,0) \geq 0. \tag{2}$$

Essentially it has been known since Weitzman (1973) that in this case, any (interior) optimal path $\{x_t^*\}$ satisfies

$$\lim_{t \uparrow \infty} \beta^t [-v_2(x_t^*, x_{t+1}^*)] x_{t+1}^* = 0. \tag{3}$$

This condition means that the value of optimal stocks must be zero at infinity. We call (3) the standard TVC (STVC).[2]

Perhaps the easiest way to derive (3) is to consider the Bellman equation

$$V(x_t^*) = v(x_t^*, x_{t+1}^*) + \beta V(x_{t+1}^*). \tag{4}$$

The finiteness of lifetime utility for all feasible paths implies

$$\lim_{t \uparrow \infty} \beta^t V(x_t^*) = 0. \tag{5}$$

Assuming interiority, one obtains the Euler equation

$$v_2(x_t^*, x_{t+1}^*) + \beta v_1(x_{t+1}^*, x_{t+2}^*) = 0. \tag{6}$$

By the envelope theorem,

$$V'(x_t^*) = v_1(x_t^*, x_{t+1}^*). \tag{7}$$

---

[2]An alternative form of the condition, $\lim_{t \uparrow \infty} \beta^t v_1(x_t^*, x_{t+1}^*) x_t^* = 0$, is often used. Though this condition is equivalent to (3) under the Euler equation (6), it has no counterpart in the finite-horizon case as well as in the continuous-time case.

From (6) and (7),

$$-v_2(x_t^*, x_{t+1}^*)x_{t+1}^* = \beta V'(x_{t+1}^*)x_{t+1}^* \le \beta V(x_{t+1}^*), \qquad (8)$$

where the last inequality holds since $V(0) \ge 0$ by (2) and V is concave by the concavity of v. Since v is nonincreasing in the second argument, the leftmost side of (8) is nonnegative. Now TVC (3) follows from (8) and (5) by the "squeezing" argument.[3]

2.2 Counterexamples

The crucial conditions used in the above argument are (2) and (5), which are obtained from the assumed finiteness of lifetime utility for all feasible paths. If this assumption is violated, however, the STVC may fail to be necessary for optimality. Let us study two simple examples in which the STVC is indeed not necessary.

We first consider the discounted Ramsey model:

$$\left. \begin{array}{ll} \max_{\{x_t\}} & \sum_{t=0}^{\infty} \beta^t u(f(x_t) - x_{t+1}), \\ \text{s.t.} & x_0 = \overline{x}_0, \quad \forall t \in \mathbb{Z}_+, 0 \le x_{t+1} \le f(x_t), \end{array} \right\} \qquad (9)$$

where u is the utility function, and f is the production function. We assume that f satisfies the standard neoclassical properties. Since the main focus of this paper is on the CRRA case, let us assume

$$\exists \alpha \in (-\infty, 1] \setminus \{0\}, \quad u(c) = \frac{c^{\alpha}}{\alpha}. \qquad (10)$$

It is well-known that in the discounted Ramsey model, the optimal path $\{x_t^*\}$ from any nonzero initial capital stock converges to the modified golden rule capital stock x^{β}, which is given by $\beta f'(x^{\beta}) = 1$. Hence

$$\lim_{t \uparrow \infty} \beta^t u'(f(x_t^*) - x_{t+1}^*)x_{t+1}^* = \lim_{t \uparrow \infty} \beta^t u'(f(x^{\beta}) - x^{\beta})x^{\beta}. \qquad (11)$$

The right-hand side is of course zero since $\beta \in (0, 1)$. Thus the STVC holds.

It is also well-known that optimal paths converge to the golden rule capital stock x^1 when $\beta = 1$. In this case, however, the right-hand side of (11) is

---

[3]See Kamihigashi (2002) for an elementary proof that does not use the envelop theorem.

strictly greater than zero, so that the STVC is violated. Perhaps partly for this reason, the violation of the STVC is often associated with the absence of discounting. But the actual factor that determines necessity of the STVC is not whether discounting is present or not.

To see this point clearly, consider the following exogenous growth model:

$$\left.\begin{array}{ll} \max\limits_{\{x_t\}} & \sum\limits_{t=0}^{\infty} \beta^t u(\theta^t(x_t)^\gamma - x_{t+1}), \\[2mm] \text{s.t.} & x_0 = \overline{x}_0, \quad \forall t \in \mathbb{Z}_+, 0 \le x_{t+1} \le \theta^t(x_t)^\gamma, \end{array}\right\} \tag{12}$$

where $\theta > 1$ is the growth factor, and $\gamma \in (0,1)$. Let $g = \theta^{1/(1-\gamma)}$. Note that g satisfies

$$\theta g^\gamma = g. \tag{13}$$

For $t \in \mathbb{Z}_+$, define

$$\tilde{x}_t = x_t/g^{t-1}. \tag{14}$$

The maximization problem (12) can now be written as follows:[4]

$$\left.\begin{array}{ll} \max\limits_{\{x_t\}} & \sum\limits_{t=0}^{\infty} \tilde{\beta}^t u(g^{-\gamma}(\tilde{x}_t)^\gamma - \tilde{x}_{t+1}), \\[2mm] \text{s.t.} & \tilde{x}_0 = g\overline{x}_0, \quad \forall t \in \mathbb{Z}_+, 0 \le \tilde{x}_{t+1} \le g^{-\gamma}(\tilde{x}_t)^\gamma, \end{array}\right\} \tag{18}$$

where $\tilde{\beta} = \beta g^\alpha$. If $\tilde{\beta} = 1$, this problem is a special case of the undiscounted Ramsey model (i.e., (9) with $\beta = 1$). Hence in this case, the STVC is not necessary for optimality. But the original problem (12) has strictly positive discounting since $\tilde{\beta} = 1$ is equivalent to $\beta = g^{-\alpha} < 1$.

