Economic Theory (2006) 29: 341–362
DOI 10.1007/s00199-005-0023-0

**SPECIAL ISSUE ARTICLE**

**M. Ali Khan · Tapan Mitra**

# Undiscounted optimal growth in the two-sector Robinson-Solow-Srinivasan model: a synthesis of the value-loss approach and dynamic programming

Received: 22 December 2004 / Accepted: 22 August 2005 / Published online: 19 October 2005


**Abstract** We highlight two features of undiscounted optimal growth in the context of a two-sector model due to Robinson, Solow and Srinivasan. First, we use the value-loss approach of Radner-Gale-McKenzie to show a multiplicity of optimal programs in situations when optimality does not coincide with value-loss minimization. Second, we use a theory of undiscounted dynamic programming, not available in the literature, to derive properties of the optimal policy correspondence. In terms of a methodological perspective, we suggest a synthesis of the two methods for the analysis of problems of optimal intertemporal resource allocation.

**Keywords** Undiscounted optimal program · Full-employment program · Golden rule · Value loss · Dynamic programming · Optimal policy function · Transition dynamics

**JEL Classification Numbers** C62 · D90 · O21

This essay is dedicated to Mukul Majumdar on the occasion of his sixtieth birthday, with affection and admiration. We would like to thank Robert Becker, Minako Fujio and Ron Jones for useful discussions and to a referee of this journal for very insightful comments. We are grateful to the *Center for Analytic Economics* at Cornell and to the *Center for a Livable Future* at Johns Hopkins for research support.

M. Ali Khan
Department of Economics, The Johns Hopkins University,
Baltimore, MD 21218, USA
E-mail: akhan@jhu.edu

T. Mitra
Department of Economics, Cornell University,
Ithaca, NY 14853, USA
E-mail: tm19@cornell.edu

## 1 Introduction

The value-loss approach to problems of intertemporal resource allocation, pioneered by Radner (1961) and brought to maturity at the hands of McKenzie (1968, 2002), exploits the price implications of efficiency, basing its essential argumentation on the fact that the planner loses by deviating from price-supported activities. In the context of intertemporal allocation problems without discounting, Brock (1970) showed how the price-support property of a special stationary program (known as the golden-rule) alone can be used to establish the existence of an optimal program as one minimizing value-losses among all "good" programs. This insight has by now been exploited in a variety of publications; indeed, Brock's theorem has been shown to go beyond mere existence to delineate the qualitative properties of optimal programs.[1]

Dynamic programming represents an alternative approach to problems of intertemporal resource allocation. This approach is based on the determination of a value function in the first instance, and subsequently, on the optimal policy correspondence that sustains it. The value function is obtained as the unique solution to the functional equation of dynamic programming, and the optimal policy correspondence can be seen as a solution to the associated maximization problem, involving the solved value function. Thus, while the value-loss approach relies on the methods of convex analysis, the dynamic programming approach belongs in the province of functional analysis.

In this essay, we argue for a synthesis of these two approaches to problems of optimal intertemporal allocation of resources. To be sure, this is not a novel aspiration,[2] but the literature lacks an example of a simple model, in which a complete description of transition dynamics can be effectively obtained by a simultaneous and complementary use of the two approaches.

The model we work with is a special case due to Robinson, Solow and Srinivasan (termed the RSS model in Khan and Mitra 2005a) in which there is only one type of machine, wherein, after normalization of a linear felicity function, the constant amount of labor available in each period, and the sole technique in the consumption-goods sector, are specified by only two real numbers: the labor-capital ratio $a > 0$, and the depreciation rate of the machines $d \in (0, 1)$. In fact, it turns out that optimal behavior in this model is governed effectively by the value of one parameter, $\xi = (1/a) - (1 - d)$, which measures the rate of transformation between machines today and machines tomorrow, while maintaining full-employment of both labor and capital. This two-sector RSS model is also a special case of the two-sector Leontief model considered in Nishimura and Yano (1995), which in itself is a special case of the standard neoclassical two-sector model of Srinivasan (1964) and Uzawa (1964).[3]

We provide an analysis of optimal dynamic behavior in the two-sector RSS model when future utilities are not discounted. This leads us to a thorough study

[1] See, for example, the Faustmann solution in Mitra and Wan (1986), and the Stiglitz policy in Khan and Mitra (2005a,b).

[2] The work of Weitzman (1973) is an outstanding example of the point of view that we advocate.

[3] Since the investment goods sector uses labor as its only factor of production, the two-sector RSS model is an extreme instance of a setting in which the consumption sector is capital-intensive and there is no reversal of factor intensities, assumptions which have been emphasized in recent work pertaining to cycles and chaos in economic theory.

of the value-loss approach in the context of our simple model (see Section 3) and reveals some aspects of the general theory of optimal intertemporal allocation that might not be fully appreciated. Specifically, it highlights the fact that while value-loss minimization implies optimality (a fact implicit in Brock's demonstration of the existence of an optimal program), optimality does not always imply value-loss minimization. In the case in which the equivalence breaks down ($\xi = 1$), there exist optimal programs that yield *higher* value-losses at golden-rule prices than other (also optimal) programs, and this phenomenon is therefore seen to be intimately tied to the non-uniqueness (in fact, indeterminacy) of optimal programs. In the case in which the equivalence holds ($\xi \neq 1$), a turnpike property of good programs can be established, leading to an appropriate transversality condition being satisfied.

The turnpike property of good programs (when $\xi \neq 1$) is the key to developing a completely satisfactory theory of dynamic programming in the undiscounted case (see Section 4). The results on dynamic programming which we present are, consequently, based (indirectly) on the value-loss approach. [4] However, once derived, the dynamic programming results can be employed to solve for (some regions of) the optimal policy correspondence in a more streamlined and transparent manner than the value-loss approach. On the other hand, the application of the value-loss approach provides a more direct route to identify optimal behavior for other regions. Thus, we show how a combination of the value-loss and dynamic programming approaches can be employed effectively to solve completely for the optimal policy correspondence in this simple model.

The explicit solution of the optimal policy correspondence allows us to examine the nature of optimal transition dynamics in a way that the general theory of intertemporal allocation does not allow us to do (since it can only discuss aspects of long-run optimal behavior). Two features of the transition dynamics (discussed in Section 5) stand out. First, starting from initial stocks below the golden-rule, optimality always requires an over-building phase, leading to stocks above the golden-rule stock, even when the long-run behavior warrants a convergence to the golden-rule stock. Second, even when it is feasible to fully utilize the available capital stock, it is not always optimal to do so, making it possible to observe excess-capacity and production of machines in the same period.

The first feature is not only quite distinct from the monotone convergence result in the corresponding continuous-time model[5]. It makes this extremely simple two-sector model qualitatively different from discrete-time aggregative models in which monotone optimal behavior is a consequence of the "non-crossing property". The second feature is especially noteworthy because it directly contradicts a central policy prescription of Stiglitz (1968), obtained from a continuous-time analysis of the RSS model: a Stiglitz policy would require that all the existing capital stock (in the consumption good sector) be manned, if possible, and the remaining labor (if any) be used to produce new machines.

[4] In contrast to the usual approach to dynamic programming in the discounted case, which relies on a contraction mapping theorem in an appropriate function space, our approach to undiscounted dynamic programming makes use of the methods of convex analysis.

[5] See Stiglitz (1968) and Khan and Mitra (2003).

## 2 Preliminaries

### 2.1 The model

We work with a special case of a two-sector model in which the production of machines requires only labor. In the consumption-good sector, a single consumption good is produced by infinitely divisible labor and machines with the further Leontief specification that a unit of labor and a unit of a machine produce a unit of the consumption good. In the investment-good sector, only labor is required to produce machines, with $a > 0$ units of labor producing a single machine. Machines depreciate at the rate $0 < d < 1$. A constant amount of labor, normalized to unity, is available in each time period $t \in \mathbb{N}$, where $\mathbb{N}$ is the set of non-negative integers. Thus, in the canonical formulation of McKenzie (1968, 2002), the collection of production plans $(x, x')$, the amount $x'$ of machines at the end of the next period (tomorrow) from the amount $x$ available at the end of the current period (today), is given by the *transition possibility set*:

$$\Omega = \{(x, x') \in \mathbb{R}^2_+ : x' - (1-d)x \geq 0, \text{ and } a(x' - (1-d)x) \leq 1\}$$

where $z \equiv (x' - (1-d)x)$ is the number of machines that are produced during the next period, and $z \geq 0$ and $az \leq 1$ respectively formalize constraints on reversiblity of investment and the use of labor. For any $(x, x') \in \Omega$, one can consider the amount $y$ of the machines available for the production of the consumption goods, leading to a correspondence $\Lambda : \Omega \longrightarrow \mathbb{R}_+$ with $\Lambda(x, x') = \{y \in \mathbb{R}_+ : 0 \leq y \leq x \text{ and } y \leq 1 - a(x' - (1-d)x)\}$. Welfare is derived only from the consumption good and is represented by a linear function, normalized so that $y$ units of the consumption good yields a welfare level $y$. A *reduced form utility function,* $u : \Omega \to \mathbb{R}_+$ with $u(x, x') = \max\{y \in \Lambda(x, x')\}$ indicates the maximum welfare level that can be obtained today, if one starts with $x$ of machines today, and ends up with $x'$ of machines tomorrow, where $(x, x') \in \Omega$.

