# Necessity of Transversality Conditions for Stochastic Problems

(Forthcoming in the *Journal of Economic Theory*)

Takashi Kamihigashi*

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

Phone/Fax: +81-78-803-7015
E-mail: tkamihig@rieb.kobe-u.ac.jp

April 26, 2002

**Abstract**

This paper shows stochastic versions of (i) Michel's (1990, *Econometrica* **58**, 705–723, Theorem 1) necessity result , (ii) a generalization of the TVC results of Weitzman (1973, *Manage. Sci.* **19**, 783–789) and Ekeland and Scheinkman (1986, *Math. Oper. Res.* **11**, 216–229), and (iii) Kamihigashi's (2001, *Econometrica* **69**, 995–1012, Theorem 3.4) result, which is useful particularly in the case of homogeneous returns. These stochastic extensions are established for an extremely general stochastic reduced-form model that assumes neither differentiability nor continuity. *Journal of Economic Literature Classification Numbers:* C61, D90, G12.

**Keywords:** Transversality condition, stochastic optimization, stochastic reduced-form model, homogeneous returns, stochastic growth model.

---

*I would like to thank an associate editor of this *Journal* and an anonymous referee for helpful comments and suggestions. Remaining errors are mine.

# 1 Introduction

Although general results on necessity of transversality conditions (TVCs) for deterministic problems are shown in Kamihigashi [4], widely applicable results for stochastic problems are not available in the literature. As mentioned below, the stochastic versions of Weitzman's [9] theorem shown by Zilcha [10] and Takekuma [8] are not easily applicable.[1] Furthermore the existing literature does not offer stochastic versions of the results of Michel [7] and Ekeland and Scheinkman [2].[2]

We establish three results in this paper. Our first result is a stochastic version of the necessity part of Michel [7, Theorem 1]. Our second result is a stochastic version of Kamihigashi's [4, Theorem 3.3] result, which is a generalization of the TVC results of Weitzman [9] and Ekeland and Scheinkman [2]. Our second result also generalizes the TVC results of Zilcha [10] and Takekuma [8]. Our third result is a stochastic version of Kamihigashi's [4, Theorem 3.4] result, which is useful particularly in the case of homogeneous returns.

We follow Ekeland and Scheinkman [2] in using directional derivatives instead of support prices. The results of Zilcha [10] and Takekuma [8] are not easily applicable since they use support prices and thus rely heavily on the infinite-dimensional separation theorem, which requires severe restrictions.[3] We do not require such restrictions since, instead of constructing support prices, we simply use lower Dini directional derivatives, which are well-defined even if the return functions are discontinuous or nondifferentiable. This allows us to concentrate on conditions directly related to TVCs.

The proofs of our main results are based on our realization that necessary TVCs for a stochastic problem can be derived from a one-dimensional deterministic problem. This approach enables us to obtain TVCs for an extremely general stochastic reduced-form model with very few technical restrictions. More specifically, our main results are proved in Appendix B by simple applications of the general results established in Appendix A for a one-dimensional deterministic problem. Section 2 presents the main results. Section 3 concludes the paper.

[1]See Zilcha [11] for results specific to an undiscounted stationary model.

[2]For discussions on these and other related results for deterministic problems, see Kamihigashi [3, 4, 6].

[3]It should be mentioned that the difficult part in establishing their results is the construction of support prices, not the proof of the necessity of TVCs.

## 2 Main Results

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 $\omega$, we write $Ez(\omega)$ instead of $Ez$. Consider the following problem.

$$\begin{cases} \text{“}\max\limits_{\{x_t\}_{t=0}^{\infty}} & \sum\limits_{t=0}^{\infty} Ev_t(x_t(\omega), x_{t+1}(\omega), \omega)\text{”} \\ \text{s.t.} & x_0 = \overline{x}_0, \quad \forall t \in \mathbb{Z}_+, (x_t, x_{t+1}) \in X_t.^4 \end{cases} \tag{2.1}$$

For any pair of sets $Y$ and $Z$, let $F(Y, Z)$ denote the set of all functions from $Y$ to $Z$. We assume the following.