What is common in the preceding two "counterexamples" is not that there is no discounting, but that lifetime utility is not finite at the optimum. Note that the finiteness of lifetime utility at the optimum is a much weaker assumption than the finiteness of lifetime utility for all feasible paths. In fact,

---

[4]To see this, note that

$$\theta^t(x_t)^\gamma - x_{t+1} = \theta^t(g^{t-1}\tilde{x}_t)^\gamma - g^t\tilde{x}_{t+1} \tag{15}$$

$$= (\theta g^\gamma)^t g^{-\gamma}(\tilde{x}_t)^\gamma - g^t\tilde{x}_{t+1} \tag{16}$$

$$= g^t[g^{-\gamma}(\tilde{x}_t)^\gamma - \tilde{x}_{t+1}], \tag{17}$$

where the last equality uses (13).

if $\alpha < 0$, the latter assumption is violated in the above models regardless of β, even when lifetime utility is finite at the optimum.

It can therefore be expected that the sharpest result one can hope to prove is that the STVC is necessary in the CRRA case if lifetime utility is finite at the optimum. This is exactly what we show in Section 4 for a general stochastic growth model. This result generalizes similar results known for deterministic models with CRRA utility (e.g., Kamihigashi 2000, 2001).

It should be mentioned that the finiteness of lifetime utility at the optimum is a meaningless assumption unless one focuses on a specific class of utility functions. In fact, it can always be made true by subtracting an appropriate constant from the utility function, as is usually done in the undiscounted Ramsey model. This type of normalization is prohibited by (10), so that the finiteness of lifetime utility is a meaningful assumption here.

2.3   Related Results in the Literature

The argument in Section 2.1 is essentially due to Benveniste and Scheinkman (1982, p. 15), but Weitzman (1973) used a similar argument based on support prices instead of derivatives. The difficult part of the actual proof is to show (7).[5] The key assumptions used in the argument are (2) and (5), as mentioned above.

The first attempt to relax (2) was made by Ekeland and Scheinkman (1986), who studied a model with CRRA-type utility. Instead of assuming the finiteness of lifetime utility for all feasible paths, they assumed that the feasible sequences of utility are bounded above by a common summable sequence. This assumption was relaxed by Kamihigashi (2000) to the finiteness of lifetime utility at the optimum.

Michel (1990) extended Weitzman's (1973) analysis to a discrete-time optimal control problem with possibly unbounded returns. Instead of the STVC, however, he focused on TVCs that an optimal path must satisfy against all feasible paths that do not incur an infinite loss.

Kamihigashi (2001) unified the approaches of Ekeland and Scheinkman (1986) and Michel (1990), generalizing earlier results and obtaining new results in a continuous-time reduced-form model. The basic idea of Kamihigashi (2001) is that the STVC is necessary if the optimal path can be shifted proportionally downward without causing an infinite loss. See Kamihigashi

---

[5]See Stokey and Lucas (1989, p. 85) for a proof of (7).

(2001) for a more detailed literature review on necessity of TVCs for deterministic models.

Stochastic versions of Weitzman's (1973) result were obtained by Zilcha (1976) and Takekuma (1992) using infinite-dimensional separation arguments.[6] Their results on the STVC were generalized by Kamihigashi (2003) for a discrete-time stochastic reduced-form model using the approach of Kamihigashi (2001).

As suggested in the Introduction, the contribution of this paper is to provide a fairly complete set of solutions regarding necessity of the STVC for stochastic growth models in applications. Though our results are based on a general theorem shown in Kamihigashi (2003), they are more directly applicable. Indeed, since unbounded utility functions that do not belong to the CRRA class are rather rare in applications, our results seem to cover most cases in practice.

3   The Model

This section presents our general model and discusses our assumptions. Some of the material here is borrowed from Kamihigashi (2003), which offers further discussions on the general structure of the model.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let E denote the associated expectation operator; i.e., $Ez = \int z(\omega)dP(\omega)$ for any random variable $z : \Omega \to \overline{\mathbb{R}}$. When it is important to make explicit the dependence of z on ω, we write $Ez(\omega)$ instead of Ez. Consider the following problem:

$$
\left.
\begin{array}{rl}
\text{``} \max\limits_{\{x_t\}_{t=0}^\infty} & \sum\limits_{t=0}^\infty \beta^t Eu(g_t(x_t(\omega), x_{t+1}(\omega), \omega))\text{''} \\[2mm]
\text{s.t.} & x_0 = \overline{x}_0, \\[1mm]
& \forall t \in \mathbb{Z}_+, (x_t, x_{t+1}) \in X_t.^7
\end{array}
\right\}
\tag{19}
$$

In what follows, we list and discuss all our assumptions.

Let $n \in \mathbb{N}$.[8] Let F be the set of all functions from Ω to $\mathbb{R}^n$. The following assumption means that each x_t is a random variable in $\mathbb{R}^n$.

---

[6] Foldes (1978) extended Weitzman's result to a one-sector stochastic growth model using directional derivatives and martingale arguments.

[7] $\mathbb{Z}_+ \equiv \{0, 1, 2, \ldots\}$.