An *economy* consists of a pair $(a, d)$, and the following concepts apply to it. A *feasible program* from $x_o$ is a sequence $\{x(t), y(t)\}$ such that $x(0) = x_o$, and for all $t \in \mathbb{N}$, $(x(t), x(t+1)) \in \Omega$ and $y(t) \in \Lambda((x(t), x(t+1))$. A *program* from $x_o$ is a sequence $\{x(t), y(t)\}$ such that $x(0) = x_o$, and for all $t \in \mathbb{N}$, $(x(t), x(t+1)) \in \Omega$ and $y(t) = \max \Lambda((x(t), x(t+1)) = u(x(t), x(t+1))$. A *program* $\{x(t), y(t)\}$ is a program from $x(0)$, and associated with it is a *gross investment sequence* $\{z(t+1)\}$, defined by $z(t+1) = (x(t+1) - (1-d)x(t))$ and a *consumption sequence* $\{c(t+1)\}$ defined by $c(t+1) = y(t)$ for all $t \in \mathbb{N}$. It is easy to check that every program $\{x(t), y(t)\}$ is bounded by $M(x(0)) \equiv \max\{x(0), \bar{x}\}$, where $\bar{x} = (1/ad)$ is the *maximum sustainable capital stock* . A program $\{x(t), y(t)\}$ is called *stationary* if $(x(t), y(t)) = (x(t+1), y(t+1))$ for all $t \in \mathbb{N}$. For a stationary program $\{x(t), y(t)\}$, we have $x(t) \leq \bar{x}$ for all $t \in \mathbb{N}$.

A program $\{x^*(t), y^*(t)\}$ from $x_o$ is called *optimal* if:

$$\lim_{T\to\infty} \inf \sum_{t=0}^{T} [u(x(t), x(t+1)) - u(x^*(t), x^*(t+1))] \leq 0$$

for every program $\{x(t), y(t)\}$ from $x_o$. A *stationary optimal program* is a program that is stationary and optimal.

### 2.2 Full employment programs

Our analysis in the following sections will be facilitated by focusing on *full-employment programs*, $\{x(t), y(t)\}$, which are programs satisfying: $y(t) + a[x(t+1) - (1-d)x(t)] = 1$ for all $t \in \mathbb{N}$. In this subsection, we note some basic properties relating to programs and full-employment programs, which will be useful in this connection.

It is fairly obvious that given any program, there is a full-employment program which has the *same* consumption sequence. Slightly more subtle is the property that higher initial stocks always allow programs which have *dominating* consumption sequences (at least as high in all periods, higher in at least one period). It is a consequence of this property that given any program which is *not* a full-employment program, there is a full-employment program (from the same initial stock) that has a dominating consumption sequence. We formally summarize these results in the following Proposition.[6]

**Proposition 1** *(1) If $\{x(t), y(t)\}$ is a program from $x$, then there is a full-employment program $\{x'(t), y'(t)\}$ from $x$, such that $y'(t) = y(t)$ for $t \geq 0$.*

*(2) If $\{x(t), y(t)\}$ is a program from $x$, and $x' > x$ then there is a full-employment program $\{x'(t), y'(t)\}$ from $x'$, such that $y'(t) \geq y(t)$ for all $t \in \mathbb{N}$, and $y'(t) > y(t)$ for some $t \in \mathbb{N}$.*

*(3) If $\{x(t), y(t)\}$ is a program from $x$, which is not a full-employment program, then there is a full-employment program $\{x'(t), y'(t)\}$ from $x$, such that $y'(t) \geq y(t)$ for all $t \in \mathbb{N}$, and $y'(t) > y(t)$ for some $t \in \mathbb{N}$.*

### 2.3 A golden rule

The concept of a *golden-rule* plays a fundamental role in our analysis. Formally, we define a stock $\hat{x} \in \mathbb{R}_+$ as a *golden-rule stock* if $(\hat{x}, \hat{x}) \in \Omega$ and $u(\hat{x}, \hat{x}) \geq u(x, x')$ for all $(x, x') \in \Omega$ with $x' \geq x$. We record in the following result the existence and uniqueness of a golden-rule stock, and its "price support". It is useful to note that the golden-rule stock is *unique*, even though there is no strict concavity-convexity feature in the specification of our model.

**Proposition 2** *(1)The pair $(\hat{x}, \hat{p}) = (1/(1+ad), a/(1+ad))$ satisfies $(\hat{x}, \hat{x}) \in \Omega$, and:*

$$u(\hat{x}, \hat{x}) \geq u(x, x') + \hat{p}x' - \hat{p}x \quad for\ all\ (x, x') \in \Omega \tag{1}$$

*(2) $\hat{x}$ is a unique golden-rule stock.*

We refer to $\hat{p}$, given by Proposition 2, as the *golden-rule price* and to the pair $(\hat{x}, \hat{p})$ as the *golden-rule*. The golden-rule price provides a *price-support* to the golden-rule stock in the sense conveyed precisely by (1). Note that the golden-rule consumption level is $\hat{y} = \hat{x}$. The *von Neumann facet* is the set of those $(x, x') \in \Omega$ for which equality holds in (1).

[6] The proofs, being fairly straightforward, are omitted.

The *value-loss* (relative to the golden-rule) from operating at $(x, x') \in \Omega$ and $y = u(x, x')$ is:

$$\delta(x, x', y) = u(\hat{x}, \hat{x}) - [u(x, x') + \hat{p}x' - \hat{p}x] \geq 0 \tag{2}$$

Given that there are two factors of production, intuition suggests only two possible sources of value-loss: excess-capacity of capital and unemployment of labor. This can be formalized by splitting up the value-loss in (2) into two parts as follows:

$$\left.\begin{array}{l} \alpha(x, x', y) = \hat{p}d(x - y) \geq 0 \\ \beta(x, x', y) = (\hat{p}/a)(1 - y - a(x' - (1 - d)x)) \geq 0 \\ \delta(x, x', y) = \alpha(x, x', y) + \beta(x, x', y) \geq 0 \end{array}\right\} \tag{3}$$

Let $\{x(t), y(t)\}$ be a program; then, $(x(t), x(t+1)) \in \Omega$ and $y(t) = u(x(t), x(t+1))$ for $t \in \mathbb{N}$. Thus, using (2) and (3), we have for $t \in \mathbb{N}$ :

$$\begin{aligned} u(\hat{x}, \hat{x}) &= u(x(t), x(t+1)) + \hat{p}x(t+1) - \hat{p}x(t) + \alpha(t) + \beta(t) \\ &\equiv u(x(t), x(t+1)) + \hat{p}x(t+1) - \hat{p}x(t) + \delta(t) \end{aligned} \tag{4}$$

where $\alpha(t) = \alpha(x(t), x(t+1), y(t))$, $\beta(t) = \beta(x(t), x(t+1), y(t))$ and $\delta(t) = \delta(x(t), x(t+1), y(t))$ for $t \in \mathbb{N}$. This yields a useful identity, relating the sum of value losses to the sum of utility differences from the golden rule utility level, along any program $\{x(t), y(t)\}$ and any $T \in \mathbb{N}$ :

$$\sum_{t=0}^{T}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] = \hat{p}x(0) - \hat{p}x(T+1) - \sum_{t=0}^{T}\delta(t) \tag{SVL}$$

We can eliminate the variable $y(t) = u(x(t), x(t+1))$ from (4), by using (3), to obtain for all $t \in \mathbb{N}$ :

$$x(t+1) = (1/a) - \xi x(t) + A(t) - B(t)$$

where $\xi = [(1/a) - (1 - d)]$, and $A(t) = (1/ad\hat{p})\alpha(t)$, $B(t) = (1/\hat{p})\beta(t)$ for $t \in \mathbb{N}$. On measuring capital stocks relative to the golden-rule stock, $X(t) = (x(t) - \hat{x})$, and on noting that $\hat{x} = (1/a) - \xi\hat{x}$, we can obtain the basic dynamic equation of this essay:

$$X(t+1) = (-\xi)X(t) + A(t) - B(t) \quad for\ all\ t \in \mathbb{N} \tag{5}$$

The parameter, $\xi$, which figures prominently in our analysis, represents the rate of transformation between machines today and machines tomorrow, while maintaining full-employment of both labor and capital. Without further explicit mention, $\xi$ will be assumed to be positive in what follows. [7]

[7] The reader can look at Khan and Mitra (2005b) for an analysis of the case in which $\xi \leq 0$.

### 2.4 Long-run dynamics of good programs

A program $\{x(t), y(t)\}$ is called *good* if there is a real number $G$ such that for all $T \in \mathbb{N}$, we have $\sum_{t=0}^{T}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] \geq G$. The general theory of intertemporal resource allocation emphasizes the average turnpike property as the basic asymptotic result of good programs. The two-sector RSS model is fully affirmative (as it must be) of the general theory in its abstract form, but its particular structure yields an interesting parametric classification, depending on the value of our key parameter, $\xi$, which we now describe.

In the case $\xi \neq 1$, every good program is seen to converge to the golden-rule stock, a *turnpike* property. When $\xi = 1$, in contrast, there are good programs which exhibit a period-two cycle, and every good program is seen to either converge to the golden-rule stock or to converge to a period-two cycle around the golden-rule (see Theorem 1). These results are derived by using the price-support property in Proposition 2, and the basic dynamic equation (5). They nicely illustrate why the *average turnpike* property of good programs is indeed the general result of the subject.