**Assumption 2.1.** There exists a sequence of real vector spaces $\{B_t\}_{t=0}^{\infty}$ such that $\overline{x}_0 \in F(\Omega, B_0)$ and $\forall t \in \mathbb{Z}_+, X_t \subset F(\Omega, B_t) \times F(\Omega, B_{t+1})$.

**Assumption 2.2.** $\forall t \in \mathbb{Z}_+, \forall (y, z) \in X_t$, (i) $\forall \omega \in \Omega, v_t(y(\omega), z(\omega), \omega) \in [-\infty, \infty)$, (ii) the mapping $v_t(y(\cdot), z(\cdot), \cdot) : \Omega \to [-\infty, \infty)$ is measurable, and (iii) $Ev_t(y(\omega), z(\omega), \omega)$ exists in $[-\infty, \infty)$.

Assumption 2.1 means that $x_t$ is a random variable whose realization lies in a real vector space. Assumption 2.2 simply means that the expression $Ev_t(x_t(\omega), x_{t+1}(\omega), \omega)$ makes sense.

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 [1]) as our optimality criterion. We say that a feasible path $\{x_t^*\}$ is *optimal* if for any feasible path $\{x_t\}$,

$$\varliminf_{T \uparrow \infty} \sum_{t=0}^{T} [Ev_t(x_t(\omega), x_{t+1}(\omega), \omega) - Ev_t(x_t^*(\omega), x_{t+1}^*(\omega), \omega)] \leq 0.^5 \tag{2.2}$$

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

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

[5] To be precise, this inequality requires that the left side is well-defined. This means that the left side does not involve expressions like "$\infty - \infty$" and "$-\infty + \infty$." An implication of this requirement is that $\forall t \in \mathbb{Z}_+, Ev_t(x_t^*, x_{t+1}^*)$ is finite; for otherwise the left side of (2.2) is undefined for $\{x_t\} = \{x_t^*\}$.

criterion fails, and (iii) is weaker than the similar criterion with $\overline{\lim}$ replacing $\underline{\lim}$ in (2.2). Our results therefore apply to virtually any problem of the form (2.1).

In stochastic optimization problems, feasible paths are usually required to be adapted to a filtration (e.g., $x_{t+1}$ can depend only on information available in period $t$). Though such an information structure could be imposed here, it is already covered by Assumptions 2.1 and 2.2. To be more specific, let $\{\mathcal{F}_t\}$ be a filtration; e.g., $\mathcal{F}_t$ may be the $\sigma$-field generated by all possible histories of shocks up to period $t$. Note that since $\mathcal{F}_t$-measurability implies measurability (i.e., $\mathcal{F}$-measurability), Assumption 2.2 allows $v_t(y(\cdot), z(\cdot), \cdot)$ to be $\mathcal{F}_t$-measurable. Let $M_t$ be the set of $\mathcal{F}_{t-1}$-measurable functions from $\Omega$ to $B_t$ (assuming $B_t$ is a measurable space). Then one can require feasible paths to be adapted to $\{\mathcal{F}_t\}$ by assuming $X_t \subset M_t \times M_{t+1}$, i.e., $x_{t+1}(\cdot)$ is $\mathcal{F}_t$-measurable. But this is only a special case of Assumption 2.1 since obviously $M_t \subset F(\Omega, B_t)$. Likewise other information structures like this are covered by Assumptions 2.1 and 2.2. The rest of the section assumes the following.

**Assumption 2.3.** There exists an optimal path $\{x_t^*\}$.

Since we are only interested in necessary conditions for optimality, this assumption imposes no restriction on the model. For simplicity, for $(x_t, x_{t+1}) \in X_t$, let $v_t(x_t, x_{t+1})$ denote the random variable $v_t(x_t(\cdot), x_{t+1}(\cdot), \cdot) : \Omega \to [-\infty, \infty)$. For $t \in \mathbb{Z}_+$ and $d \in F(\Omega, B_{t+1})$ such that $(x_t^*, x_{t+1}^* + \epsilon d) \in X_t$ for sufficiently small $\epsilon > 0$, we define the random variable $v_{t,2}(x_t^*, x_{t+1}^*; d)$ by