[8] $\mathbb{N} \equiv \{1, 2, 3, \ldots\}$.

Assumption 3.1. $\overline{x}_0 \in F$ and $\forall t \in \mathbb{Z}_+, X_t \subset F \times F$.

Since F consists of all functions from Ω to $\mathbb{R}^n$, and since X_t is only a subset of $F \times F$, X_t can be chosen in such a way that x_t and x_{t+1} must be measurable with respect to the information available in period t and period $t + 1$, respectively. Restricting X_t this way is not necessary, however, since our results require no such information structure.

The next assumption collects our general restrictions on the preference side of the model.

Assumption 3.2. (i) $\beta \in (0, 1)$; and (ii) $u : \mathbb{R}_+ \to [-\infty, \infty)$ is C^1 on $\mathbb{R}_{++}$, concave, and strictly increasing.

The next assumption simply means that the expression

$$Eu(g_t(x_t(\omega), x_{t+1}(\omega), \omega)) \tag{20}$$

makes sense for all $(x_t, x_{t+1}) \in X_t$.

Assumption 3.3. $\forall t \in \mathbb{Z}_+, \forall (y, z) \in X_t$, (i) $\forall \omega \in \Omega, g_t(y(\omega), z(\omega), \omega) \geq 0$, (ii) the mapping $g_t(y(\cdot), z(\cdot), \cdot) : \Omega \to \mathbb{R}_+$ is measurable, and (iii) the expectation $Eu(g_t(y(\omega), z(\omega), \omega))$ exists in $[-\infty, \infty)$.

To simplify expressions like (20), let $g_t(x_t, x_{t+1})$ denote the random variable $g_t(x_t(\cdot), x_{t+1}(\cdot), \cdot) : \Omega \to \mathbb{R}$ for $(x_t, x_{t+1}) \in X_t$. We say that a sequence $\{x_t\}_{t=0}^{\infty}$ is a *feasible path* if $x_0 = \overline{x}_0$ and $\forall t \in \mathbb{Z}_+, (x_t, x_{t+1}) \in X_t$. Since in applications the objective function is often not guaranteed to be finite or well-defined for all feasible paths, we use weak maximality (Brock, 1970) as our optimality criterion. We say that a feasible path $\{x_t^*\}$ is *optimal* if for any feasible path $\{x_t\}$,

$$\varlimsup_{T \uparrow \infty} \sum_{t=0}^{T} \beta^t [Eu(g_t(x_t, x_{t+1})) - Eu(g_t(x_t^*, x_{t+1}^*))] \leq 0.^9 \tag{21}$$

Our optimality criterion (i) reduces to the standard maximization criterion whenever the latter makes sense, (ii) applies even when the standard

---

[9] To be precise, this inequality requires that the left-hand side be well-defined. This means that the left-hand side may not involve expressions like "$\infty - \infty$" and "$-\infty + \infty$." Thus optimality implies that $Eu(g_t(x_t^*, x_{t+1}^*))$ is finite for all $t \in \mathbb{Z}_+$; for otherwise the left-hand side of (21) is undefined for $\{x_t\} = \{x_t^*\}$.

criterion fails, and (iii) is weaker than the similar criterion with $\overline{\lim}$ replacing $\underline{\lim}$ in (21). Using this optimality criterion does not limit the applicability of our results in any way since we are only interested in necessary conditions for optimality.

For the same reason, assuming the existence of an optimal path imposes no restriction on the model.

Assumption 3.4. There exists an optimal path $\{x_t^*\}$.

For $t \in \mathbb{Z}_+$ and $d : \Omega \to \mathbb{R}^n$ such that $(x_t^*, x_{t+1}^* + \epsilon d) \in X_t$ for sufficiently small $\epsilon > 0$, we define the random variable $g_{t,2}(x_t^*, x_{t+1}^*; d)$ as follows:

$$g_{t,2}(x_t^*, x_{t+1}^*; d) = \underline{\lim_{\epsilon \downarrow 0}} \frac{g_t(x_t^*, x_{t+1}^* + \epsilon d) - g_t(x_t^*, x_{t+1}^*)}{\epsilon}, \tag{22}$$

where $\underline{\lim}_{\epsilon \downarrow 0}$ is applied pointwise (i.e., for each $\omega \in \Omega$ separately). The right-hand side of (22) is called a lower Dini directional derivative. Note that if g_t is differentiable in the second argument, then

$$g_{t,2}(x_t^*, x_{t+1}^*; d) = D_2 g_t(x_t^*, x_{t+1}^*) d, \tag{23}$$

where $D_2 g_t$ denotes the derivative of g_t with respect to the second argument. We use the general derivative defined in (22) only because assuming differentiability does not offer any simplification and, at the same time, requires additional assumptions that are somewhat cumbersome to state.

As mentioned in Section 2.1, the STVC is the condition that the value of optimal stocks at infinity must be zero. It takes the following form in the current framework:

$$\lim_{t \uparrow \infty} \beta^t E u'(g_t(x_t^*, x_{t+1}^*)) g_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) = 0. \tag{24}$$

Recalling (23), one can see that (24) reduces to the familiar TVC (e.g., (3)) in the differentiable case. The following assumption is a minimum requirement for (24) to make sense.

Assumption 3.5. The mapping $g_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) : \Omega \to \overline{\overline{\mathbb{R}}}$ is measurable.