**Theorem 1** *(1) There exists a good program $\{x(t), y(t)\}$ from every $x_o \in \mathbb{R}_+$.*

*(2) If $\{x(t), y(t)\}$ is a good program, then $\sum_{t=0}^{\infty} \delta(t) < \infty$.*

*(3) If $\xi \neq 1$, and $\{x(t), y(t)\}$ is a good program, then $(x(t), y(t)) \to (\hat{x}, \hat{y})$ as $t \to \infty$, and:*

$$\sum_{t=0}^{\infty}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] = \hat{p}x(0) - \hat{p}\hat{x} - \sum_{t=0}^{\infty} \delta(t) \qquad \text{(US)}$$

*(4) Suppose $\xi = 1$, and $\{x(t), y(t)\}$ is a good program, such that $\tilde{x}$ is an arbitrary limit point of $\{x(t)\}$. Then, either $x(t)$ converges to $\tilde{x}$ for all odd periods, and to $(2\hat{x} - \tilde{x})$ for all even periods, or $x(t)$ converges to $\tilde{x}$ for all even periods, and to $(2\hat{x} - \tilde{x})$ for all odd periods.*

## 3 Optimality and value-loss minimization

In this section we study the implications of the value-loss approach to dynamic optimization problems in the context of the RSS model. The focus of the study is the relation between optimality and value-loss minimization.

We start with the observation that value-loss minimization is *sufficient* for optimality, and make this observation operational (in checking for optimality) by obtaining a formula which enables us to keep track of the sum of value-losses along (full-employment) programs. These results are useful in characterizing the optimal policy correspondence in section 4.

We then show that value-loss minimization is *necessary* for optimality when $\xi \neq 1$, and this becomes the key ingredient in developing the dynamic programming approach in section 4. When $\xi = 1$, value-loss minimization is not necessary for optimality, and this observation can be seen to be closely related to the result that there are initial stocks from which optimal programs are indeterminate.

### 3.1 Value-loss minimization implies optimality

Brock (1970) showed that value-loss minimization is *sufficient* for optimality; therefore, establishing the existence of a program which minimizes the sum of value-losses enables one to establish the existence of an optimal program. For easy reference, we state his result for our framework.

**Proposition 3** *(1) If* $\{x(t), y(t)\}$ *is a program from* $x \in \mathbb{R}_+$, *such that:*

$$\sum_{t=0}^{\infty} \delta(t) \leq \sum_{t=0}^{\infty} \delta'(t) \tag{VLM}$$

*for every program* $\{x'(t), y'(t)\}$ *from* $x$, *then* $\{x(t), y(t)\}$ *is optimal from* $x$.

*(2) If* $x \in \mathbb{R}_+$, *there is a program* $\{x(t), y(t)\}$ *from* $x$, *such that:*

$$\sum_{t=0}^{\infty} \delta(t) \leq \sum_{t=0}^{\infty} \delta'(t)$$

*for every program* $\{x'(t), y'(t)\}$ *from* $x$.

*(3) If* $x \in \mathbb{R}_+$, *there is an optimal program from* $x$.

Statement (1) of the above Proposition can provide a useful criterion for *characterizing* optimality, if there is a convenient way to express the magnitude of the sum of value losses along a program. Our next result provides the central result in this direction by showing that the sum of value-losses along full-employment programs is related to the variation in the value of capital stocks over time.

**Lemma 1** *(1) For any full-employment program* $\{x(t), y(t)\}$,

$$\sum_{t=0}^{T}[\alpha(t)/(-\xi)^t] = ad\xi\{\hat{p}X(0) - \hat{p}[X(T+1)/(-\xi)^{T+1}]\} \quad for\ all\ T > 1 \tag{6}$$

*(2) Any optimal program is a full-employment program, and satisfies (6).*

### 3.2 Optimality does not imply value-loss minimization

In our framework, good programs exist from every initial stock. Consequently, optimal programs are necessarily good; so in providing the parametric classification of long-run dynamics of good programs (stated in Theorem 1 above), the RSS model illustrates an important point that ought to be emphasized. When $\xi \neq 1$, optimality implies value-loss minimization, since the relevant "transversality condition":

$$\lim_{t \to \infty} \hat{p}(x(t) - \hat{x}) = 0$$

is satisfied along an optimal program $\{x(t), y(t)\}$.

**Proposition 4** *If* $\{x(t), y(t)\}$ *is an optimal program from* $x \in \mathbb{R}_+$, *then it must be good. If* $\xi \neq 1$, *then:*

$$\lim_{t\to\infty} \hat{p}(x(t) - \hat{x}) = 0 \tag{7}$$

*and:*

$$\sum_{t=0}^{\infty} \delta(t) \leq \lim_{T\to\infty} \sum_{t=0}^{T} \delta'(t) \tag{VLM}$$

*for every program* $\{x'(t), y'(t)\}$ *from* $x$. *Furthermore, we have:*

$$\sum_{t=0}^{\infty} [u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] = \hat{p}x(0) - \hat{p}\hat{x} - \sum_{t=0}^{\infty} \delta(t) \tag{US}$$

A consequence of the above Proposition is that one can develop a completely satisfactory theory of undiscounted dynamic programming in the case $\xi \neq 1$, comparable to the theory available in the discounted case, a topic we address fully in the next section.

However, in the case $\xi = 1$, the transversality condition (7) is not necessarily satisfied along optimal programs, and optimality does not imply value-loss minimization. It is a consequence of this observation, that when $\xi = 1$, optimal programs are no longer unique and, in fact, we can have indeterminacy of optimal programs (see Theorem 2 below).

This is most conveniently demonstrated by using Lemma 1 to show that if $\xi = 1$, then at least for initial stocks above and "close to" the golden-rule stock, there is a continuum of optimal programs. Two of these optimal programs can be described simply. The program with zero value loss in every period generates period-two cycles around the golden-rule and is optimal since it clearly minimizes the sum of value-losses. However, the ("straight down the turnpike") program which reaches the golden-rule stock in one period, and thereafter stays there, is also optimal, and it does not minimize the sum of value-losses.

**Theorem 2** *Suppose* $\xi = 1$ *and* $x_o \in (\hat{x}, 1)$.

*(1) The full-employment program* $\{x(t), y(t)\}$ *from* $x_o$, *satisfying* $x(t) = \hat{x}$ *for all* $t \geq 1$, *is optimal from* $x_o$.

*(2) The full-employment program* $\{x'(t), y'(t)\}$ *from* $x_o$, *satisfying* $x'(t) = x_o$ *for all even* $t$ *and* $x'(t) = 2\hat{x} - x_o$ *for all odd* $t$, *is optimal from* $x_o$.

*(3) There is a continuum of optimal programs from* $x_o$.

## 4 Optimal policy function

In this section, we provide an explicit solution of the optimal policy correspondence for our model, which provides a complete characterization of optimal behavior. One can infer from it the nature of optimal behavior in the long-run (asymptotic properties) as well as in the short-run (transition dynamics).

The *optimal policy correspondence* for our model can be defined as a correspondence, $h$, from $\mathbb{R}_+$ to subsets of $\mathbb{R}_+$ such that, given any $x \in \mathbb{R}_+$, there is an optimal program $\{x(t), y(t)\}$ from $x$, with $x(1) \in h(x)$. Since $\Omega$ is convex, any convex combination of optimal programs from $x$ is a program from $x$, and since $u$ is concave on $\Omega$, it too must be optimal from $x$. Thus, $h$ is a convex-valued correspondence.

The optimal policy correspondence is a concept belonging principally to the dynamic programming approach to optimization problems, and so naturally we rely on the theory of dynamic programming to solve it. In this connection, let us make two observations. First, a theory of dynamic programming in the undiscounted case, exactly analogous to that in the discounted case, can be developed when $\xi \neq 1$, using the turnpike property of good programs (Theorem 1). Thus our basic results on dynamic programming (stated in the first subsection below) are derived using the value-loss approach. Second, while the dynamic programming approach is ideally suited to solving for the optimal policy correspondence for certain domains ("high" and "low" capital stocks), the value-loss approach (especially Propositions 3, 4 and Lemma 1) to optimization problems appears to be better suited to solving it for other domains (the "middle section" of capital stocks). Solving completely for the optimal policy correspondence essentially uses a synthesis of the two approaches. This method, independent of the result itself, seems to us to be of potential use in the study of other dynamic optimization problems.

## 4.1 Dynamic programming

In this subsection, we develop the principal results of dynamic programming for our model with the standing hypothesis of $\xi \neq 1$. For this purpose, we need to define the central concept of a *value function* for our model. To this end, note that, given $x \in \mathbb{R}_+$, any optimal program $\{x(t), y(t)\}$ from $x$ is good (by Proposition 4), has $\sum_{t=0}^{\infty} \delta(t) < \infty$ (by Theorem 1), and minimizes the sum of value losses (by Proposition 4). Therefore, it has the same sum of value losses as any other optimal program from $x$. Denote this common value (finite by Theorem 1) by $\delta(x)$. Thus, (US) of Proposition 4 can be rewritten as:

$$\sum_{t=0}^{\infty}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] = \hat{p}x - \hat{p}\hat{x} - \delta(x) \tag{8}$$

We can therefore define a value function, $V : \mathbb{R}_+ \to \mathbb{R}$, by:

$$V(x) \equiv \sum_{t=0}^{\infty}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] \tag{9}$$

where $\{x(t), y(t)\}$ is any optimal program from $x$, since the right-hand side of (8) is independent of the optimal program chosen.[8] Note that for every good program $\{x'(t), y'(t)\}$ from $x$, we have by using (US) of Theorem 1(3), and (8) and (9),

[8] A similar device is used by Brock and Majumdar (1988) to define the value function in the undiscounted case. However, in their framework, they follow Gale (1967) and assume strict concavity of the utility function, so that optimal programs are unique by assumption. In our framework, the utility function is not strictly concave. However, when $\xi \neq 1$, the turnpike property holds, and this ensures that the limit in (9) is well-defined and independent of the actual optimal program chosen.

$$V(x) \geq \sum_{t=0}^{\infty}[u(x'(t), x'(t+1)) - u(\hat{x}, \hat{x})] \tag{10}$$

We define the transition correspondence from $\mathbb{R}_+$ to subsets of $\mathbb{R}_+$ by $\Omega(x) = \{x' \in \mathbb{R}_+ : (x, x') \in \Omega\}$. We can now summarize the principal result on dynamic programming as follows.

