Classical Turnpike Theory and the Economics of Forestry

M. Ali Khan and Adriana Piazzaz, School of Economics Discussion Paper No. 397, August 2009, School of Economics, The University of Queensland. Australia.

Abstract

Classical turnpike theory, as originally conceived by Samuelson, pertains to optimal growth theory over a large but infinite time horizon with given initial and terminal stocks. In this paper, we present two turnpike results in the context of the economics of forestry with given initial and terminal forest configurations. Our results depart from the general theory in that they pertain to a transitional production set which does not satisfy the assumptions of inaction and free disposal, and rely on a recently-discovered non-interiority assumption on concave (not necessarily differentiable) benefit functions that implies, and is implied by, the asymptotic convergence of good programs.

EPrint Type:	Departmental Technical Report
Keywords:	Turnpike theory, asymptotic convergence, good program, approximately optimal program, time horizon, large but infinite, non-interiority, non-differentiability, forest management.
Subjects:	
ID Code:	JEL Classification: C62, D90, Q23
Deposited By:	

Classical Turnpike Theory and the Economics of Forestry*

M. Ali Khan†
Adriana Piazza‡

20 July, 2009

Abstract

Classical turnpike theory, as originally conceived by Samuelson, pertains to optimal growth theory over a large but finite time horizon with given initial and terminal stocks. In this paper, we present two turnpike results in the context of the economics of forestry with given initial and terminal forest configurations. Our results depart from the general theory in that they pertain to a transitional production set which does not satisfy the assumptions of *inaction* and *free disposal,* and rely on a recently-discovered non-interiority assumption on concave (not necessarily differentiable) benefit functions that implies, and is implied by, the asymptotic convergence of good programs. (102 words)

Keywords Turnpike theory, asymptotic convergence, good program, approximately optimal program, time horizon, large but finite, non-interiority, non-differentiability, forest management.

JEL Classification Numbers C62, D90, Q23

*This work was initiated during Piazza's visit to Johns Hopkins in May 2008, continued during her visit to the University of Illinois at Urbana-Champaign in April-May, 2009, and completed when Khan held the position of Visiting Professor at the University of Queensland, June-July 2009, and at the Nanyang Technological University, August, 2009. The authors thank Joseph Harrington, Flavio Menezes, Euston Quah, Anne Villamil and Nicholas Yannelis, and each department, for their hospitality. They are also grateful to Tapan Mitra for correspondence and conversation, and in particular for his result indicated in Footnote 18 below. Adriana Piazza gratefully acknowledges the financial support of FONDECYT under project 3080059.

†Department of Economics, The Johns Hopkins University, Baltimore, MD 21218 and Department of Economics, University of Queensland, Brisbane, QLD 4072. **E-mail** akhan@jhu.edu

‡Centro de Modelamiento Matemático, Universidad de Chile. Blanco Encalada 2120, 7 piso. Santiago, Chile. **E-mail** apiazza@dim.uchile.cl

The [turnpike] results while simple and concise could probably not be predicted in advance.

(Gale 1970)

1 Introduction

The results presented in this paper can perhaps be introduced best by asking what is "classical turnpike theory"? and what has it to do with the "economics of forestry"? We begin this introduction with the first question.

The origin of the subject is easily dated to Paul Samuelson's 1949 Rand Memorandum, though a more public presentation had to await the 1958 publication of DOSSO and the culminating summary description in the 1971 Nobel lecture.[1] An essential motivation was to show the relevance of classical methods of optimization, the calculus of variations in particular, to the (then) more recent minimax and fixed-point mathematics of von-Neumann's.[2] In a less technical and more substantive register, it was to show that the solutions of an intertemporal allocation problem over a large but finite time horizon, with given initial and final terminal stocks, stayed approximately close, most of the time, to maximal balanced growth programs that are divorced from any considerations pertaining to a terminal horizon. In other words, it was to give a normative microfoundation, one based on optimization over a large but finite time interval, to a descriptive solution represented by a saddlepoint, one that paired a maximal balanced growth rate of commodities with an identical minimal balanced growth rate of the associated prices. To revert to the technical, it was to show that the Euler-Lagrange differential or difference equations delineating the solution of a finite-horizon variational program, when linearized around the maximally balanced growth program, furnished real roots that came in opposite signed pairs.[3] In sum, turnpike theory, as conceived by Samuelson, went beyond the characterization of necessary and sufficient conditions in the calculus of variations to an investigation of the qualitative properties of the solutions satisfying such conditions for large but finite time intervals.[4]