$$v_{t,2}(x_t^*, x_{t+1}^*; d) = \underline{\lim}_{\epsilon \downarrow 0} \frac{v_t(x_t^*, x_{t+1}^* + \epsilon d) - v_t(x_t^*, x_{t+1}^*)}{\epsilon}, \qquad (2.3)$$

where $\underline{\lim}_{\epsilon \downarrow 0}$ is applied pointwise (i.e., for each $\omega \in \Omega$ separately). Even if $v_t$ is nondifferentiable or discontinuous, the right side of (2.3), which is a lower Dini directional derivative, is well-defined (with probability one) as long as $(x_t^*, x_{t+1}^* + \epsilon d) \in X_t$ for sufficiently small $\epsilon > 0$. Note that if $v_t$ is differentiable in the second argument (in an appropriate sense), then $v_{t,2}(x_t^*, x_{t+1}^*; d) = [\partial v_t(x_t^*, x_{t+1}^*)/\partial x_{t+1}^*]d$. Note also that $v_{t,2}(x_t^*, x_{t+1}^*; d)$ reduces to the usual directional derivative if $v_{t,2}$ is concave in the second argument.

**Remark 2.1.** Theorems 2.1–2.3 below hold even if $\overline{\lim}$ replaces $\underline{\lim}$ in (2.3).[6]

[6] The working paper version of this paper [5] explains why this remark is true. See Assumption 2.5 and Remark 2.4 for why we use $\underline{\lim}$ in (2.3).

Our basic strategy in deriving necessary TVCs for our stochastic problem is to consider the following *deterministic* problem of choosing $\{y_t\}_{t=0}^{\infty} \subset \mathbb{R}$:

$$\begin{cases} \text{“}\max\limits_{\{y_t\}_{t=0}^{\infty}} & \sum\limits_{t=0}^{\infty} Ev_t(x_t^* + y_t e_t, x_{t+1}^* + y_{t+1}e_{t+1})\text{”} \\ \text{s.t.} & y_0 = 0, \quad \forall t \in \mathbb{Z}_+, (x_t^* + y_t e_t, x_{t+1}^* + y_{t+1}e_{t+1}) \in X_t, \end{cases} \tag{2.4}$$

where $\forall t \in \mathbb{Z}_+, e_t \in F(\Omega, B_t)$. For $t \in \mathbb{Z}_+$, let $y_t^* = 0$. Then $\{y_t^*\}$ is optimal for (2.4) since $\{x_t^*\}$ is optimal for (2.1). The proofs of our main results are based on our realization that necessary TVCs for (2.1) can be derived from (2.4) with $\{e_t\}$ appropriately chosen. Indeed our main results are established by applying the results shown in Appendix A for a one-dimensional deterministic problem of which (2.4) is a special case.

**Theorem 2.1.** *Assume Assumptions 2.1–2.3. Suppose* $\forall t \in \mathbb{Z}_+, X_t$ *is convex and* $\forall \omega \in \Omega, v_t(\cdot, \cdot, \omega)$ *is concave. Then*

$$\varliminf_{t\uparrow\infty} Ev_{t,2}(x_t^*, x_{t+1}^*; x_{t+1} - x_{t+1}^*) \leq 0^7 \tag{2.5}$$

*for any feasible path* $\{x_t\}$ *such that*

$$\varliminf_{T\uparrow\infty} \sum_{t=1}^{T} [Ev_t(x_t, x_{t+1}) - Ev_t(x_t^*, x_{t+1}^*)] > -\infty, \tag{2.6}$$

$$\begin{cases} \forall t \in \mathbb{Z}_+, \exists \epsilon > 0, \\ \zeta_t(\epsilon) \equiv (x_t^*, x_{t+1}^* + \epsilon(x_{t+1} - x_{t+1}^*)) \in X_t, Ev_t(\zeta_t(\epsilon)) > -\infty. \end{cases} \tag{2.7}$$

Theorem 2.1 is a stochastic version of the necessity part of Michel [7, Theorem 1]. Condition (2.7) is needed here for $Ev_{t,2}(x_t^*, x_{t+1}^*; x_{t+1} - x_{t+1}^*)$ to be well-defined. The rest of this section assumes the following.