It is difficult to think of a situation in which this assumption is violated. It must be assumed nonetheless since the other assumptions do not imply it. It is satisfied, for example, if g_t depends on ω through a technology shock

and is continuously differentiable in x_t, x_{t+1}, and the technology shock (and if these variables are measurable).

The assumptions stated thus far are preliminary requirements that are in many cases assumed implicitly. In what follows, we state three assumptions that are actually needed to prove the necessity of the STVC. Each of them is fairly general though it has a specific role. The following is our assumption on X_t.

Assumption 3.6. $\forall t \in \mathbb{Z}_+, \exists \underline{\lambda}_t \in [0, 1), \forall \lambda \in [\underline{\lambda}_t, 1), (x_t^*, \lambda x_{t+1}^*) \in X_t$ and $\forall \tau \geq t + 1, (\lambda x_\tau^*, \lambda x_{\tau+1}^*) \in X_\tau$.

This assumption means that the optimal path $\{x_t^*\}$ can be shifted proportionally downward starting from any period. The assumption is common to most results on the STVC, which basically means that no gain should be achieved by shifting the optimal path proportionally downward.

Assumption 3.6 is satisfied if X_t is convex, $(0, 0) \in X_t$, and $(x_t^*, 0) \in X_t$ for all $t \in \mathbb{Z}_+$. We use Assumption 3.6 in order to accommodate cases in which it is not always feasible to reduce capital stocks to zero in one step, as in models with irreversible investment (e.g., Olson, 1989). In a deterministic model, Ekeland and Scheinkman (1986, Proposition 5.1(iii), (iv)) used a special case of Assumption 3.6 in which $\underline{\lambda}_t = 0$ for all $t \in \mathbb{Z}_+$.

Assumption 3.6 with $\underline{\lambda}_t = 0$ implies that the convex combinations of the optimal path and the constant path at zero are feasible (except for the initial condition). It is natural to consider such comparison paths to derive the STVC, which is essentially the optimality requirement that no gain be achieved by shifting the optimal path proportionally downward, as mentioned above. Such comparison paths however need not be feasible in nonconvex models (e.g., Skiba, 1978), but in some cases other comparison paths can be constructed to obtain a different TVC. For example, in Skiba's model, under some conditions, the convex combinations of the optimal path and a certain constant path are feasible. Ekeland and Scheinkman (1986, Proposition 5.3) used this idea to show a result that can accommodate Skiba's model. This paper, on the other hand, is mainly concerned with convex models.

The next two assumptions impose some structure on g_t. In what follows, any property on a random variable such as $g_t(\lambda x_t^*, \lambda x_{t+1}^*)$ is understood as holding with probability one.

Assumption 3.7. $\forall t \in \mathbb{Z}_+$, (i) $g_t(x_t^*, x_{t+1}^*) > 0$ and (ii) $\forall t \in \mathbb{N}, g_t(\lambda x_t^*, \lambda x_{t+1}^*)$ is concave in $\lambda \in [\underline{\lambda}_0, 1]$, where $\underline{\lambda}_0$ is given by Assumption 3.6.

Part (i) above holds if $u(\cdot)$ satisfies the Inada condition at zero, as long as there is a feasible path satisfying (i). This part is not needed in the case of bounded utility. Part (ii) is used to find a lower bound on $g_t(\lambda x_t^*, \lambda x_{t+1}^*)$ that depends only on λ and $g_t(x_t^*, x_{t+1}^*)$ (see Lemma A.2). This part can be weakened considerably though it seems already general enough for most purposes.

In our notation, Ekeland and Scheinkman (1986, Proposition 5.1(iii)) assumed that

$$\forall t \in \mathbb{N}_+, \forall \lambda \in [0, 1], \quad \lambda g_t(x_t^*, x_{t+1}^*) \leq g_t(\lambda x_t^*, \lambda x_{t+1}^*). \tag{25}$$

Our results hold even if Assumption 3.7(ii) is replaced by (25). In fact, Lemma A.2 in Appendix A shows that Assumption 3.7 implies a weaker version of (25), and that property is all we need from Assumption 3.7(ii) to show our main results.

The following is our last assumption.

Assumption 3.8. $\forall t \in \mathbb{Z}_+, g_t(x_t^*, \lambda x_{t+1}^*)$ is nonincreasing and continuous in $\lambda \in (\underline{\lambda}_t, 1]$, where $\underline{\lambda}_t$ is given by Assumption 3.6.

That $g_t(x_t^*, \lambda x_{t+1}^*)$ is nonincreasing in λ basically means that there is a tradeoff between consumption and investment. Such a tradeoff is typically required for the existence of an optimal path. In this paper, this restriction is used to express the STVC as an equality condition as in (24). The continuity of $g_t(x_t^*, \lambda x_{t+1}^*)$ in λ plays a technical role that allows us to write $g_t(x_t^*, x_{t+1}^*)$ as the argument of $u'(\cdot)$ in (24). Without this continuity requirement, $g_t(x_t^*, x_{t+1}^*)$ must be replaced by $\underline{\lim}_{\lambda \uparrow 1} g_t(x_t^*, \lambda x_{t+1}^*)$ in (24); see the proof of Lemma A.3 for details.

4 Main Results

We are now ready to state our main results, which assume all the assumptions stated in the preceding section. Let us start from the case of bounded utility.