**Proposition 5** *(1) The value function $V$ is a concave and strictly increasing function on $\mathbb{R}_+$ and continuous on $\mathbb{R}_{++}$, satisfying $V(\hat{x}) = 0$ ; (2) $V$ satisfies the functional equation of dynamic programming $V(x) = max_{x' \in \Omega(x)}\{[u(x, x') - u(\hat{x}, \hat{x})] + V(x')\}$;(3) $\{x(t), y(t)\}$ is an optimal program if and only if for all $t \in \mathbb{N}$, $V(x(t)) = [u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] + V(x(t+1))$.*

Since any program starting from a point in $X \equiv [0, 1/ad]$ remains in $X$, we will use $X$ as the state space, and confine our solution of the optimal policy correspondence to this domain.[9]

The above result on dynamic programming helps us to pin down the optimal policy correspondence for certain ranges of stocks in $X$, when $\xi \neq 1$. To state this precisely, let us denote:

$$k \equiv \hat{x}/(1-d),\ \ A = [0, \hat{x}],\ \ B = (\hat{x}, k),\ \ C = [k, 1/ad] \tag{11}$$

We refer to $A$ as the range of "low stocks" and to $C$ as the range of "high stocks". For these two ranges of stocks, Proposition 5 yields the following characterization of the optimal policy correspondence (when $\xi \neq 1$).

**Corollary 1** *Suppose $\xi \neq 1$. Then the optimal policy correspondence, $h : X \to \mathbb{R}_+$ satisfies:*

$$h(x) = \begin{cases} (1/a) - \xi x & for\ all\ x \in A \\ (1-d)x & for\ all\ x \in C \end{cases} \tag{12}$$

### 4.2 The value-loss approach

The value-loss approach essentially utilizes the relation between optimality and value-loss minimization, that we developed in section 3. Using this approach, and still confining ourselves to the case $\xi \neq 1$, one can solve for the optimal policy function for the middle range of stocks, denoted by $B$ in (11), thereby completing the full description of the optimal policy function on $X$, when $\xi \neq 1$.

Specifically, in the case in which $\xi < 1$, one can define a program, starting from any initial stock in $B$, which has minimum value loss in the initial period (among all programs starting from the same initial stock) and zero value-loss in all subsequent periods, and is therefore optimal. Furthermore, any first-period deviation from this program produces a larger initial value-loss, so that the policy function is given by exactly the first-period stock along the specified program.

[9] After the solution has been presented on $X$, the reader should have no difficulty in checking that $(1-d)x \in h(x)$ for $x > (1/ad)$.

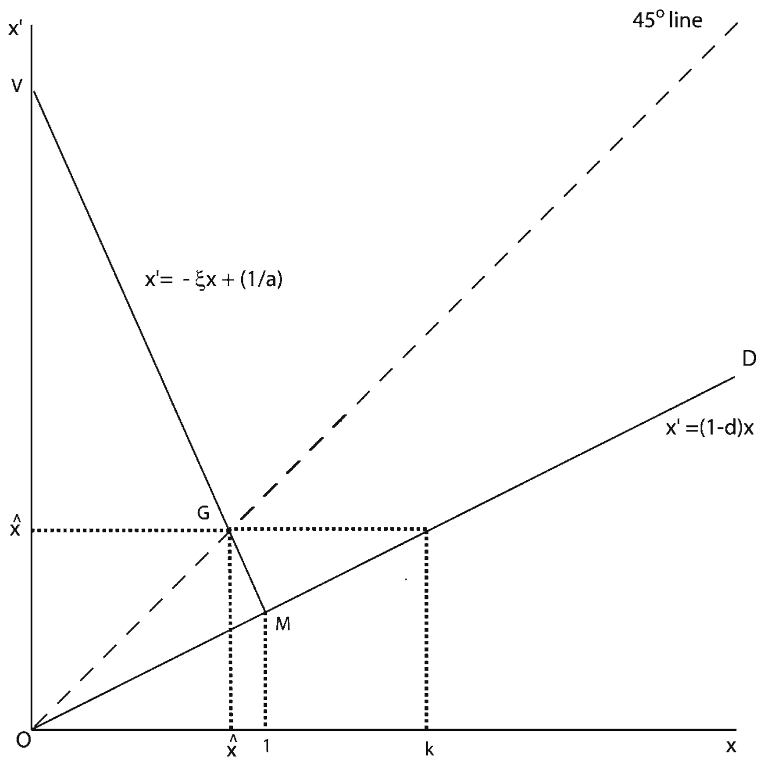


**Fig. 1** Optimal policy function

In the case $\xi > 1$, the program which reaches the golden-rule stock in the first period and stays there forever afterwards can be shown to have the minimum *sum* of value losses by a direct application of Lemma 1, although it clearly does not have minimum value loss in the initial period. The bifurcation in optimal behavior seen by comparing the case $\xi < 1$ with $\xi > 1$ is the key to understanding the optimal dynamics in this model.

**Corollary 2** *(1) Suppose* $\xi > 1$. *Then the optimal policy correspondence,* $h : X \to \mathbb{R}_+$ *satisfies:*

$$h(x) = \hat{x} \quad for\ x \in (\hat{x}, k) \tag{13}$$

*(2) Suppose* $\xi < 1$.*Then the optimal policy correspondence,* $h : X \to \mathbb{R}_+$ *satisfies:*

$$h(x) = \begin{cases} (1/a) - \xi x & for\ x \in (\hat{x}, 1) \\ (1-d)x & for\ x \in [1, k) \end{cases} \tag{14}$$

We treat the borderline case of $\xi = 1$ last. In this case, we expect (in view of the bifurcation mentioned above, as well as the result in Theorem 2) to observe non-unique optimal behavior, and this is indeed the case. In fact, any first-period stock between the first-period optimal stocks in the above two cases (of $\xi < 1$ and $\xi > 1$) can be seen to be optimal in this case.

**Corollary 3** *Suppose $\xi = 1$. Then the optimal policy correspondence, $h : X \to \mathbb{R}_+$ satisfies:*

$$\begin{array}{ll} (1/a) - \xi x \in h(x) & for\ x \in A \\ [(1/a) - \xi x, \hat{x}] \subset h(x) & for\ x \in (\hat{x}, 1) \\ [(1-d)x, \hat{x}] \subset h(x) & for\ x \in [1, k) \\ (1-d)x \in h(x) & for\ x \in C \end{array} \tag{15}$$

## 5 Discussion of transition dynamics

Given the full description of optimal behavior, through an explicit solution of the optimal policy correspondence, we devote this section to a discussion of some aspects of optimal transition dynamics.

Probably the aspect of optimal behavior that stands out most clearly in this model is the phenomenon of "overbuilding" starting from initial stocks below the golden-rule. We *never* have monotonic convergence to the golden-rule starting from such initial stocks, as one would have in a *continuous-time* version of this model. But a continuous-time version misses out on a crucial ingredient, the "time to build" aspect of durable capital goods.[10] After the phase of over-building, optimality requires (except in the borderline case of $\xi = 1$) that there be convergence to the golden-rule stock (either immediately, as in the case of $\xi > 1$, or asymptotically, as in the case $\xi < 1$, when overbuilding leads to stocks in the "middle range").

This overbuilding phenomenon also distinguishes optimal behavior in the two-sector RSS model from other *discrete-time* models, in which a "non-crossing" property holds.[11] The non-crossing property implies monotonic optimal behavior over time in all dynamic models with a stationary structure. Although our model appears to be as simple as (in some respects, even simpler than) some of the aggregative models in which the non-crossing property holds, it generates considerably richer optimal dynamics.

The other aspect of optimal behavior that is noteworthy stems from the fact that our framework is a stripped-down version of the RSS model, used in studying optimal choice of technique in the process of development. In that framework, a continuous-time analysis by Stiglitz (1968) revealed an important policy prescription. In the case of a single type of machine (the special case of the RSS model that we are dealing with) this policy warrants that labor be first allocated to man the machines in the consumption goods sector, and only the remaining labor (if any) be allocated to produce new machines. Thus, it would never be optimal to

[10] This aspect has been emphasized as an important ingredient in producing optimal fluctuations in other frameworks with durable capital; see, among others, Kydland and Prescott (1982) and Mitra and Wan (1986).