**Assumption 2.4.** $\forall t \in \mathbb{Z}_+, \exists \underline{\lambda}_t \in [0, 1), \forall \lambda \in [\underline{\lambda}_t, 1), (x_t^*, \lambda x_{t+1}^*) \in X_t$ and $\forall \tau \geq t + 1, (\lambda x_\tau^*, \lambda x_{\tau+1}^*) \in X_\tau$.

**Remark 2.2.** Assumption 2.4 holds if $\forall t \in \mathbb{Z}_+, X_t$ is convex and $(x_t^*, 0), (0, 0) \in X_t$.

[7] By this inequality, we also mean that the mapping $v_{t,2}(x_t^*, x_{t+1}^*; x_{t+1} - x_{t+1}^*) : \Omega \to \overline{\mathbb{R}}$ is measurable and $Ev_{t,2}(x_t^*, x_{t+1}^*; x_{t+1} - x_{t+1}^*)$ exists in $\overline{\mathbb{R}}$. The same remark applies to (2.10).

Assumption 2.4 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 standard TVC, which basically means that no gain should be achieved by shifting the optimal path proportionally downward.

For $t \in \mathbb{N} \equiv \{1, 2, 3, \ldots\}$ and $\lambda \in \mathbb{R} \setminus \{1\}$ with $(\lambda x_t^*, \lambda x_{t+1}^*) \in X_t$, define

$$w_t(\lambda) = \frac{Ev_t(x_t^*, x_{t+1}^*) - Ev_t(\lambda x_t^*, \lambda x_{t+1}^*)}{1 - \lambda}, \tag{2.8}$$

$$\hat{w}_t(\lambda) = \sup_{\tilde{\lambda} \in [\lambda, 1)} w_t(\tilde{\lambda}), \tag{2.9}$$

where $\hat{w}_t(\lambda)$ is defined for $\lambda < 1$ such that $\forall \tilde{\lambda} \in [\lambda, 1), (\lambda x_t^*, \lambda x_{t+1}^*) \in X_t$. Note from Assumption 2.4 that $\hat{w}_t(\lambda)$ is defined at least for $\lambda \in [\underline{\lambda}_0, 1)$.

**Remark 2.3.** If $v_t(\lambda x_t^*, \lambda x_{t+1}^*)$ is concave in $\lambda \in [\underline{\lambda}_0, 1]$,[8] then $\forall \lambda \in [\underline{\lambda}_0, 1), \hat{w}_t(\lambda) = w_t(\lambda)$.

**Assumption 2.5.** $\forall t \in \mathbb{Z}_+$,

$$Ev_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) \leq \varliminf_{\epsilon \downarrow 0} E \frac{v_t(x_t^*, x_{t+1}^* - \epsilon x_{t+1}^*) - v_t(x_t^*, x_{t+1}^*)}{\epsilon}. \tag{2.10}$$

**Remark 2.4.** (2.10) holds by Fatou's lemma (recall (2.3)) if the mapping $v_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) : \Omega \to \overline{\mathbb{R}}$ is measurable and if $v_t(x_t^*, \lambda x_{t+1}^*)$ is nonincreasing in $\lambda \in [\underline{\lambda}_t, 1]$ (which is the case in most economic models).

**Remark 2.5.** (2.10) holds with equality by the monotone convergence theorem if $v_t(x_t^*, \lambda x_{t+1}^*)$ is concave in $\lambda \in [\underline{\lambda}_t, 1]$ and if $\exists \lambda \in [\underline{\lambda}_t, 1), Ev_t(x_t^*, \lambda x_{t+1}^*) > -\infty$.