Proposition 4.1. *Let Assumptions 3.1–3.8 hold. Suppose u is bounded. Then TVC (24) holds.*

Hence the STVC is necessary as long as u is bounded. As discussed in Section 2.1, Proposition 4.1 is a well-known result in the deterministic case. If

additional assumptions are made, Proposition 4.1 follows from Zilcha (1976, Theorem 1) or Takekuma (1992, Theorem 5.2). The main advantage of our result is that it does not require any technical condition that is difficult to verify in applications. For example, while the results of Zilcha (1976) and Takekuma (1992) require the state variable x_t to be essentially bounded, our result imposes no such restriction.

Proposition 4.1 makes it clear that the STVC may fail to be necessary only if u is unbounded. Thus we next turn to the case of unbounded utility. But since unbounded utility functions that do not belong to the CRRA class are rather rare in applications, we restrict attention to this class. The following result deals with the case of logarithmic utility.

Proposition 4.2. *Let Assumptions 3.1–3.8 hold. Suppose*

$$u(\cdot) = \ln(\cdot). \tag{26}$$

Then TVC (24) holds.

Hence the STVC is necessary in the case of logarithmic utility. Proposition 4.2 generalizes a similar result shown in Kamihigashi (2001) for a deterministic model. It makes it clear that once logarithmic utility is assumed, there is no issue about necessity of the STVC even for stochastic models.

The following result deals with the case of non-logarithmic CRRA utility.

Proposition 4.3. *Let Assumptions 3.1–3.8 hold. Suppose*

$$\exists \alpha \in (-\infty, 1] \setminus \{0\}, \quad u(\cdot) = \frac{(\cdot)^\alpha}{\alpha}, \tag{27}$$

$$-\infty < \sum_{t=0}^\infty \beta^t E u(g_t(x_t^*, x_{t+1}^*)) < \infty. \tag{28}$$

Then TVC (24) holds.

Thus under (27), the STVC is guaranteed to be necessary unless lifetime utility is allowed to be infinite at the optimum. Such cases are usually ruled out in practice, and (28) is often assumed implicitly or taken for granted in applied studies. Proposition 4.3 generalizes similar results shown in Kamihigashi (2000, 2001) for deterministic models.

It is worth emphasizing that (28) requires lifetime utility to be finite only for the optimal path. It does not require lifetime utility to be finite for any

other feasible path. This is important since lifetime utility is often $-\infty$ for many feasible paths when $\alpha < 0$. Proposition 4.3 makes it clear that once non-logarithmic CRRA utility is assumed, there is no issue about necessity of the STVC for all practical purposes.

5   Endogenous Labor Supply

One can easily apply the results shown in the preceding section to models with endogenous labor supply such as RBC models (e.g., King and Rebelo, 1999). To do so, one can take an optimal labor (or, equivalently, leisure) path as given and consider the maximization problem over consumption and capital paths.

To be more specific, consider the following problem:

$$\left.\begin{array}{rl} \text{``} \max\limits_{\{x_t\}_{t=0}^{\infty}} & \sum\limits_{t=0}^{\infty} \beta^t E r(c_t, l_t) \text{''} \\[2mm] \text{s.t.} & x_0 = \overline{x}_0, \\ & \forall t \in \mathbb{Z}_+, c_t + x_{t+1} = f_t(x_t, 1 - l_t), \\ & c_t, x_{t+1} \geq 0, \quad 0 \leq l_t \leq 1, \end{array}\right\} \tag{29}$$

where c_t is consumption, l_t is leisure, r is the utility function, and f_t is the production function (which depends on ω in the same way as g_t does).[10] In this section we assume $n = 1$, i.e., $x_t \in \mathbb{R}$ for all $t \in \mathbb{Z}_+$.

First consider the case in which utility is additively separable in consumption and leisure, i.e.,

$$r(c_t, l_t) = u(c_t) + n(l_t) \tag{30}$$

for some functions $u(\cdot)$ and $n(\cdot)$. In this case one can take an optimal leisure path $\{l_t^*\}$ as given and simply set $g_t(x_t, x_{t+1}) = f_t(x_t, 1 - l_t^*) - x_{t+1}$. The model then reduces to (19), ignoring utility from leisure (which is taken as given).

When utility is not additively separable, it is typically the case that utility can be written as

$$r(c_t, l_t) = u(c_t e(l_t)) \tag{31}$$

---

[10] More precisely, $f_t(x_t, 1 - l_t)$ denotes the random variable $f_t(x_t(\omega), 1 - l_t(\omega), \omega)$ in the same way as $g_t(x_t, x_{t+1})$ denotes the random variable $g_t(x_t(\omega), x_{t+1}(\omega), \omega)$; recall the paragraph below Assumption 3.3.

for some functions $u(\cdot)$ and $e(\cdot)$ (e.g., King and Rebelo, 1999, p. 945). Typically, $u(\cdot)$ is assumed to satisfy (27) (or (26), which leads to (30)), and $e(\cdot)$ is required to satisfy certain regularity conditions (King and Rebelo, 1999, footnote 69). In the case of (31), one can take an optimal leisure path $\{l_t^*\}$ as given and set $g_t(x_t, x_{t+1}) = [f_t(x_t, 1 - l_t^*) - x_{t+1}]e(l_t^*)$. The model then reduces to (19) again.

We have shown that maximization problems of the form (29) reduce to (19) in typical parametric cases if an optimal leisure path is taken as given. There is no problem in taking an optimal leisure path as given, since necessary optimality conditions such as Euler equations and TVCs are expressed taking an optimal path as (implicitly) given anyway. The discussions here along with our main results indicate that there is practically no issue about necessity of the STVC for standard parametric stochastic growth models, whether labor supply is exogenous or endogenous.