[11] See, for example, Majumdar and Nermuth (1982) and Dechert and Nishimura (1983).

have excess-capacity, so long as it is feasible to man the stock of capital in the consumption-goods sector (that is, so long as the initial stock did not exceed unity). However, in our discrete-time formulation, we clearly see that when $\xi > 1$, and the initial stock is between the golden-rule stock and unity, it is optimal to decumulate to precisely the golden-rule stock (and no further), thereby using less than the full capacity of the consumption good sector.[12]

## 6 Proofs

*Proof of Proposition 2* (1) It is straightforward to check that $(\hat{x}, \hat{x}) \in \Omega$ and $u(\hat{x}, \hat{x}) = \hat{x}$. Now, denote for $(x, x') \in \Omega$ and $y \in \Lambda(x, x')$, $d\hat{p}(x - y)$ by $\alpha(x, x', y)$, and $(\hat{p}/a)\{1-y-a[x'-(1-d)x]\}$ by $\beta(x, x', y)$. Then, $\alpha(x, x', y) \geq 0$ and $\beta(x, x', y) \geq 0$. The result now follows from the following computation:

$$
\begin{aligned}
y + \hat{p}x' - \hat{p}x &= y + \hat{p}[x' - (1-d)x] - d\hat{p}x \\
&= [1/(1+ad)]y + \hat{p}[x' - (1-d)x] - \alpha(x, x', y) \\
&= [1/(1+ad)]\{y + a[x' - (1-d)x]\} - \alpha(x, x', y) \\
&= \hat{x}\{y + a[x' - (1-d)x]\} - \alpha(x, x', y) \\
&= \hat{x} - \hat{x}\{1 - y - a[x' - (1-d)x]\} - \alpha(x, x', y) \\
&= \hat{x} - \beta(x, x', y) - \alpha(x, x', y)
\end{aligned} \tag{16}
$$

(2) If $x$ is a golden-rule stock, then $(x, x) \in \Omega$ and $y \in \Lambda(x, x)$, with $y = \hat{x}$. Thus, using (16), we must have $\beta(x, x, y) = \alpha(x, x, y) = 0$, so that $x = y = \hat{x}$. □

*Proof of Theorem 1* (1) Define $y(0) = 0$, and $y(t+1) = (1-d)y(t) + d\hat{x}$ for $t \geq 0$. Then, $y(t)$ is monotonically non-decreasing, and converges to $\hat{x} = \hat{y}$ as $t \to \infty$. Define $z(t+1) = d\hat{x}$ for $t \geq 0$. Given an arbitrary initial stock, $x$, define $x(0) = x$, and $x(t+1) = (1-d)x(t) + z(t+1)$ for $t \geq 0$. Then, it is easy to check that $\{x(t), y(t)\}$ is a program from $x$. Given the definition of the sequence $\{y(t)\}$, we have $(y(t) - \hat{y}) = (1-d)^t(y(0) - \hat{y})$ for $t \geq 0$. Thus, the sequence $\{\hat{y} - y(t)\}$ is summable, and so $\{x(t), y(t)\}$ is a good program from $x$.

(2) Using (SVL), we have, for $T \in \mathbb{N}$ :

$$\sum_{t=0}^{T}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] = \hat{p}x(0) - \hat{p}x(T+1) - \sum_{t=0}^{T}\delta(t)$$

Since $\{x(t), y(t)\}$ is good, there is a real number $G$, such that $\sum_{t=0}^{T}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] \geq G$, for all $T \in \mathbb{N}$. So, $\sum_{t=0}^{T}\delta(t) \leq \hat{p}x(0) - G$ for all $T \in \mathbb{N}$, and consequently $\sum_{t=0}^{\infty}\delta(t) < \infty$.

[12] This aspect is explored more fully in Khan and Mitra (2005a), where Stiglitz policies are explicitly formulated (in a discrete-time version of the RSS model, with many types of machines) and compared to optimal policies.

(3) Note that since $\{x(t), y(t)\}$ is good, we can infer from (2) above and equation (3) that $\sum_{t=0}^{\infty} A(t) < \infty$, and $\sum_{t=0}^{\infty} B(t) < \infty$. This implies that given any $\varepsilon > 0$, we can choose a positive integer $\tau \in \mathbb{N}$, such that:

$$\sum_{t=\tau}^{\infty} A(t) + \sum_{t=\tau}^{\infty} B(t) < \varepsilon/3 \tag{17}$$

Further, defining $D = 2M(x)$, we have $|X(t)| \leq D$ for $t \in \mathbb{N}$, since $x(t) \leq M(x)$ for $t \geq 0$, and $\hat{x} \leq \bar{x} \leq M(x)$. Note that since $0 < \xi = (1/a) - (1-d)$, we have two possible cases: (a) $0 < \xi < 1$, and (b) $\xi > 1$.Thus, we can find a positive integer $\nu$ such that:

$$\min[D\xi^{\nu}, D/\xi^{\nu}] < \varepsilon/3 \tag{18}$$

Now, pick any positive integer, $T \geq \tau$. Then, for all positive integers $n \geq \nu$, we have by iterating on (5):

$$X(T+n) = (-\xi)^n X(T) + \sum_{s=0}^{n-1} (-\xi)^{n-1-s} A(T+s) - \sum_{s=0}^{n-1} (-\xi)^{n-1-s} B(T+s) \tag{19}$$

Now, under case (a), we use (17) and (18) in (19) to get $|X(T+n)| < \varepsilon$. In case (b), we divide (19) by $(-\xi)^n$ to obtain:

$$\begin{aligned}(1/(-\xi)^n)X(T+n) = X(T) &+ \sum_{s=0}^{n-1} (1/(-\xi)^{1+s}) A(T+s) \\ &- \sum_{s=0}^{n-1} (1/(-\xi)^{1+s}) B(T+s)\end{aligned} \tag{20}$$

Using (17) and (18) in (20), we obtain $|X(T)| < \varepsilon$. Since $T \geq \tau$, and $n \geq \nu$ are arbitrary, we have in either case, $|x(t) - \hat{x}| = |X(t)| \to 0$ as $t \to \infty$. Using Proposition 1, we have $y(t) = 1 - a[x(t+1) - (1-d)x(t)]$ for all $t \in \mathbb{N}$, and so $y(t) \to 1 - a[\hat{x} - (1-d)\hat{x}] = \hat{x} = \hat{y}$ as $t \to \infty$. Using (2) above and the fact that $x(t) \to \hat{x}$ as $t \to \infty$, we note that the right hand side of (SVL) has a limit as $T \to \infty$, and so the left-hand side of (SVL) also has a limit as $T \to \infty$. Taking limits in (SVL), by letting $T \to \infty$, we obtain (US).

(4) As in the proof of (3) above, given any $\varepsilon > 0$, we can choose a positive integer $\tau \in \mathbb{N}$, such that (17) holds. Now, given any integer $T \geq \tau$, we sum (5) from $T$ to $T+s$, except that for $s$ odd, we use the equation in (5) with a negative sign. This yields:

$$\begin{aligned}-(\varepsilon/3) \leq -\left[\sum_{t=T}^{\infty} A(t) + \sum_{t=T}^{\infty} B(t)\right] &\leq X(T) + (-1)^s X(T+s+1) \\ \leq \left[\sum_{t=T}^{\infty} A(t) + \sum_{t=T}^{\infty} B(t)\right] &\leq (\varepsilon/3)\end{aligned} \tag{21}$$

Let $\tilde{x}$ be an arbitrary limit point of $\{x(t)\}$. Denote $(\tilde{x} - \hat{x})$ by $\tilde{X}$; then, there is a subsequence $\{t_r\}$, such that $\{X(t_r)\}$ converges to $\tilde{X}$. Pick $\mu$ such that for $r \geq \mu$, we have $t_r \geq \tau$, and $\tilde{X} - (\varepsilon/3) \leq X(t_r) \leq \tilde{X} + (\varepsilon/3)$. Thus, using (21), we have for $r \geq \mu$, and $s \in \mathbb{N}$,

$$\begin{aligned} -\tilde{X} - (2\varepsilon/3) \leq -X(t_r) - (\varepsilon/3) &\leq (-1)^s X(t_r + s + 1) \\ &\leq -X(t_r) + (\varepsilon/3) \leq -\tilde{X} + (2\varepsilon/3) \end{aligned}$$

Thus, for even positive integers $s$, we obtain:$-\tilde{X} - (2\varepsilon/3) \leq X(t_\mu + s + 1) \leq -\tilde{X} + (2\varepsilon/3)$, so that $X(t_\mu + s)$ converges to $(-\tilde{X})$ for all odd integers $s$. Also, for odd positive integers $s$, we obtain: $-\tilde{X} - (2\varepsilon/3) \leq -X(t_R + s + 1) \leq -\tilde{X} + (2\varepsilon/3)$. We can write this as: $\tilde{X} - (2\varepsilon/3) \leq X(t_R + s + 1) \leq \tilde{X} + (2\varepsilon/3)$, so that $X(t_\mu + s)$ converges to $\tilde{X}$ for all even integers $s$.

If $t_\mu$ is odd, we have $X(t)$ converging to $(-\tilde{X})$ for even integers and $X(t)$ converging to $\tilde{X}$ for odd integers. If $t_\mu$ is even, we have $X(t)$ converging to $(-\tilde{X})$ for odd integers and $X(t)$ converging to $\tilde{X}$ for even integers. □

*Proof of Lemma 1* (1) If $\{x(t), y(t)\}$ is a full-employment program, then $B(t) = 0$ for all $t \in \mathbb{N}$. Thus, we can use (5) to write:

$$[A(t)/\xi^t] = [X(t+1)/\xi^t] + \xi[X(t)/\xi^t] \quad for\ t \in \mathbb{N} \tag{22}$$

Given any positive integer $T \in \mathbb{N}$, we now sum (22) from $t = 0$ to $t = T$, except that for $t$ odd, we use the equation in (22) with a negative sign. This yields:

$$\sum_{t=0}^{T} [A(t)/(-\xi)^t] = \xi\{X(0) - [X(T+1)/(-\xi)^{T+1}]\} \tag{23}$$

Using the fact that $A(t) = [\alpha(t)/\hat{p}ad]$ for $t \in \mathbb{N}$, (23) yields (6).