**Theorem 2.2.** *Assume Assumptions 2.1–2.5. Suppose*

$$\exists \{b_t\}_{t=1}^{\infty} \subset \mathbb{R}, \exists \lambda \in [\underline{\lambda}_0, 1), \forall t \in \mathbb{N}, \quad \hat{w}_t(\lambda) \leq b_t. \tag{2.11}$$

*Then* (i) *(2.12) ⇒ (2.13) and* (ii) *(2.14) ⇒ (2.15), where (2.12)–(2.15) are*

[8] To be precise, by "$v_t(\lambda x_t^*, \lambda x_{t+1}^*)$ is concave in $\lambda$," we mean that with probability one, $v_t(\lambda x_t^*, \lambda x_{t+1}^*)$ is a concave function of $\lambda$. Likewise any condition involving random variables is understood to hold with probability one.

*given by*

$$\varlimsup_{T\uparrow\infty} \sum_{t=1}^{T} b_t < \infty, \tag{2.12}$$

$$\varliminf_{t\uparrow\infty} Ev_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) \leq 0, \tag{2.13}$$

$$\sum_{t=1}^{\infty} b_t \textit{ exists in } [-\infty, \infty), \tag{2.14}$$

$$\varlimsup_{t\uparrow\infty} Ev_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) \leq 0. \tag{2.15}$$

Conclusion (ii) of Theorem 2.2 generalizes the TVC results of Zilcha [10] and Takekuma [8] except that our result uses lower Dini directional derivatives instead of support prices. Theorem 2.2 is also a stochastic version of Kamihigashi's [4, Theorem 3.3] result, which is a generalization of the TVC results of Weitzman [9] and Ekeland and Scheinkman [2]. Our last result uses the following assumption.

**Assumption 2.6.** $\exists \overline{\mu} > 1, \forall \mu \in (1, \overline{\mu}]$, (i) $(x_0^*, \mu x_1^*) \in X_0$, (ii) $\forall t \in \mathbb{N}, (\mu x_t^*, \mu x_{t+1}^*) \in X_t$, (iii) $Ev_0(x_0^*, \mu x_1^*) > -\infty$, and (iv) $\forall t \in \mathbb{N}, Ev_t(\mu x_t^*, \mu x_{t+1}^*) \geq Ev_t(x_t^*, x_{t+1}^*)$.

Assumption 2.6 means that the optimal path can be shifted proportionally upward ((i) and (ii)) and that a sufficiently small such shift entails a finite loss in period 0 and nonnegative gains in the subsequent periods ((iii) and (iv)). The assumption is innocuous at least for standard models with homogeneous returns.

**Theorem 2.3.** *Assume Assumptions 2.1–2.6. Suppose*

$$\exists \lambda \in [\underline{\lambda}_0, 1), \exists \mu \in (1, \overline{\mu}], \exists \theta \geq 0, \forall t \in \mathbb{N}, \qquad \hat{w}_t(\lambda) \leq \theta w_t(\mu). \tag{2.16}$$

*Then TVC (2.15) holds.*

Theorem 2.3 is similar to Kamihigashi's [4, Theorem 3.4] result, but the proof here is quite different. Basically Kamihigashi's [4, Theorem 3.4] result uses $\lim_{\mu\downarrow 1} w_t(\mu)$ in (2.16) instead of $w_t(\mu)$ and its proof relies on differentiability and the Euler equation. The proof of Theorem 2.3 by contrast verifies (2.11) and (2.14) using (2.16) and Assumption 2.6. Theorem 2.3 is useful particularly in the case of homogenous returns.

# 3 Concluding Remarks

This paper has shown stochastic versions of (i) Michel's [7, Theorem 1] necessity result, (ii) a generalization of the TVC results of Weitzman [9] and Ekeland and Scheinkman [2], and (iii) Kamihigashi's [4, Theorem 3.4] result, which is useful particularly in the case of homogeneous returns. These stochastic extensions have been established for an extremely general stochastic reduced-form model that assumes neither differentiability nor continuity.

Our results are significant as well as useful, not only because the previous literature does not provide widely applicable results on necessity of TVCs for stochastic problems, but also because our results require very few technical restrictions in addition to those needed for the corresponding deterministic results. Our results suggest that as far as necessity of TVCs is concerned, there is little difference between deterministic and stochastic cases.