6   Other Models with Control Variables

The preceding discussions suggest that our results apply even to models that do not take the exact form of (19), which contains no explicit "control" variable. Since many models involve both state and control variables, we briefly discuss here how a model with control variables can be reduced to (19).[11] To be specific, consider the following problem:

$$\left.\begin{array}{c} \text{"} \max\limits_{\{x_t\}_{t=0}^{\infty}} \quad \sum\limits_{t=0}^{\infty} \beta^t E u(h_t(x_t, a_t)) \text{"} \\[2mm] \text{s.t.} \quad x_0 = \overline{x}_0, \\[2mm] \forall t \in \mathbb{Z}_+, x_{t+1} = f_t(x_t, a_t), \quad x_{t+1}, a_t \geq 0, \end{array}\right\} \tag{32}$$

where $a_t \equiv (a_t^1, \ldots, a_t^N) \in \mathbb{R}^N$ for some $N \in \mathbb{N}$, and h_t and f_t depend on ω in the same way as g_t does (recall footnote 10).

If $N = n$ (where n is the dimension of x_t), it is often possible to reduce (32) to (19) by solving the constraint equation for a_t and substituting into the objective function. If there are multiple solutions for a_t (or if $N \neq n$), one way to reduce (32) to (19) is to define

$$g_t(x_t, x_{t+1}) = \max\{h_t(x_t, a_t) \mid x_{t+1} = f_t(x_t, a_t), a_t \geq 0\}, \tag{33}$$

---

[11]This section owes much to the comments made by an anonymous referee.

where the maximum is taken over a_t with x_t and x_{t+1} taken as given.[12] This is a special case of the standard reduced-form approach, which is to express the return function as a function of state variables by maximizing it in terms of control variables.

An alternative approach is to eliminate some control variables and to assume that the rest take optimal values. This is the approach used in Section 5. If applicable, it seems easier in most cases than the standard reduced-form approach. For example, the standard approach applied to (29) means expressing c_t and l_t as functions of x_t and x_{t+1}. This procedure does not reduce (29) to (19) under (30), in which case (29) does not take the exact form of (32) due to the additive term $n(l_t)$. But if we take an optimal path for $\{l_t\}$ as given, then (29) effectively becomes a special case of (32). Although the standard approach does reduce (29) to (19) under (31), the resulting STVC is rather complicated (though it may be simplified using the envelope theorem). Thus even when the standard approach is applicable, the alternative approach seems easier.

7 Concluding Remarks

In this paper we showed the necessity of the standard transversality condition (STVC) for a multi-sector growth model with a single consumption good under fairly general conditions. We considered the case of constant-relative-risk-aversion (CRRA) utility as well as that of bounded utility. In particular, it was shown that the STVC is necessary if utility is either bounded or logarithmic. It was also shown that the STVC is necessary in the case of non-logarithmic CRRA utility as long as lifetime utility is finite at the optimum. These results apply to various stochastic growth models, including RBC models with endogenous labor supply. Because of their general applicability, our results provide a fairly complete set of solutions regarding necessity of the STVC for stochastic growth models in applications. Indeed, since utility functions that do not belong to the CRRA class are rather rare in applications, our results seem to cover most cases in practice.

---

[12]The existence of the maximum in (33) requires some assumptions, such as upper semicontinuity of $h_t(x_t, a_t)$ in a_t and nonemptiness and compactness of the set $\{a_t \geq 0 \mid x_{t+1} = f_t(x_t, a_t)\}$.

Appendix A Preliminary Lemmas

The proofs of our propositions use a result shown in Kamihigashi (2003) for a general reduced-form model. We state it without proof after introducing some notation. After stating the result, we show two further lemmas. All the assumptions stated in Section 3 are maintained here though not all of them are used.

For $t \in \mathbb{Z}_+, (y, z) \in X_t$, and $\omega \in \Omega$, define

$$v_t(y(\omega), z(\omega), \omega) = \beta^t u(g_t(y(\omega), z(\omega), \omega)). \tag{A.1}$$

Like $g_t(x_t, x_{t+1})$, $v_t(x_t, x_{t+1})$ denotes the random variable $v_t(x_t(\cdot), x_{t+1}(\cdot), \cdot) : \Omega \to \mathbb{R} \cup \{-\infty\}$ for $(x_t, x_{t+1}) \in X_t$. For $t \in \mathbb{N}$ and $\lambda < 1$ with $(\lambda x_t^*, \lambda x_{t+1}^*) \in X_t$, define

$$w_t(\lambda) = \frac{Ev_t(x_t^*, x_{t+1}^*) - Ev_t(\lambda x_t^*, \lambda x_{t+1}^*)}{1 - \lambda}, \tag{A.2}$$

$$\hat{w}_t(\lambda) = \sup_{z \in [\lambda, 1)} w_t(z), \tag{A.3}$$

where $\hat{w}_t(\lambda)$ is defined for $\lambda < 1$ such that $\forall z \in [\lambda, 1), (z x_t^*, z x_{t+1}^*) \in X_t$.

Lemma A.1. *Suppose*

$$\exists \{b_t\}_{t=1}^{\infty} \subset \mathbb{R}, \exists \lambda \in [\underline{\lambda}_0, 1), \forall t \in \mathbb{N}, \quad \hat{w}_t(\lambda) \le b_t, \tag{A.4}$$

$$\sum_{t=1}^{\infty} b_t \ \text{exists in} \ [-\infty, \infty), \tag{A.5}$$

where $\underline{\lambda}_0$ is given by Assumption 3.6. Then

$$\overline{\lim_{t \uparrow \infty}} \, Ev_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) \le 0, \tag{A.6}$$

where $v_{t,2}$ is defined as in (22).