In the first decade (1958-1968) that followed the publication of DOSSO, the initial contributions to the theory all retained this emphasis on large but finite programs: the literature simply broadened its scope from the relatively narrow confines of the von-Neumann's model, at least from an economic point of view, to those of Ramsey's where time paths of consumption, rather than the utility from terminal capital stocks, were given prominence and primacy.[5] Other than this, the essential qualitative outlines of the problematic, and the forging of the theory to address it, remained unchanged. What did change by the end of the decade, however, was the acceptance of an infinite-horizon variational problem as a legitimate vehicle to address issues of intertemporal allocation, something already committed to in Ramsey (1928),[6] but with the von Weizsäcker-Atsumi 1965 circumvention of his methods to deal with the possible non-convergence of the objective. Gale (1967), and the symposium in which his article appeared, and Cass (1965) and Koopmans (1965) in Ramsey's single-stock aggregative setting,[7] are perhaps the most evident benchmarks of this acceptance and the changed emphasis.[8] Gale focusses on the problem of the existence of an optimal

[1]For the first, see Figure 14 in [23], also referred to in [13] with the precise date of June 29, 1949. DOSSO is now a time-honored acronym for [2]. It is of interest that there is no reference to [23] in [2], or to the word *turnpike* in its index. The 1956 publication [31] also touches on the topic and already has references to a forthcoming DOSSO. For the second, see the seven paragraphs in the section titled "Dynamics and Maximizing" in the Nobel lecture [26].

[2]See [26] and the reference to the von-Neumann Samuelson exchange in [13, p. 2]. This episode is also focussed on in [10, Footnote 8] and the accompanying text which has a discussion of how this exchange seems to have colored the subsequent development of the subject.

[3]The undiscounted setting deserves emphasis; see [24, Section 8], the subsequent [25], and the minimax interpretations of the Ramsey "bliss points" in [31]. This point receives extended elaboration, also for the discounted setting, in a series of papers in the volumes [27, 30].

[4]Chapter 223 in [30] is an important one in this connection.

[5]For the attendant literature, one can hardly do better than to refer to [29, Footnote 1].

[6]Also pursued in a multi-sectoral context in a somewhat isolated and in a recently neglected contribution of Samuelson-Solow (1956).

[7]It is therefore perhaps fitting that modern macroeconomics, in its neglect of the undiscounted multi-sectoral setting, references only Ramsey, Cass and Koopmans and intensively studies only the so-called RCK model.

[8]It is also of some interest for historians of economic analysis that of Gale's six references, Ramsey, Samuelson-Solow, Koopmans and von Weizsäcker (in that order) are the only substantive ones. See Chapters on the Turnpike in Hicks (1965) for

program, and resolves it by conditions ensuring the asymptotic convergence of what he terms *good* programs, a methodological procedure that he refers to as a "round-about method."[9] This asymptotic property of good programs is referred to, in passing, as a "turn-pike" theorem.[10]

By the end of the subsequent decade, Lionel McKenzie (1976) had incorporated the Gale-Koopmans-Cass asymptotics into turnpike theory, and substantially broadened its scope to include within it all intertemporal allocation problems. By distinguishing three types of turnpikes, he shifted the emphasis from large but finite programs to allocation problems involving an undetermined (infinite) time horizon. Classical turnpike theory, as conceived by Samuelson, became the *middle,* as opposed to the *early* and *late* turnpikes. Yet another decade later, in his introduction to the subject, McKenzie (1986; p. 1282) was to write:

> The theory that I will present will cover both discounted and undiscounted utility. We will seek to determine the asymptotic behavior of maximal paths, which display a tendency to cluster in the sufficiently distant future from whatever capital stocks they start. Other types of turnpike behavior that have been studied are clustering in early periods for finite optimal paths that start from the same initial stocks, but have different terminal stocks, and clustering in the middle parts of paths that may start and end with different stocks.

With this linguistic turn, the classical conception becoming the "other", it was inevitable that the late turnpike became *the* turnpike to be investigated. There is no middle turnpike theorem in McKenzie's Handbook chapter,[11] and already in 1982, he refers to a "turnpike result" as being synonymous to a result on "asymptotic convergence."[12] It is therefore only fitting that in his state-of-the-art survey of the aggregative model, Mitra (2005) writes

> It should be clarified that it is only in the sense of "global asymptotic stability" that the term "turnpike property" is used in this paper.

Ironically, in his expository paper, Gale (1970) reverts to the classical conception, and along with monotonicity properties of optimal programs, situates classical turnpike theory within the broader rubric of qualitative properties of optimal programs.[13] Thus what is at heart a semantic issue can nevertheless be a source of some confusion at best, and a facile lack of analytical discrimination at worst.[14] At any rate, the point that is being made in this introductory description of the evolution of the subject is simply that in this paper we revert to Samuelson's classical conception pertaining to large but finite optimal programs, and investigate McKenzie's middle turnpike.

The resurgence of interest in the economics of forestry can also be dated originally to Samuelson (1976) and to his sighting of Faustman's (1849) analysis. However, it remained for the remarkable articles of Mitra-Wan (1985, 1986) to take his market analysis and recast it into an optimal planning, Ramseyian framework. Once this link is forged, and the conceptual markers of the theory of intertemporal resource allocation are identified in the forestry model – the golden-rule stocks and prices and corresponding maximal sustainable timber yields as Ramsey's bliss point – we are naturally lead to ask whether forest management programs for large but finite time horizons follow a turnpike. To go to the epigraph from Gale, there appears to be no reason why they should, and this scepticism leads directly into our second question concerning the

---

this turn.