# A General Results

This appendix establishes two general results for a one-dimensional deterministic problem. By applying those results, Appendix B proves Theorems 2.1–2.3. Consider the following problem.

$$\begin{cases} \text{“}\displaystyle\max_{\{y_t\}_{t=0}^\infty} & \displaystyle\sum_{t=0}^\infty r_t(y_t, y_{t+1})\text{”} \\ \quad\text{s.t.} & y_0 = \overline{y}_0, \quad \forall t \in \mathbb{Z}_+, (y_t, y_{t+1}) \in Y_t. \end{cases} \tag{A.1}$$

**Assumption A.1.** $\overline{y}_0 \in \mathbb{R}$ and $\forall t \in \mathbb{Z}_+, Y_t \subset \mathbb{R} \times \mathbb{R}$.

**Assumption A.2.** $r_t : Y_t \to [-\infty, \infty)$.

Assumption A.1 says that $y_t$ is one-dimensional. Feasible paths and optimal paths are defined as in Section 2.

**Assumption A.3.** There exists an optimal path $\{y_t^*\}$.

**Assumption A.4.** $\forall t \in \mathbb{Z}_+, \exists \overline{\epsilon}_t > 0, \forall \epsilon \in (0, \overline{\epsilon}_t], (y_t^*, y_{t+1}^* - \epsilon) \in Y_t$ and $\forall \tau \geq t+1, (y_\tau^* - \epsilon, y_{\tau+1}^* - \epsilon) \in Y_\tau$.

Assumption A.4 means that the optimal path can be shifted uniformly downward starting from any period. For $t \in \mathbb{Z}_+$, define

$$q_t = r_{t,2}(y_t^*, y_{t+1}^*; -1), \tag{A.2}$$

where the right side is defined as in (2.3). For $t \in \mathbb{N}$ and $\epsilon \in \mathbb{R} \setminus \{0\}$ with $(y_t^* - \epsilon, y_{t+1}^* - \epsilon) \in Y_t$, define

$$m_t(\epsilon) = \frac{r_t(y_t^*, y_{t+1}^*) - r_t(y_t^* - \epsilon, y_{t+1}^* - \epsilon)}{\epsilon}, \tag{A.3}$$

$$\hat{m}_t(\epsilon) = \sup_{\tilde{\epsilon} \in (0,\epsilon]} m_t(\tilde{\epsilon}), \tag{A.4}$$

where $\hat{m}_t(\epsilon)$ is defined for $\epsilon > 0$ such that $\forall \tilde{\epsilon} \in (0, \epsilon], (y_t^* - \epsilon, y_{t+1}^* - \epsilon) \in Y_t$. Note from Assumption A.4 that $\hat{m}_t(\epsilon)$ is defined at least for $\epsilon \in (0, \overline{\epsilon}_0]$. Define

$$\Psi = \{\{f_t\}_{t=1}^{\infty} \subset \mathbb{R} \,|\, \overline{\lim}_{T \uparrow \infty} \textstyle\sum_{t=1}^{T} f_t \in [-\infty, \infty)\}, \tag{A.5}$$

$$\Phi = \{\{f_t\}_{t=1}^{\infty} \subset \mathbb{R} \,|\, \lim_{T \uparrow \infty} \textstyle\sum_{t=1}^{T} f_t \text{ exits in } [-\infty, \infty)\}. \tag{A.6}$$

The following result can be shown by the argument of Kamihigashi [4, Theorem 3.2]. The working paper version of this paper [5] contains the details.

**Theorem A.1.** *Assume Assumptions A.1–A.4. Suppose*

$$\exists \{b_t\}_{t=1}^{\infty} \subset \mathbb{R}, \exists \epsilon \in (0, \overline{\epsilon}_0], \forall t \in \mathbb{N}, \qquad \hat{m}_t(\epsilon) \leq b_t. \tag{A.7}$$

*Then* (i) $\{b_t\} \in \Psi \Rightarrow \underline{\lim}_{t \uparrow \infty} q_t \leq 0$ *and* (ii) $\{b_t\} \in \Phi \Rightarrow \overline{\lim}_{t \uparrow \infty} q_t \leq 0$.