Proof. See Kamihigashi (2003, Theorem 2.2). $\qquad\qquad\square$

Conditions (A.4) and (A.5) basically mean that the optimal path can be shifted proportionally downward without causing an infinite loss. Using these conditions, one can show (A.6) by a generalized version of the squeezing argument (recall Section 2.1).

For simplicity, we use the following notations for $t \in \mathbb{Z}_+$:

$$g_t^* = g_t(x_t^*, x_{t+1}^*), \quad g_t(\lambda) = g_t(\lambda x_t^*, \lambda x_{t+1}^*). \tag{A.7}$$

16

Lemma A.2. *We have*

$$\forall t \in \mathbb{N}, \forall \lambda \in [\underline{\lambda}_0, 1), \quad g_t(\lambda) \geq m(\lambda) g_t^*, \tag{A.8}$$

where $\underline{\lambda}_0$ is given by Assumption 3.6 and

$$m(\lambda) = \frac{\lambda - \underline{\lambda}_0}{1 - \underline{\lambda}_0}. \tag{A.9}$$

Proof. It is easy to see that

$$m(\lambda) + [1 - m(\lambda)]\underline{\lambda}_0 = \lambda. \tag{A.10}$$

Let $t \in \mathbb{N}$. Since $g_t(\lambda)$ is concave in $\lambda \in [\underline{\lambda}_0, 1]$ by Assumption 3.7, for $\lambda \in [\underline{\lambda}_0, 1]$,

$$g_t(\lambda) \geq m(\lambda) g_t^* + [1 - m(\lambda)] g_t(\underline{\lambda}_0) \tag{A.11}$$
$$\geq m(\lambda) g_t^*, \tag{A.12}$$

where the last inequality holds since $g_t(\underline{\lambda}_0) \geq 0$ by Assumptions 3.6 and 3.3. Now (A.8) follows. $\qquad\square$

Lemma A.3. *TVC (A.6) is equivalent to TVC (24).*

Proof. Since Assumption 3.8 implies $v_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) \geq 0$, TVC (A.6) is equivalent to

$$\lim_{t \uparrow \infty} E v_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) = 0. \tag{A.13}$$

Thus it suffices to verify the following for $t \in \mathbb{Z}_+$:

$$v_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) = \beta^t u'(g_t^*) g_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*). \tag{A.14}$$

Let $t \in \mathbb{Z}_+$. Let $\epsilon > 0$ be such that $(1 - \epsilon) \in [\underline{\lambda}_0, 1)$. Note that

$$v_t(x_t^*, x_{t+1}^* - \epsilon x_{t+1}^*) - v_t(x_t^*, x_{t+1}^*) \tag{A.15}$$
$$= \beta^t [u(g_t(x_t^*, x_{t+1}^* - \epsilon x_{t+1}^*)) - u(g_t^*)] \tag{A.16}$$
$$= \beta^t u'(\tilde{g})[g_t(x_t^*, x_{t+1}^* - \epsilon x_{t+1}^*) - g_t^*], \tag{A.17}$$

where (A.17) holds by the mean value theorem for some

$$\tilde{g} \in [g_t^*, g_t(x_t^*, (1 - \epsilon)x_{t+1}^*)]. \tag{A.18}$$

Since $g_t(x_t^*, (1 - \epsilon)x_{t+1}^*)$ is continuous in $\epsilon \in [0, 1 - \underline{\lambda}_t)$ by Assumption 3.8, $\tilde{g}$ converges to g_t^* as $\epsilon \downarrow 0$. Thus dividing (A.15) and (A.17) by ϵ and applying $\underline{\lim}_{\epsilon \downarrow 0}$ yields (A.14). $\qquad\square$

Appendix B   Proofs

By Lemma A.3, to conclude TVC (24) from Lemma A.1, it suffices to verify (A.4) and (A.5). These conditions are verified in each of the following three proofs.

B.1   Proof of Proposition 4.1

Let $\underline{\lambda} \in (\underline{\lambda}_0, 1), \lambda \in [\underline{\lambda}, 1)$, and $t \in \mathbb{N}$. By Assumption 3.7(ii) and the concavity of u, $w_t(\lambda)$ is nonincreasing in $\lambda \in [\underline{\lambda}, 1)$. Thus

$$\hat{w}_t(\lambda) \leq w_t(\underline{\lambda}) = \beta^t \frac{u(g_t(x_t^*, x_{t+1}^*)) - u(g_t(\underline{\lambda} x_t^*, \underline{\lambda} x_{t+1}^*))}{1 - \underline{\lambda}}. \tag{B.1}$$

Let b_t equal the rightmost side above. Then $\{b_t\}$ clearly satisfies (A.4). It also satisfies (A.5) by the boundedness of u.