[9]Gale [3, p.1] writes, "It may well be that there is a more direct way of obtaining our existence theorem, but even if this should turn out to be true, the present round-about approach would not be an entirely wasted effort." The continuing relevance of this "round-about" method is noted, and utilized, in [10, 11].

[10]The phrase is used twice: in the third paragraph of the introduction, and to introduce Section 9. Given that the concept of *good* programs originates in Gale, his reference to the concept as being "central to much of the recent literature on dynamic production models" is initially somewhat puzzling, but insightful on reflection.

[11]Other than Theorem 2 for the von-Neumann production model, the same is true of in Chapter 7 of McKenzie's 2002 text; see Theorems 4 and 6.

[12]The first sentence in [12] reads "In virtually all proofs of turnpike results, that is, proofs of theorems on the asymptotic convergence of optimal paths of capital accumulation, duality considerations play a critical role." Also see the meaning given to the term in Bewley (1982).

[13]See paragraph 3 and Section 5 in [4]. (The reader is warned that there are two Sections 4 in the paper.)

[14]Such a lack of discrimination revolves around the generality of turnpike theory, a point to which we return below.

relevance of classical turnpike theory to the Mitra-Wan (MW) forestry model. Here the usual justifications of turnpike theory as resolving qualitative puzzles and computational difficulties take on an added force. At least since Faustman, periodicity seems to be the rule in the economics of forestry, and the charting of an optimal policy correspondence with a non-linear benefit (felicity) function remains, even until now, totally uncharted.[15] Furthermore, a forest, unlike a given stock of capital, is much more than a durable input for the production of desired commodities – it is desirable in itself, and if not a "way of life" of so-called endogenous and native communities, a stock imbricated by externality considerations and entrusted by one generation to another.[16] As such, a forest configuration also enters as an argument in the benefit (felicity) function, and thereby further complicates the difficulties of determining what the planner has to do "tomorrow and the day after" rather than the long-run. This lack of knowledge of the optimal policy correspondence and of transition dynamics leads only to an increased reliance on turnpike results.[17] But more to the point, such results allay fears and furnish a reassurance that a planner's departure from an initial forest configuration to one yielding maximal sustained timber yields, even when he has to return to future generations the forest in the same state that he was given it, is not betraying this trust. Thus, turnpike theory attains a normative significance perhaps even greater than that in capital theory in the abstract and in the large, and that the economics of forestry can only gain when supplemented by a suitably formulated turnpike theory.

With the relevance of turnpike theory to the economics of forestry established and out of the way, one can turn to the more immediately antecedent literature and delineate the precise contribution of this paper. Given the extensive work on turnpike theory associated with the names of Samuelson, McKenzie, Gale and their followers, why can one not simply appeal to the standard results? What is the need for additional modifications? In particular, why are the results presented in this paper not straightforward applications of the recent extension of the theory sketched in Khan-Zaslavski [10]? The answers to these questions require a technically more focussed discussion. We turn to this.

Even though the RSS model and the MW model are different models with entirely different interpretive registers, the subtle analytical connections between them are undeniable, and the recent RSS revisitation of turnpike theory in [11] worth noting. It involves at least four disparate elements: (i) the irrelevance, in principle, of necessary, first order Euler-Lagrange conditions, and indeed of differentiability of the felicity function at the golden-rule stock, (ii) the identification of asymptotic convergence of good programs as a sufficient condition for classical turnpike theory, and therefore for the asymptotic convergence of optimal programs, (iii) the derivation of asymptotic stability of optimal programs from the classical turnpike result, which is to say, the derivation of results on the early and late turnpikes as a *consequence* of a result on the middle turnpike, (iv) a focus on approximately optimal large but finite programs. Only points (i) and (ii) need further supplementation in the context of results that we report here, and we take them in turn. As regards (i), it is now well-understood that the golden-rule stock in the RSS model, is not in the interior of the transition set, and even for the case with a single type of machine when it is in the interior, the reduced form utility function is not differentiable at it even with a linear felicity function. In part, this is precisely what gives the RSS model its continuing interest. The same issue occurs in the MW model. As regards (ii), it allows us to move away from the dichotomy of linear and strictly concave felicity functions to a more productive sufficiency condition, something essential for the RSS model where even strictly concave felicity functions do *not* lead to strictly concave reduced-form utility functions as is required by the theory. In particular, such a condition allows a turnpike theorem when the felicity function is linear and the marginal rate of transformation $\xi_\sigma \neq 1$. It is this issue that finds its most satisfactory culmination in the MW model.