The following assumption means that the optimal path can be shifted uniformly upward with a finite loss in period 0 and nonnegative gains in the subsequent periods.

**Assumption A.5.** $\exists \overline{\delta} > 0, \forall \delta \in (0, \overline{\delta}]$, (i) $(y_0^*, y_1^* + \delta) \in Y_0$, (ii) $\forall t \in \mathbb{N}, (y_t^* + \delta, y_{t+1}^* + \delta) \in Y_t$, (iii) $r_0(y_0^*, y_1^* + \delta) > -\infty$, and (iv) $\forall t \in \mathbb{N}, r_t(y_t^* + \delta, y_{t+1}^* + \delta) \geq r_t(y_t^*, y_{t+1}^*)$.

**Theorem A.2.** *Assume Assumptions A.1–A.5. Suppose*

$$\exists \epsilon \in (0, \overline{\epsilon}_0], \exists \delta \in (0, \overline{\delta}], \exists \theta \geq 0, \forall t \in \mathbb{N}, \qquad \hat{m}_t(\epsilon) \leq \theta m_t(-\delta). \tag{A.8}$$

*Then* $\overline{\lim}_{t \uparrow \infty} q_t \leq 0$.

*Proof.*[9] By the optimality of $\{y_t^*\}$ and Assumption A.5,

$$r_0(y_0^*, y_1^* + \delta) - r_0(y_0^*, y_1^*) + \sum_{t=1}^{\infty} [r_t(y_t^* + \delta, y_{t+1}^* + \delta) - r_t(y_t^*, y_{t+1}^*)] \leq 0, \tag{A.9}$$

[9]The proof of Kamihigashi [4, Theorem 3.4], which assumes differentiability and an interior optimal path, does not apply to Theorem A.2.

where the infinite sum exists by Assumption A.5(iv). By Assumption A.5(iii) (and footnote 5), for (A.9) to hold, we must have

$$\sum_{t=1}^{\infty}[r_t(y_t^* + \delta, y_{t+1}^* + \delta) - r_t(y_t^*, y_{t+1}^*)] < \infty. \tag{A.10}$$

Dividing through by $\delta$ and recalling (A.3), we get $\sum_{t=1}^{\infty} m_t(-\delta) < \infty$. Thus by (A.8) and Theorem A.1(ii), $\overline{\lim}_{t\uparrow\infty} q_t \leq 0$. □

# B Proofs of Theorems 2.1–2.3

Consider (2.1). Assume Assumptions 2.1–2.3. For $t \in \mathbb{Z}_+$, let $e_t \in F(\Omega, B_t)$ and define

$$r_t(y_t, y_{t+1}) = Ev_t(x_t^* + y_t e_t, x_{t+1}^* + y_{t+1} e_{t+1}), \tag{B.1}$$

$$Y_t = \{(y_t, y_{t+1}) \in \mathbb{R}^2 \,|\, (x_t^* + y_t e_t, x_{t+1}^* + y_{t+1} e_{t+1}) \in X_t\}. \tag{B.2}$$

It follows from (A.2) and (2.3) that under (B.1), $\forall t \in \mathbb{Z}_+$,

$$q_t = \underline{\lim}_{\epsilon\downarrow 0} E\frac{v_t(x_t^*, x_{t+1}^* - \epsilon e_t) - v_t(x_t^*, x_{t+1}^*)}{\epsilon}. \tag{B.3}$$

**Remark B.1.** (B.1) and (B.2) satisfy Assumptions A.1 and A.2. Furthermore the path $\{y_t^*\}_{t=0}^{\infty}$ defined by $\forall t \in \mathbb{Z}_+, y_t^* = 0$, is optimal for (2.4).