B.2   Proof of Proposition 4.2

Let $\underline{\lambda} \in (\underline{\lambda}_0, 1), \lambda \in [\underline{\lambda}, 1)$, and $t \in \mathbb{N}$. We have

$$(1 - \lambda)\beta^{-t} w_t(\lambda) = E[\ln g_t^* - \ln g_t(\lambda)] \tag{B.2}$$

$$\leq E\left[\ln g_t^* - \ln\left(m(\lambda)g_t^*\right)\right] \tag{B.3}$$

$$= -\ln m(\lambda), \tag{B.4}$$

where (B.3) uses Lemma A.2. It follows that

$$\beta^{-t} w_t(\lambda) \leq \frac{-\ln m(\lambda)}{1 - \lambda} \leq A, \tag{B.5}$$

where

$$A = \sup_{\lambda \in [\underline{\lambda}, 1)} \frac{-\ln m(\lambda)}{1 - \lambda}. \tag{B.6}$$

Note that A is finite since the supremand is continuous on $[\underline{\lambda}, 1)$ and converges to $1/(1 - \underline{\lambda}_0)$ as $\lambda \uparrow 1$. Now (A.4) and (A.5) hold with $b_t = \beta^t A$.

B.3   Proof of Proposition 4.3

Let $\underline{\lambda} \in (\underline{\lambda}_0, 1), \lambda \in [\underline{\lambda}, 1)$, and $t \in \mathbb{N}$. We have

$$(1 - \lambda)\beta^{-t}w_t(\lambda) = E[u(g_t^*) - u(g_t(\lambda))] \tag{B.7}$$

$$\leq E[u(g_t^*) - u(m(\lambda)g_t^*)] \tag{B.8}$$

$$= [1 - (m(\lambda))^\alpha]Eu(g_t^*), \tag{B.9}$$

where (B.8) holds by Lemma A.2, and (B.9) holds by (27). It follows that

$$\beta^{-t}w_t(\lambda) \leq \frac{1 - (m(\lambda))^\alpha}{(1 - \lambda)\alpha}\alpha Eu(g_t^*) \tag{B.10}$$

$$\leq A\alpha Eu(g_t^*), \tag{B.11}$$

where

$$A = \sup_{\lambda \in [\underline{\lambda}, 1)} \frac{1 - (m(\lambda))^\alpha}{(1 - \lambda)\alpha}. \tag{B.12}$$

Note that A is finite since the supremand is continuous in $\lambda \in [\underline{\lambda}, 1)$ and converges to $1/(1 - \underline{\lambda}_0)$ as $\lambda \uparrow 1$. Now (B.10), (B.11), and (28) imply (A.4) and (A.5) with $b_t = \beta^t A\alpha Eu(g_t^*)$.

References

Alvarez, F., Stokey, N.L., 1998, Dynamic programming with homogeneous functions, *Journal of Economic Theory* 82, 167–189.

Barro, R.J., Sala-i-Martin, X., 1995, Economic Growth. McGraw-Hill, New York.

Benhabib, J., and R.E.A. Farmer, 1999, Indeterminacy and sunspots in macroeconomics, in: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics, Vol. 1A. Elsevier, Amsterdam, pp. 387–448.

Benveniste, L.M., Scheinkman, J.A., 1982, Duality theory for dynamic optimization models of economics: the continuous time case, Journal of Economic Theory 27, 1–19.

Brock, W.A., 1970, On existence of weakly maximal programmes in a multi-sector economy," Review of Economic Studies 37, 275–280.

Brock, W.A., Mirman, L.J., 1972, Optimal economic growth and uncertainty: the discounted case, Journal of Economic Theory 4, 479–513.

Ekeland, I., Scheinkman, J.A., 1986, Transversality conditions for some infinite horizon discrete time optimization problems, Mathematics of Operations Research 11, 216–229.

Foldes, L., 1978, Martingale conditions for optimal saving - discrete time, Journal of Mathematical Economic 5, 83–96.

Halkin, H., 1974, Necessary conditions for optimal control problems with infinite horizons, Econometrica 42, 267–272.

Kamihigashi, T., 2000, A simple proof of Ekeland and Scheinkman's result on the necessity of a transversality condition, Economic Theory 15, 463–468.

Kamihigashi, T., 2001, Necessity of transversality conditions for infinite horizon problems, Econometrica 69, 995–1012.

Kamihigashi, T., 2002, A simple proof of the necessity of the transversality condition, Economic Theory 20, 427–433.

Kamihigashi, T., 2003, Necessity of transversality conditions for stochastic problems, Journal of Economic Theory 109, 140–149.

King, R.G., Rebelo, S.T., 1999, Resuscitating real business cycles," in: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics, Vol. 1A. Elsevier, Amsterdam, pp. 927–1007.

Martínez-García, M.P., 2003, The general instability of balanced paths in endogenous growth models: the role of transversality conditions, Journal of Economic Dynamics and Control 27, 599–618.

Michel, P., 1990, Some clarifications on the transversality condition, Econometrica 58, 705–723.

Olson, L.J., 1989, Stochastic growth with irreversible investment, Journal of Economic Theory 47, 101–129.

Shell, K., 1969, Applications of Pontryagin's maximum principle to economics, in: Kuhn, H.W., Szegö, G.P. (Eds.) Mathematical Systems Theory and Economics I. Springer-Verlag, New York, pp. 241–292.

Skiba, A.K., 1978, Optimal growth with a convex-concave production function, Econometrica 46, 527–539.

Stokey, N., Lucas, R.E., Jr., 1989, Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, MA.

Takekuma, S., 1992, Optimal growth under uncertainty: a complete characterization of weakly maximal programs, Hitotsubashi Journal of Economics 33, 169–182.

Weitzman, M.L., 1973, Duality theory for infinite horizon convex models, Management Science 19, 783–789.

Zilcha, I., 1976, Characterization by prices of optimal programs under uncertainty, Journal of Mathematical Economics 3, 173–181.