In recent work, Khan-Piazza [9] furnish for the MW forestry model a non-interiority condition that is necessary *and* sufficient for asymptotic convergence of good programs when the benefit (felicity) function is assumed only to be concave and not necessarily differentiable.[18] And so the natural question arises as to

[15]Mitra (2006) is a notable exception for the strictly concave case, but even his 22-page analysis is limited to a dual-aged forest and requires additional assumptions.

[16]For externality considerations in the economics of forestry, see [28]; and for the larger implications reaching into political theory, see Kant-Berry (2005), and perhaps also the discussion of the references in [7].

[17]Also see the last paragraph of Gale's (1967) introduction. Its somewhat "paradoxical" defense of infinite-horizon problems as pertaining to the very immediate future – the "guidance of a ship on a long journey" – is premised on precisely the ability to compute this policy correspondence.

[18]Subsequent to obtaining the results presented here, Tapan Mitra provided a transparent equivalent formulation of the

whether one can construct a robust turnpike theory of the classical type for the economics of forestry, as is done in [10] for the (RSS) choice of technique problem in development planning. The results presented in Section 3 below answer this question in the affirmative, and constitute the principal results of this work. They also yield as direct corollaries results on the asymptotic convergence of optimal programs. Unlike the RSS model, there is no natural ordering on the transition production set in the MW model, and this necessitates novel and different arguments, and as it happens, more constructive ones than those presented in [10]. Sections 4 and 5 present the statements of several results that are both, needed in the proofs of the results of Section 3, and interesting on their own right. We comment especially aspects in which they depart from the corresponding arguments in [10]; the formalities of the proofs themselves are confined to the Appendix. Section 6 concludes the paper with a delineation of three open problems. In the next section, by way of introducing the reader to the notation and the terminology, we present the basic analytics of the model. This material is by now well-understood, but for the sake of completeness, we present results that being phrased in terms of our non-interiority condition, generalize corresponding results in [8].

2 The Mitra-Wan Tree Farm and the Non-Interiority Condition

We begin by introducing some notation. Let $\mathbb{N}$ be the set of non-negative integers and $\mathbb{R}$ ($\mathbb{R}_+$) the set of real (non-negative) numbers. We shall work in the $n-1$-dimensional simplex $\Delta = \{x \in \mathbb{R}_+ / \sum_{i=1}^n x_i = 1\}$. For any $x, y \in \mathbb{R}^n$ we denote the inner product by $xy = \sum_{i=1}^n x_i y_i$ and the supreme norm of x by $\|x\|_\infty$.

In addition to its original formulation [14, 15], an outline of the Mitra-Wan forestry model is also available in [17], and of the special dual-aged case, in [16, 18]. Here we depart from the original specification and work with the reformulation presented in [21] and pursued in [7, 8, 22]. Under this specification, the model consists simply of the pair (b, w), where b is a non-negative vector of biomass coefficients $(b_1, \ldots, b_n)$, and $w : [0, \infty) \to \mathbb{R}$ the benefit (felicity) function of timber yields. A forest (farm) configuration is an element of Δ, representing the fact that trees of ages ranging from one to n cover completely a homogeneous plot of land of normalized unit size.

Note that we do not use their timber-content function $f(\cdot)$, and make no assumptions on the biomass coefficients other than the following Brock-Mitra-Wan uniqueness condition.

Standing Hypothesis (BMW): There exists $\sigma \in \{1, \ldots, n\}$ such that $(b_\sigma/\sigma > b_i/i)$ for all $i \in \{1, \ldots, n\} \setminus \{\sigma\}$.

In addition to this, we very much follow the original conception and assume that there are no costs of plantation, that the timber content per unit of area is related only to the age of the trees, and that n is the age after which a tree dies or losses its economic value. However, one difference should be noted. In their treatment, Mitra-Wan take N to be the age at which the biomass per unit of land is maximized, claiming that "for any reasonable objective function for the economy, trees will never be allowed to grow beyond age N; we therefore take this as a condition of feasibility itself."[19] It is this reasoning that allows the authors to limit themselves to an N-dimensional state vector. However, given the fact that a concavity benefit function favors a homogeneously configured forest, the planner may well adopt the trade-off of postponing harvesting beyond age N in order to reshape the forest into a more homogeneous state. We circumvent this by simply assuming n to be the age at which a tree dies.[20]

For each period $t \in \mathbb{N}$ we denote $x_i(t) \geq 0,\ i = 1, \ldots, n$, the surface occupied by trees of age i at time t. We represent the state of the forest by the vector $x(t) = (x_1(t), \ldots, x_n(t)) \in \Delta$.

At every stage we must decide how much land to harvest of every age-class, $c(t) = (c_1(t), \ldots, c_n(t))$ where $c_i(t) \in [0, x_i(t)]$. As we know that after n a tree has no value, we assume that $c_n(t) = x_n(t)$ for all t. By the end of period $t+1$, the state will be exactly

$$x(t+1) = \big(\textstyle\sum_{i=1}^n c_i(t), x_1(t) - c_1(t), \ldots, x_{n-1}(t) - c_{n-1}(t)\big).$$

necessary and sufficient condition presented as the non-interiority condition 2.1 below.