## B.1 Proof of Theorem 2.1

Let $\{x_t\}$ be a feasible path satisfying (2.6) and (2.7). Then

$$\{b_t\}_{t=1}^{\infty} \equiv \{Ev_t(x_t^*, x_{t+1}^*) - Ev_t(x_t, x_{t+1})\}_{t=1}^{\infty} \in \Psi, \tag{B.4}$$

where $\Psi$ is defined by (A.5). Consider (2.4) with $\{e_t\} = \{x_t^* - x_t\}$. By Remark B.1, Assumptions A.1–A.3 hold. Assumption A.4 follows from (2.7) and the convexity of $X_t$. Note that $\forall \epsilon \in (0, 1], \forall t \in \mathbb{N}$,

$$m_t(\epsilon) = \frac{Ev_t(x_t^*, x_{t+1}^*) - Ev_t(x_t^* - \epsilon e_t, x_{t+1}^* - \epsilon e_{t+1})}{\epsilon} \tag{B.5}$$

$$\leq Ev_t(x_t^*, x_{t+1}^*) - Ev_t(x_t^* - e_t, x_{t+1}^* - e_{t+1}) = b_t, \tag{B.6}$$

where the inequality holds by concavity. (A.7) now follows. It is easy to see from (B.3), (2.7), the concavity of $v_t$, and the monotone convergence theorem that $\forall t \in \mathbb{Z}_+, q_t = Ev_{t,2}(x_t^*, x_{t+1}^*; -e_t)$. Thus TVC (2.5) holds by Theorem A.1(i).

## B.2 Proof of Theorem 2.2

Consider (2.4) with $\{e_t\} = \{x_t^*\}$. By Remark B.1, Assumptions A.1–A.3 hold. Assumption A.4 follows from Assumption 2.4. Note that $\forall t \in \mathbb{N}, \hat{m}_t(1-\lambda) = \hat{w}_t(\lambda)$. Thus (A.7) holds by (2.11). By (B.3) and Assumption 2.5,

$$\forall t \in \mathbb{Z}_+, \quad Ev_{t,2}(x_t^*, x_{t+1}^*; -x_{t+1}^*) \leq q_t. \tag{B.7}$$

Hence both conclusions hold by Theorem A.1.

## B.3 Proof of Theorem 2.3

Consider (2.4) with $\{e_t\} = \{x_t^*\}$ again. As in the proof of Theorem 2.2, Assumptions A.1–A.4 hold. Assumption A.5 follows from Assumption 2.6. Note that $\forall t \in \mathbb{N}, \hat{m}_t(1-\lambda) = \hat{w}_t(\lambda)$ and $m_t(1-\mu) = w_t(\mu)$. Thus (A.8) holds by (2.16). Recalling (B.7), we see that TVC (2.15) holds by Theorem A.2.

# References


[1] W.A. Brock, On Existence of Weakly Maximal Programmes in a Multi-Sector Economy, *Rev. Econ. Stud.* **37** (1970), 275–280.

[2] I. Ekeland and J.A. Scheinkman, Transversality Conditions for Some Infinite Horizon Discrete Time Optimization Problems, *Math. Oper. Res.* **11** (1986), 216–229.

[3] T. Kamihigashi, A Simple Proof of Ekeland and Scheinkman's Result on the Necessity of a Transversality Condition, *Econ. Theory* **15** (2000), 463–468.

[4] T. Kamihigashi, Necessity of Transversality Conditions for Infinite Horizon Problems, *Econometrica* **69** (2001), 995–1012.

[5] T. Kamihigashi, Necessity of Transversality Conditions for Stochastic Problems, SUNY-Stony Brook, 2001.

[6] T. Kamihigashi, A Simple Proof of the Necessity of the Transversality Condition, *Econ. Theory* **20** (2002), 427–433.

[7] P. Michel, Some Clarifications on the Transversality Condition, *Econometrica* **58** (1990), 705–723.

[8] S. Takekuma, Optimal Growth under Uncertainty: A Complete Characterization of Weakly Maximal Programs, *Hitotsubashi J. Econ.* **33** (1992), 169–182.

[9] M.L. Weitzman, Duality Theory for Infinite Horizon Convex Models, *Manage. Sci.* **19** (1973), 783–789.

[10] I. Zilcha, Characterization by Prices of Optimal Programs under Uncertainty, *J. Math. Econ.* **3** (1976), 173–181.

[11] I. Zilcha, Transversality Condition in a Multi-sector Economy under Uncertainty, *Econometrica* **46** (1978), 515–525.