(2) Using (3) of Proposition 1, if $\{x(t), y(t)\}$ is an optimal program, then it must be a full-employment program. Hence, it must satisfy (6), by (1) above. □

*Proof of Proposition 4* If $\{x(t), y(t)\}$ is an optimal program from $x$, and $\{x'(t), y'(t)\}$ is *any* program from $x$, then using (4), we can write for every positive integer $T \in \mathbb{N}$,

$$\sum_{t=0}^{T-1} [u(x'(t), x'(t+1)) - u(\hat{x}, \hat{x})] = \hat{p}x'(0) - \hat{p}x'(T) - \sum_{t=0}^{T-1} \delta'(t) \tag{24}$$

and:

$$\begin{aligned} &\sum_{t=0}^{T-1} [u(x'(t), x'(t+1)) - u(x(t), x(t+1))] \\ &\quad = \hat{p}x(T) - \hat{p}x'(T) - \sum_{t=0}^{T-1} \delta'(t) + \sum_{t=0}^{T-1} \delta(t) \end{aligned} \tag{25}$$

By Theorem 1 (1), there is a good program $\{\bar{x}(t), \bar{y}(t)\}$ from $x$, and $\sum_{t=0}^{\infty} \bar{\delta}(t) < \infty$. So, using this program in place of $\{x'(t), y'(t)\}$ in (25), together with the fact that $\{x(t), y(t)\}$ is an optimal program from $x$, we can infer that $\sum_{t=0}^{\infty} \delta(t) < \infty$, since $\bar{x}(t) \leq M(x)$ for $t \geq 0$. Now, using $\{x(t), y(t)\}$ in place of $\{x'(t), y'(t)\}$ in (24), we see that $\{x(t), y(t)\}$ must be good, since $x(t) \leq M(x)$ for $t \geq 0$. Now, (7) follows from Theorem 1 (3).

Since we have verified that $\{x(t), y(t)\}$ is good, we have $\sum_{t=0}^{\infty} \delta(t) < \infty$ by Theorem 1(2). If $\{x'(t), y'(t)\}$ is not good, then since $x'(t) \leq M(x)$ for $t \geq 0$, we can use (24) to infer that $\sum_{t=0}^{T-1} \delta'(t) \to \infty$ as $T \to \infty$, so that (VLM) is trivially satisfied. If $\{x'(t), y'(t)\}$ is good, then $\sum_{t=0}^{\infty} \delta'(t) < \infty$ by Theorem 1(2). Further, from Theorem 1(3), we have $\lim_{t\to\infty} \hat{p}x(t) = \hat{p}\hat{x}$, and $\lim_{t\to\infty} \hat{p}x'(t) = \hat{p}\hat{x}$. Thus, we see that the right hand side of (25) has a limit as $T \to \infty$. So, the left-hand side of (25) has a limit as $T \to \infty$, and this limit equals $[\sum_{t=0}^{\infty} \delta(t) - \sum_{t=0}^{\infty} \delta'(t)]$. Now, (VLM) follows from the optimality of $\{x(t), y(t)\}$. Finally, note that since $\{x(t), y(t)\}$ is an optimal program from $x$, it is a good program by (1), and so (US) follows from Theorem 1 (3). □

*Proof of Theorem 2* Let us note first that since $\xi = 1$, we have $\hat{x} = 1/2a$, $\hat{p} = 1/2$, and $(1/a) = 1 + (1 - d) > 1$, so that $a \in (0, 1)$ and $ad \in (0, 1)$.

(1) Since $\{x(t), y(t)\}$ is a full-employment program from $x_o$, we have $\beta(t) = 0$ for $t \in \mathbb{N}$, and $\delta(t) = \alpha(t)$ for $t \in \mathbb{N}$. Since $x(t) = \hat{x}$ for $t \geq 1$, we also have $X(t) = 0$ and $\delta(t) = 0$ for $t \geq 1$. Thus, using Lemma 1 and $\xi = 1$, we obtain for all $T \geq 1$,

$$\sum_{t=0}^{T} \delta(t) = \delta(0) = \alpha(0) = \sum_{t=0}^{T} [\alpha(t)/(-\xi)^t] = ad\hat{p}X(0) \tag{26}$$

If $\{x(t), y(t)\}$ is not optimal, then there is $\theta > 0$ and $N \in \mathbb{N}$, and a program $\{x'(t), y'(t)\}$ from $x_o$ such that for all $T > N$,

$$\theta \leq \sum_{t=0}^{T-1} [u(x'(t), x'(t+1)) - u(x(t), x(t+1))] \tag{27}$$

By Proposition 1, $\{x'(t), y'(t)\}$ can be taken to be a full-employment program from $x_o$. So, we have $\beta'(t) = 0$ for $t \in \mathbb{N}$, and $\delta'(t) = \alpha'(t)$ for $t \in \mathbb{N}$. Thus, using Lemma 1 and $\xi = 1$, we obtain for all $T \geq 1$,

$$\begin{aligned} \sum_{t=0}^{T} \delta'(t) &\geq \sum_{t=0}^{T} [\alpha'(t)/(-1)^t] \\ &= ad\hat{p}X(0) - ad\hat{p}[X'(T+1)/(-1)^{T+1}] \end{aligned} \tag{28}$$

Using (26) and (28), we obtain:

$$\sum_{t=0}^{T-1} \delta(t) - \sum_{t=0}^{T-1} \delta'(t) \leq ad\hat{p}[X'(T)/(-1)^T] \tag{29}$$

Now, using (27) and (29) in (25), we get for all $T > N$,

$$-\hat{p}X'(T) + ad\hat{p}[X'(T)/(-1)^T] \geq \theta \tag{30}$$

Note that by (27), $\{x'(t), y'(t)\}$ must be good, since $\{x(t), y(t)\}$ is clearly good. Using (30), we see that when $T > N$ is odd, we must have $[-X'(T)] \geq \theta/\hat{p}(1 + ad) = (\theta/a)$. Thus, by Theorem 1 (4), for $T > N$ and even, $X'(T)$ must converge to a positive number as $T \to \infty$. But, for $T > N$ and even, we have $\hat{p}X'(T)[ad - 1] \geq \theta$ by (30), so that $X'(T) \leq -\theta/\hat{p}[1 - ad] < 0$ (since $ad \in (0, 1)$ ), a contradiction. Thus, $\{x(t), y(t)\}$ is optimal from $x_o$.

(2) It can be checked that the program $\{x'(t), y'(t)\}$ has $\alpha'(t) = \beta'(t) = 0$ for all $t \in \mathbb{N}$, so that it satisfies condition (VLM). Thus, by Proposition 3, it must be optimal.

(3) Since $h$ is convex valued, there is a continuum of optimal programs from $x_o$. □

*Proof of Proposition 5* The golden-rule program, defined by $(\hat{x}(t), \hat{y}(t)) = (\hat{x}, \hat{x})$ has zero value loss in each period, and is therefore optimal by Proposition 3. Thus, $V(\hat{x}) = 0$ by (10).

If $\{x(t), y(t)\}$ is an optimal program from $x$, and $x' > x$, then we can define: $x'(0) = x'$, $x'(t+1) = x'(t)+z(t+1)$ for $t \geq 0$, and $y'(t) = y(t)$ for $t \geq 0$, where $\{z(t+1)\}$ is the investment sequence associated with the program $\{x(t), y(t)\}$. Then, it is easy to check that $x'(t) > x(t)$ for all $t \geq 0$, and so $\{x'(t), y'(t)\}$ is a program from $x'$. Thus, $V$ is non-decreasing on $\mathbb{R}_+$.

A slight refinement of this argument shows that, in fact, $V$ is strictly increasing on $\mathbb{R}_+$. First, note that we must have $z(t + 1) > 0$ for some $t \geq 0$. Otherwise, if $z(t + 1) = 0$ for all $t \geq 0$, then $x(t) \to 0$ as $t \to \infty$, which would contradict the turnpike property of optimal programs. Let $T$ be the first period for which $z(T + 1) > 0$. If $x' > x$, we can define: $x'(0) = x'$, $x'(t + 1) = x'(t) + z(t + 1)$ for $t \neq T$, $x'(T+1) = x'(T)+z'(T+1)$, where $z'(T+1) = z(T+1)-\varepsilon$, where $0 < \varepsilon < z(T + 1)$, and $\varepsilon$ is sufficiently close to 0 so that $x'(T + 1) > x(T + 1)$. Finally, define $y'(t) = y(t)$ for $t \neq T$, and $y'(T) = y(T) + \varepsilon$. It is easy to check that $x'(t) > x(t)$ for all $t \geq 0$, and that $\{x'(t), y'(t)\}$ is a program from $x'$. Thus, $V$ is strictly increasing on $\mathbb{R}_+$.

Since $\Omega$ is convex, and $u$ is concave on $\Omega$, the value function is concave on $\mathbb{R}_+$. It follows that $V$ is continuous on $\mathbb{R}_{++}$.