[19] See [15, p. 232]. The same point is made in [17, Section 4, Paragraph 5].

[20] This is simply a somewhat subtle point of interpretation; the technicalities of the two analyses remain the same.

Definition 2.1 *A sequence* $\{x(t)\}_{t=0}^{\infty}$ *is called a program if for each* $t \geq 0$

$$\begin{cases} x(t) \in \Delta, \\ x_{i+1}(t+1) \leq x_i(t) \quad i = 1, \dots, n{-}1 \end{cases} \tag{1}$$

Definition 2.2 *Let* T_1 *and* T_2 *be integers such that* $0 \leq T_1 < T_2$. *A sequence* $\{x(t)\}_{t=T_1}^{t=T_2}$ *is called a program if* $x(T_2) \in \Delta$ *and relations* (1) *hold for each* t *satisfying* $T_1 \leq t < T_2$.

Define the transition possibility set Ω as the collection of pairs $(x, x') \in \Delta \times \Delta$ such that it is possible to go from the state x in the current period (today) to the state of the forest x' in the next period (tomorrow) fulfilling relations (1). Formally,

$$\Omega = \{(x, x') \in \Delta \times \Delta \,/ x_i \geq x'_{i+1} \text{ for all } i = 1, \dots, n{-}1\}$$

Definition 2.3 *The vector of harvests needed to perform this transition is given by the function* $\lambda : \Omega \to I\!R^n_+$,

$$\lambda(x, x') = (x_1 - x'_2, x_2 - x'_3, \dots, x_{n-1} - x'_n, x_n)$$

In addition, it is easy to see that

$$(x, x') \in \Omega \Leftrightarrow x, x' \in \Delta \text{ and } \lambda(x, x') \geq 0$$

The preferences of the planner are represented by a felicity function, $w : [0, \infty) \to I\!R$ which is assumed to be continuous, strictly increasing and concave. Define for any $(x, x') \in \Omega$ the function $u(x, x')$ as

$$u(x, x') = w(bc) \text{ where } c = \lambda(x, x')$$

Definition 2.4 *A* golden-rule stock $\hat{x} \in I\!R^n_+$ *is such that* $(\hat{x}, \hat{x})$ *is a solution to the problem:*

$$\begin{cases} \textit{maximize} \;\; u(x, x) \\ \textit{subject to} \;\; (x, x) \in \Omega \end{cases}$$

We now present some basic antecedent results, and except those indicated at the end of the section, they are all taken from [9].

Theorem 2.1 *There exists a unique golden-rule stock* $\hat{x} = (\underbrace{\tfrac{1}{\sigma}, \dots, \tfrac{1}{\sigma}}_{\sigma}, 0, \dots, 0)$

We denote by $\hat{c}$ the vector of harvests obtained by the pair $(\hat{x}, \hat{x})$, namely $\hat{c} = \lambda(\hat{x}, \hat{x})$. Observe that $\hat{c}_\sigma = \frac{1}{\sigma}$ and $\hat{c}_i = 0$ for all $i \neq \sigma$. The total amount of harvest obtained each period of the forest remains always at the golden rule stock is $b\hat{c} = \frac{b_\sigma}{\sigma}$. Pick any $z \in \partial^+ w(\frac{b_\sigma}{\sigma})$, [21] and set $\hat{p} \in I\!R^n_+$, $\hat{p} = z\frac{b_\sigma}{\sigma}(1, 2, \dots, n)$.

Definition 2.5 *We define the* value loss *associated with any* $(x, x') \in \Omega$ *to be*

$$\delta(x, x') = w(\tfrac{b_\sigma}{\sigma}) - w\big(b\lambda(x, x')\big) - \hat{p}(x' - x).$$

It is easy to see that the function $\delta(\cdot, \cdot)$ is convex and the following lemma asserts that $\delta(x, x') \geq 0$ for any $(x, x') \in \Omega$.

Lemma 2.1 *For any* $(x, x') \in \Omega$ *we have*

$$\delta(x, x') \geq z\big[\textstyle\sum_{i=1}^{n-1} (\frac{b_\sigma}{\sigma} - \frac{b_i}{i})\, i\, (x_i - x'_{i+1}) + (\frac{b_\sigma}{\sigma} - \frac{b_n}{n}) n x_n\big] \geq 0 \tag{2}$$

[21] We know that $\partial^+ w(\cdot) \neq \phi$ due to the concavity of w and that $z > 0$ because w is strictly increasing. [20]

A warning to the reader regarding notation: we shall consistently use the symbol $\delta(\cdot,\cdot)$ to denote a function, and it is to be distinguished from the number δ, typically assumed to be positive.

We use the following notion of *good* and *bad* programs introduced by Gale (1967)

Definition 2.6 *A program $\{x(t)\}$ is called* good *if there exists $M \in I\!R$ such that for all $T \geq 0$, $\sum_{t=0}^{T}[w(bc(t)) - w(\frac{b_\sigma}{\sigma})] \geq M$, where $c_t = \lambda\big(x(t), x(t+1)\big)$. A program is* bad *if $\lim_{T\to\infty}\sum_{t=0}^{T}[w(bc(t)) - w(\frac{b_\sigma}{\sigma})] = -\infty$.*

The following general result of Gale applies to the MW model.

Proposition 2.1 *Programs are partitioned into good and bad programs. Furthermore,*

i. $\{x(t)\}$ *is good iff* $\sum_{t=0}^{\infty}\delta\big(x(t), x(t+1)\big) < \infty$.

ii. $\{x(t)\}$ *is bad iff* $\sum_{t=0}^{\infty}\delta\big(x(t), x(t+1)\big) = \infty$.

Let $x_0 \in \Delta$. Set $\mu(x_0) = \inf\big\{\sum_{t=0}^{\infty}\delta\big(x(t), x(t+1)\big) : \{x(t)\} \text{ is a program from } x_0\big\}$. It is possible to see that there exists at least one good program from every $x_0 \in \Delta$, which in turn implies that $\mu(x_0) < \infty$. The following result can now be established.

Proposition 2.2 *From any $x_0 \in \Delta$ there exist a good program $\{x(t)\}$ such that*

$$\sum_{t=0}^{\infty}\delta\big(x(t), x(t+1)\big) = \mu(x_0). \tag{3}$$

The fact that every good program converges to the golden rule stock in the case that w is strictly concave was established in [15, Lemma 6.4]. In [9], Khan-Piazza provide a necessary and sufficient condition to assure the convergence of every good program to the golden rule stock for any concave utility function w that is not necessarily differentiable. We describe this characterization in the following terms.

Let the *discrepancy function* f be

$$f(\xi) = w(\tfrac{b_\sigma}{\sigma}) - w(b_\sigma\xi) + z(b_\sigma\xi - \tfrac{b_\sigma}{\sigma}). \tag{4}$$

We can appeal to standard results in [20] to assert that the concavity of w implies $f(\frac{1}{\sigma}) = 0$, $f(\xi) \geq 0$ for all ξ and f attains its minimum in a closed interval S_f containing $(\frac{1}{\sigma})$.

We now turn to the condition that serves as a basic standing hypothesis for our principal results.

Condition 2.1 (Non-Interiority) $(1/\sigma) \notin$ *int* S_f.

Either w coincides with the support function, $w(\frac{b_\sigma}{\sigma}) + z(\cdot - \frac{b_\sigma}{\sigma})$ only at the point $\frac{b_\sigma}{\sigma}$ or this point is one of the extremes of the interval where the two functions coincide. Of course, the non-interiority condition 2.1 is assured if w is strictly concave, but there is a broader family of functions satisfying it.

Next, we define the following subsets of $I\!R^n_+$:

$$\begin{aligned} S_c &= \{c \in I\!R^n_+ /\ c_\sigma \in S_f \ \text{ and } \ c_i = 0 \ \text{ for all } \ i \neq \sigma\} \\ V &= \{x \in \Delta /\ x_i \in S_f \ \text{ for all } \ i \leq \sigma \ \text{ and } \ x_i = 0 \text{ for all } \ i > \sigma\} \end{aligned} \tag{5}$$

As discussed in [9], the following results are obtained without the non-interiority condition 2.1.

Proposition 2.3 *The von Neumann facet is*

$$\big\{(x, x') \in \Omega /\ \delta(x, x') = 0\big\} = \big\{(x, x') \in \Omega /\ \lambda(x, x') \in S_c\big\}.$$

Remark 2.1 *Given $x \in V$, consider the σ-periodic program from x where the harvest consist of all the trees of the σ-th age class. This particular program has zero accumulated value loss, hence $\mu(x) = 0$ iff $x \in V$.*

Lemma 2.2 [22] *Every good program $\{x(t)\}$ is such that $dist(x(t), V) \to 0$*

Next, we present the optimality criteria we shall be working with.

Definition 2.7 *A program $\{x^*(t)\}$ is* optimal *if for any program $\{x(t)\}$ such that $x(0) = x^*(0)$ we have*

$$\limsup_{T\to\infty} \sum_{t=0}^{T} w(bc(t)) - w(bc^*(t)) \le 0$$

Definition 2.8 *A program $\{x^*(t)\}$ is* maximal *if for any program $\{x(t)\}$ such that $x(0) = x^*(0)$ we have*

$$\liminf_{T\to\infty} \sum_{t=0}^{T} w(bc(t)) - w(bc^*(t)) \le 0$$

If the non-interiority condition 2.1 does not hold, we cannot assure the existence of an optimal program from any $x_0 \in \Delta$, but only that of a maximal program. This follows from Proposition 2.2 that assures the existence of a minimizer of the accumulated value loss function and the following result.