(2) Let $\{x(t), y(t)\}$ be the optimal program from $x$. Then:

$$V(x) = [u(x, x(1)) - u(\hat{x}, \hat{x})] + \sum_{t=1}^{\infty} \left[u(x(t), x(t + 1)) - u(\hat{x}, \hat{x})\right]$$

Note that the sequence $\{x'(t), y'(t)\}$ defined by $(x'(t), y'(t)) = (x(t+1), y(t+1))$ for all $t \in \mathbb{N}$ defines a program from $x(1)$. Since $\{x(t), y(t)\}$ is an optimal program from $x$, it is a good program from $x$ (Proposition 4), and so $\{x'(t), y'(t)\}$ is a good program from $x(1)$.Thus, we have:

$$\sum_{t=1}^{\infty} \left[u(x(t), x(t + 1))-u(\hat{x}, \hat{x})\right]=\sum_{t=0}^{\infty} \left[u(x'(t), x'(t + 1))-u(\hat{x}, \hat{x})\right] \leq V(x(1))$$

the inequality above following from (10). This yields:

$$V(x) \leq [u(x, x(1)) - u(\hat{x}, \hat{x})] + V(x(1)) \tag{31}$$

On the other hand, if $(x, x') \in \Omega$, then by using an optimal program $\{x'(t), y'(t)\}$ from $x'$, we can define a program $\{x(t), y(t)\}$ from $x$ by: $(x(0), y(0)) = (x, \max \Lambda(x, x'))$, and $(x(t), y(t)) = (x'(t - 1), y'(t - 1))$ for $t \geq 1$.

Since the program $\{x'(t), y'(t)\}$ from $x'$ is good (by Proposition 4), so is the program $\{x(t), y(t)\}$ from $x$. Thus, by (9) and (10), we obtain:

$$\begin{aligned} V(x) &\geq [u(x, x') - u(\hat{x}, \hat{x})] + \sum_{t=1}^{\infty}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] \\ &= [u(x, x') - u(\hat{x}, \hat{x})] + \sum_{t=1}^{\infty}[u(x'(t-1), x'(t)) - u(\hat{x}, \hat{x})] \\ &= [u(x, x') - u(\hat{x}, \hat{x})] + \sum_{t=0}^{\infty}[u(x'(t), x'(t+1)) - u(\hat{x}, \hat{x})] \\ &= [u(x, x') - u(\hat{x}, \hat{x})] + V(x') \end{aligned} \tag{32}$$

Thus, if $\{x(t), y(t)\}$ is an optimal program from $x$, equality must hold in (31). There are two implications of this finding. First, we must have:

$$V(x(1)) = \sum_{t=1}^{\infty}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})]$$

and therefore [by (9)] the sequence $\{x'(t), y'(t)\}$ defined by $(x'(t), y'(t)) = (x(t+1), y(t+1))$ for all $t \in \mathbb{N}$ is an optimal program from $x(1)$. Thus, we must have for each $t \geq 0$ :

$$V(x(t)) = [u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] + V(x(t+1)) \tag{33}$$

Second, $V$ satisfies the following *functional equation of dynamic programming*:

$$V(x) = \max_{(x,x')\in\Omega} \{[u(x, x') - u(\hat{x}, \hat{x})] + V(x')\}$$

(3) We have already established one half of (iii) in (33) above. To establish the other half, let $\{x(t), y(t)\}$ be any program which satisfies for each $t \in \mathbb{N}$ :

$$V(x(t)) = [u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] + V(x(t+1))$$

Then, for every $T \in \mathbb{N}$, we have:

$$V(x(0)) = \sum\nolimits_{t=0}^{T}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] + V(x(T+1)) \tag{34}$$

Since $x(t) \leq M(x(0))$ for all $t \in \mathbb{N}$ and $V$ is increasing on $X$, we can define $m = V(M(x(0)))$, and infer that $V(x(t)) \leq V(M(x(0))) \equiv m$ for all $t \in \mathbb{N}$. On using this information in (34), we obtain:

$$\sum\nolimits_{t=0}^{T}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] = V(x(0)) - V(x(T+1)) \geq V(x(0)) - m$$

allowing us to conclude that the program $\{x(t), y(t)\}$ must be good, and that therefore $x(t) \to \hat{x}$ as $t \to \infty$. Thus, by the continuity of $V$ on $\mathbb{R}_{++}$, we obtain $V(x(t)) \to V(\hat{x}) = 0$ as $t \to \infty$. Using this information in (34), we see that $\{\lim_{T\to\infty} \sum_{t=0}^{T}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})]\}$ exists and is finite, and:

$$V(x) = \sum\nolimits_{t=0}^{\infty}[u(x(t), x(t+1)) - u(\hat{x}, \hat{x})]$$

Thus, by (8),(9) and (US), $\delta(x(0)) = \sum_{t=0}^{\infty} \delta(t)$, so that (by Proposition 3) $\{x(t), y(t)\}$ is an optimal program from $x(0)$. □

*Proof of Corollary 1* First, note that by (9), we have $V(x) - V(\hat{x}) \leq \hat{p}(x - \hat{x})$ for all $x > \hat{x}$, so that by concavity of $V$, we obtain $V'_+(\hat{x}) \leq \hat{p}$. Further, by concavity of $V$, we have $V'_+(x) \leq V'_+(\hat{x}) \leq \hat{p}$ for all $x > \hat{x}$.

We consider, first, the case in which $x \in C$. Then, for $(x, x') \in \Omega$, it must be the case that $x' \geq (1-d)x \equiv \bar{x}$. Suppose, contrary to the claim of the Corollary, there is $x' > \bar{x}$, with $x' \in h(x)$. Then, we have $x' > \hat{x}$, and so $V(x') - V(\bar{x}) \leq V'_+(\bar{x})(x' - \bar{x}) \leq \hat{p}(x' - \bar{x}) < a(x' - \bar{x})$. Using this, we can write:

$$\begin{aligned} &u(x, x') + V(x') \\ &\quad = \{1 - a[x' - (1-d)x]\} + [V(x') - V(\bar{x})] + V(\bar{x}) \\ &\quad < 1 - a(x' - \bar{x}) + a(x' - \bar{x}) + V(\bar{x}) \\ &\quad = u(x, \bar{x}) + V(\bar{x}) \leq V(x) + u(\hat{x}, \hat{x}) \end{aligned} \tag{35}$$

the last inequality in (35) following from Proposition 5(3). But, this contradicts Proposition 5(2).

Next, consider the case in which $x \in A$. Denote $[(1/a) - \xi x]$ by $\bar{x}$; note that $\bar{x} \geq \hat{x}$. Suppose, contrary to the corollary, that there is some $x' \in h(x)$ with $x' \neq \bar{x}$. If $x' < \bar{x}$, then for $y \in \Lambda(x, x')$, we have $a[x' - (1-d)x] + y < a[\bar{x} - (1-d)x] + x = 1$, so labor is not fully employed, a contradiction to Lemma 1(2). If $x' > \bar{x}$, then $x' > \bar{x} \geq \hat{x}$, and so, we get $V(x') - V(\bar{x}) \leq V'_+(\bar{x})(x' - \bar{x}) \leq \hat{p}(x' - \bar{x})$. Using this, we can write:

$$\begin{aligned} &u(x, x') + V(x') \\ &\quad = \{1 - a[x' - (1-d)x]\} + [V(x') - V(\bar{x})] + V(\bar{x}) \\ &\quad \leq \{1 - a[\bar{x} - (1-d)x]\} - a(x' - \bar{x}) + \hat{p}(x' - \bar{x}) + V(\bar{x}) \\ &\quad < x + V(\bar{x}) \\ &\quad = u(x, \bar{x}) + V(\bar{x}) \leq V(x) + u(\hat{x}, \hat{x}) \end{aligned} \tag{36}$$

the last inequality in (36) following from Proposition 5(3). But, this contradicts Proposition 5(2). □

*Proof of Corollary 2* (1) Define $\{\bar{x}(t)\}$ by $\bar{x}(0) = x$, and $\bar{x}(t) = \hat{x}$ for $t \geq 0$. Then, noting that $(\bar{x}(t), \bar{x}(t+1)) \in \Omega$ for all $t \in \mathbb{N}$, we can define $\{\bar{y}(t)\}$ by $\bar{y}(t) = u(\bar{x}(t), \bar{x}(t+1))$ for all $t \in \mathbb{N}$. Then, $\{\bar{x}(t), \bar{y}(t)\}$ is a full-employment program from $x$. Thus, we have $\bar{\beta}(t) = 0$ for $t \in \mathbb{N}$, and $\bar{\delta}(t) = \bar{\alpha}(t)$ for $t \in \mathbb{N}$. Since $\bar{x}(t) = \hat{x}$ for $t \geq 1$, we also have $\bar{X}(t) = 0$ and $\bar{\delta}(t) = 0$ for $t \geq 1$. Thus, using Lemma 1, we obtain:

$$\sum_{t=0}^{\infty} \bar{\delta}(t) = \bar{\delta}(0) = \bar{\alpha}(0) = \sum_{t=0}^{\infty} [\bar{\alpha}(t)/(-\xi)^t] = ad\xi\hat{p}\bar{X}(0) \tag{37}$$

Now, let $\{x'(t), y'(t)\}$ be an optimal program from $x$. [By Proposition 3, an optimal program exists.] Then, $\{x'(t), y'(t)\}$ is a full-employment program from $x$, and using Lemma 1 and $\xi > 1$, we obtain:

$$\begin{aligned} \sum_{t=0}^{\infty} \delta'(t) &\geq \sum_{t=0}^{\infty} \alpha'(t) \geq \sum_{t=0}^{\infty} (\alpha'(t)/\xi^t) \geq \sum_{t=0}^{\infty} [\alpha'(t)/(-\xi)^t] \\ &= ad\xi\hat{p}X(0) \end{aligned} \tag{38}$$