Proposition 2.4 *If $\{x(t)\}$ is a program from x_0 that minimizes the accumulated value loss $\left(\sum_t \delta\big(x(t), x(t+1)\big) = \mu(x_0)\right)$, then $\{x(t)\}$ is a maximal program from x_0.*

Corollary 2.1 *Every maximal program is good. Hence it converges to the set V.*

Lemma 2.3 *The non-interiority condition 2.1 holds iff $\{\hat{x}\} = V$.*

From now on, and throughout the rest of this work, we will assume that the non-interiority condition 2.1 holds. Let us spell out two preliminary results that we have with this added hypothesis. First, we present a stronger version of Lemma 2.2,

Lemma 2.4 *Any good program $\{x(t)\}$ satisfies $\lim_{t\to\infty} x(t) = \hat{x}$.*

Second, the existence of an optimal program is assured by the following equivalence,

Theorem 2.2 *Let $\{x(t)\}$ be a program from x_0. If the non-interiority condition 2.1 holds, the following conditions are equivalent:*

i. $\sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big) = \mu(x(0))$

ii. $\{x(t)\}$ *is optimal.*

iii. $\{x(t)\}$ *is maximal.*

Next, we observe that the two basic results in [8] can be generalized under the non-interiority condition 2.1 to yield the following versions whose straightforward proofs we leave to the reader. These are the only new results (so to speak) in this section.

Theorem 2.3 *Let $\epsilon > 0$. If the non-interiority condition 2.1 holds, there exists $\delta > 0$ such that for each optimal program $\{x(t)\}$ satisfying $\|x(0) - \hat{x}\|_\infty < \delta$ the following inequality holds:*

$$\|x(t) - \hat{x}\|_\infty < \epsilon \text{ for all } t \ge 0$$

Theorem 2.4 *Let $\epsilon > 0$. If the non-interiority condition 2.1 holds, there exists a natural number T_0 such that for each optimal program $\{x(t)\}$ the following inequality holds:*

$$\|x(t) - \hat{x}\|_\infty < \epsilon \text{ for all } t \ge T_0$$

[22] This lemma is the analogous to [15, Lemma 6.4].

3 Principal Results

We now present the principal results of this work. As emphasized in the introduction, both results are classical in that they pertain to the following situation: if an initial configuration of a forest is given, and a particular terminal configuration stipulated, then with enough time at his or her disposal, the planner ought to stay arbitrarily near the golden-rule configuration, and realize timber yields arbitrarily close to the maximally sustainable ones. However, the the less of an error that the planner is allowed in steering the forest configuration away from the golden-rule configuration, the larger the time horizon he or she would require.

For the rest of the paper, we will assume that the non-interiority condition 2.1 holds. We introduce notation for the aggregate value of finite optimal programs. Let $z_0, z_f \in \Delta,\ T \geq 1$, and

$$U(z_0, T) = \sup\left\{\sum_{t=0}^{T-1} w\big(bc(t)\big) \big/ \{x(t)\}_{t=0}^{T} \text{ is a program from } z_0\right\}, \tag{6}$$

$$U(z_0, z_f, 0, T) = \sup\left\{\sum_{t=0}^{T-1} w\big(bc(t)\big) \big/ \{x(t)\}_{t=0}^{T} \text{ is a program from } z_0 \text{ with } x(T) = z_f\right\}. \tag{7}$$

Observe that whenever there is no program $\{x(t)\}_{t=0}^{T}$ such that $x(0) = z_0$ and $x(T) = z_f$, we shall assume as a matter of mathematical conventions that $U(z_0, z_f, 0, T) = -\infty$.

Theorem A *Given $M > 0$ and $\epsilon > 0$ there exists $L \in I\!N$ such that for all $T > L$ and each program $\{x(t)\}_{t=0}^{T}$ satisfying*

$$\sum_{t=0}^{T-1} w(bc(t)) \geq U\big(x(0), x(T), 0, T\big) - M$$

we have

$$Card\big\{i \in [0, \ldots, T-1] : \|x(t) - \hat{x}\| > \epsilon\big\} \leq L.$$

Theorem B *Let $\epsilon > 0$. Then there exist $\tau \in I\!N$ and $M > 0$ such that for all $T > 2\tau + n + \sigma$ and each program $\{x(t)\}_{t=0}^{T}$ satisfying*

$$\sum_{t=0}^{T-1} w(bc(t)) \geq U\big(x(0), x(T), 0, T\big) - M$$

there are τ_1, τ_2 such that $\tau_1 \in [0, \tau]$, $\tau_2 \in [T-\tau, T]$ and

$$\|x(t) - \hat{x}\| \leq \epsilon \ \text{ for all } \ t = \tau_1, \ldots, \tau_2.$$

Moreover, if $\|x(0) - \hat{x}\| \leq \epsilon/n^2$ then $\tau_1 = 0$.