By Proposition 4, we have:

$$\sum_{t=0}^{\infty} \delta'(t) \leq \sum_{t=0}^{\infty} \delta(t) \tag{39}$$

for every program $\{x(t), y(t)\}$ from $x$. But, by (37) and (38), we also have $\sum_{t=0}^{\infty} \delta'(t) \geq \sum_{t=0}^{\infty} \bar{\delta}(t)$, so that we must have $\sum_{t=0}^{\infty} \delta'(t) = \sum_{t=0}^{\infty} \bar{\delta}(t)$, and using this in (39), we get:

$$\sum_{t=0}^{\infty} \bar{\delta}(t) \leq \sum_{t=0}^{\infty} \delta(t)$$

for every program $\{x(t), y(t)\}$ from $x$. Thus, by Proposition 3, $\{\bar{x}(t), \bar{y}(t)\}$ is optimal from $x$, and so $\hat{x} \in h(x)$. Now, note that since $\sum_{t=0}^{\infty} \delta'(t) = \sum_{t=0}^{\infty} \bar{\delta}(t)$, we must have equality in all the inequalities of (38). This, in turn, means that $\alpha'(t) = 0$ for all $t \geq 1$, and so $X'(t+1) = (-\xi)X'(t)$ for $t \geq 1$, by using (5). Since $x'(t) \leq M(x)$ for all $t \in \mathbb{N}$, and $\xi > 1$, we must therefore have $X'(1) = 0$, that is, $x'(1) = \hat{x}$. Thus, $\hat{x} = h(x)$.

(2) First, consider the case where $x \in (\hat{x}, 1)$. Define a sequence $\{x(t)\}$ by: $x(0) = x$, and $x(t+1) = (-\xi)(x(t) - \hat{x}) + \hat{x}$ for all $t \in \mathbb{N}$. Then, it is easy to check that $(x(t), x(t+1)) \in \Omega$ for all $t \in \mathbb{N}$, and so we can define $\{y(t)\}$ by $y(t) = u(x(t), x(t+1))$ for all $t \in \mathbb{N}$. Then, $\{x(t), y(t)\}$ is a full-employment program from $x$, with $\delta(t) = 0$ for $t \in \mathbb{N}$, by using (5). Thus, by Proposition 3, $\{x(t), y(t)\}$ is optimal from $x$, and so $x(1) = (1/a) - \xi x \in h(x)$. By Proposition 4, if $\{x'(t), y'(t)\}$ is optimal from $x$, then $\delta'(t) = 0$ for all $t \in \mathbb{N}$, and so by (5), $x'(1) = (-\xi)(x - \hat{x}) + \hat{x} = (1/a) - \xi x$. Thus, $h(x) = (1/a) - \xi x$.

Next, consider the case where $x \in [1, k)$. Define a sequence $\{x(t)\}$ by: $x(0) = x$, $x(1) = (1-d)x$ and $x(t+1) = (-\xi)(x(t) - \hat{x}) + \hat{x}$ for all $t \geq 1$. Then, it is easy to check that $(x(t), x(t+1)) \in \Omega$ for all $t \in \mathbb{N}$, and so we can define $\{y(t)\}$ by $y(t) = u(x(t), x(t+1))$ for all $t \in \mathbb{N}$. Then, $\{x(t), y(t)\}$ is a full-employment program from $x$, with $y(0) = 1$, and $\delta(t) = 0$ for $t \geq 1$, by using (5). Further, if $\{x'(t), y'(t)\}$ is any program from $x$, then we have $x'(1) \geq (1-d)x$, and so $y'(1) \leq 1 = y(1)$, so that $\delta'(0) \geq \alpha'(0) \geq \alpha(0) = \delta(0)$. Thus, by Proposition 3, $\{x(t), y(t)\}$ is optimal from $x$, and so $x(1) = (1-d)x \in h(x)$. By Proposition 4, if $\{x'(t), y'(t)\}$ is optimal from $x$, then $\delta'(0) = \delta(0)$ and $\delta'(t) = 0$ for all $t \geq 1$. Thus, $x'(1) = x(1) = (1-d)x$, and $h(x) = (1-d)x$. □

*Proof of Corollary 3* First, consider $x \in (\hat{x}, 1)$. Then, the result that $[(1/a) - \xi x, \hat{x}] \subset h(x)$ can be obtained by following the proof of Theorem 2. Next, consider $x \in [1, k)$. Then $(1-d)x \in h(x)$ by following the proof of Corollary 2(2). Also, $\hat{x} \in h(x)$ by following the proof of Theorem 2(1), noting that the argument used there applies to all $x \in (\hat{x}, k)$. Since $h$ is convex valued, $[(1-d)x, \hat{x}] \subset h(x)$.

Next, consider $x \in C$. Let $T \in \mathbb{N}$ be the smallest integer such that $\bar{x} \equiv (1-d)^T x < 1$; then $T \geq 1$, and $(1-d)^T x \geq (1-d)$. Define a sequence $\{x(t)\}$ as follows: $x(t) = (1-d)^t x$ for $t = 0, ..., T$, and $x(t+1) = 2\hat{x} - x(t)$ for $t \geq T$. Then, it is easy to check that $(x(t), x(t+1)) \in \Omega$ for all $t \in \mathbb{N}$, and so we can define $\{y(t)\}$ by $y(t) = u(x(t), x(t+1))$ for all $t \in \mathbb{N}$. Then, $\{x(t), y(t)\}$ is a full-employment program from $x$, with $\delta(t) = 0$ for $t \geq T$. Further, if $\{x'(t), y'(t)\}$ is any program from $x$, then we have $x'(t) \geq x(t)$ and $y'(t) \leq 1 = y(t)$ for $t = 0, ..., T-1$, so that $\delta'(t) \geq \alpha'(t) \geq \alpha(t) = \delta(t)$ for $t = 0, ..., T-1$.Thus, by Proposition 3, $\{x(t), y(t)\}$ is optimal from $x$, and so $x(1) = (1-d)x \in h(x)$.

Finally, consider $x \in A$. Denote $(1/a) - \xi x$ by $\bar{x}$; then $\bar{x} \geq \hat{x}$. If $\bar{x} \in (\hat{x}, 1)$, define the program $\{x(t), y(t)\}$ as in the proof of Corollary 2(2), dealing with $x \in (\hat{x}, 1)$. If $\bar{x} \in [1, k)$, define the program $\{x(t), y(t)\}$ as in the proof of Corollary 2(2), dealing with $x \in [1, k)$. If $\bar{x} \in C$, define the program $\{x(t), y(t)\}$ as in the above paragraph. Then, in each case, consider the program $\{x'(t), y'(t)\}$ defined by $(x'(0), y'(0)) = (x, x)$ and $(x'(t), y'(t)) = (x(t-1), y(t-1))$ for $t \geq 1$. It can be checked that in each case this defines an optimal program. Thus, $(1/a) - \xi x \in h(x)$ in each case. □

## References


Brock, W.A.: On the existence of weakly maximal programmes in a multi-sector economy. Rev Econ Stud **37**, 275–280 (1970)

Brock, W.A., Majumdar, M.: On characterizing optimal competitive programs in terms of decentralizable conditions. J Econ Theory **45**, 225–243 (1988)

Dechert, W.D., Nishimura, K.: A complete characterization of optimal growth paths in an aggregated model with a non-concave production function. J Econ Theory **31**, 332–354 (1983)

Gale, D.: On optimal development in a multi-sector economy. Rev Econ Studies **34**, 1–18 (1967)

Khan, M.A., Mitra, T.: Optimal growth under irreversible investment: a continuous time analysis. Mimeo, Cornell University (2003)

Khan, M.A., Mitra, T.: On choice of technique in the Robinson-Solow-Srinivasan model. Int J Econ Theory **1**, 83–110 (2005a)

Khan, M.A., Mitra, T.: Optimal growth in a two-sector model without discounting: a geometric investigation. Japanese Econ Rev (forthcoming) (2005b)

Kydland, F.E., Prescott, E.C.: Time to build and aggregate fluctuations. Econometrica **50**, 1345–1370 (1982)

Majumdar, M.K., Nermuth, M.: Dynamic optimization in non-convex models with irreversible investment: monotonicity and turnpike results. Z Nationalökonomie **42** 339–362 (1982)

McKenzie, L.W.: Accumulation programs of maximum utility and the von Neumann facet. In: Wolfe, J.N. (ed.) Value, capital and growth, pp. 353–383. Edinburgh: Edinburgh University Press 1968

McKenzie, L.W.: Classical general equilibrium theory. Cambridge: MIT Press 2002

Mitra, T., Wan, H. Jr.: On the Faustmann solution to the forest management problem. J Econ Theory **40**, 229–249 (1986)

Nishimura, K., Yano, M.: Durable capital and chaos in competitive business cycles. J Econ Behav Organ **27**, 165–181 (1995)

Radner, R.: Paths of economic growth that are optimal only with respect to final states. Rev Econ Stud **28**, 98–104 (1961)

Srinivasan, T.N.: Optimal savings in a two-sector model of growth. Econometrica **32**, 358–373 (1964)

Stiglitz, J.E.: A note on technical choice under full employment in a socialist Economy. Econ J **78**, 603–609 (1968)

Uzawa, H.: Optimal growth in a two-sector model of capital accumulation. Rev Econ Stud **31**, 1–24 (1964)

Weitzman, M. L.: Duality theory for infinite horizon convex problems. Manag Sci **19**, 783–789 (1973)