In terms of basic conception, the results are classical in the sense that the term is delineated in Khan-Zaslavski (2009). Thus, rather than the Samuelsonian triple limit alluded to in his Nobel Lecture [26], an interesting "quarter limit" is involved, and four separate considerations are quantified: the length of the time-horizon, the proximity to the the golden-rule forest configuration, the length of time that is spent within this proximity, and a "degree of slack" in the attainment of the objective. For an index of proximity quantified by ϵ, and index of slack quantified by M, Theorem A furnishes a bound L such that for time-horizon levels T greater than L, any M-optimal forest configuration lies within the golden-rule configuration for all $(T - L)$ number of periods. Since L is independent of the time-horizon, the optimal configuration lies close to the golden-rule configuration most of the time. If the planner is not allowed to choose the degree of slack M, the result can be strengthened to guarantee that the time-periods spent in proximity to the turnpike are consecutive. This is formalized in Theorem B.

Next, we observe that Theorem 2.4 above also follows as a straightforward consequence of Theorem B. To see this, simply note that any optimal infinite-horizon program $\{x(t)\}$, when truncated to T periods, is an optimal T-period program with its own initial and terminal configurations, $x(0)$ and $x(T)$; and that the Theorem furnishes us with L independent of both T *and* these forest configurations. This alternative proof is of some methodological significance in that it shows that a result on (uniform) asymptotic stability of the golden-rule forest configuration follows from the turnpike result classically conceived, and thereby establishes the primacy of McKenzie's so called *middle* turnpike over his *late* turnpike.

Theorem B also allows the deduction of the following two corollaries, and thereby establishes the primacy of McKenzie's so-called *middle* turnpike over his *early* turnpike.

Corollary 3.1 *Given $\epsilon > 0$ and $M \geq 0$, there exists $L \in \mathbb{N}$ such that for each $T > L$, and any two programs $\{x_a\}_{t=0}^{T}$ and $\{x_a\}_{t=0}^{T}$ satisfying*

$$x_a(0) = x_b(0) = x_0,\ x_a(T) = x_b(T) = x_T,\ \textstyle\sum_{t=0}^{T-1} w\big(bc(t)\big) \geq U(x_0, x_T, 0, T) - M,$$

where $c(t)$ stands alternatively for $c_a(t)$ and $c_b(t)$, following inequality holds:

$$Card\big\{i \in [0, \ldots, T-1] : \|x_a(t) - x_b(t)\|_\infty > \epsilon\big\} \leq L.$$

And if the initial forest configuration is the golden-rule configuration, Corollary 3.2 can be strengthened to yield the fact that it is the (arbitrarily large) initial interval in which the optimal program is close to the turnpike.[23]

Corollary 3.2 *Given $\epsilon > 0$, there exists $L \in \mathbb{N}$ such that for each $T > L$, and any two programs $\{x_a\}_{t=0}^{T}$ and $\{x_a\}_{t=0}^{T}$ satisfying*

$$x_a(0) = x_b(0) = \hat{x},\ x_a(T) = x_b(T) = x_T,\ \textstyle\sum_{t=0}^{T-1} w\big(bc(t)\big) \geq U(x_0, x_T, 0, T),$$

where $c(t)$ stands alternatively for $c_a(t)$ and $c_b(t)$, there is an integer $\tau \in [T - L, T]$ such that

$$\|x_a(t) - x_b(t)\|_\infty \leq \epsilon \text{ for all } t = 1, \cdots, \tau.$$

4 Preliminary Substantive Results

In this section and the next we begin developing the technical arguments needed to prove the principal results of this work. The five propositions presented here develop intuition into the basic dynamics underlying the MW model, and even though the proofs are notationally somewhat complex, the essential ideas are simple.[24]

Proposition 4.1 is a basic result of the subject that given any two forest configurations, there exists a program of n time periods that allows the planner to move from one configuration to another, n being the number of ages at which a particular tree can be tracked.[25] We shall be going into some detail as regards the comparison with [10], but for readers not particularly interested in the comparison, let us simply observe here that despite their superficial resemblance, Propositions 6.1 to 6.4 in [10] do not furnish the precise estimates that are offered here in Propositions 4.1 to 4.4, and strictly speaking there is no analogue to them in the RSS model.

Proposition 4.1 *For every $z_0, z_f \in \Delta$, there exists a program $\{x(t)\}_{t=0}^{n}$ such that $x(0) = z_0$ and $x(n) = z_f$*

The following corollary presents a refinement to programs that are not of n time periods.

Corollary 4.1 *Let $z_0, z_f \in \Delta$. (i) If $T \geq n$, there exists a program $\{x(t)\}_{t=0}^{T}$ such that $x(0) = z_0$ and $x(T) = z_f$. (ii) If $T < n$, there exists a program $\{x(t)\}_{t=0}^{T}$ such that $x(0) = z_0$ and $x(T) = z_f$ iff $z_{0,i} \geq z_{f,i+T}$ for all $i = 1, \ldots, T-n$.*

[23] This is of course the analogue of the time-honored strengthening of Radner's result by Nikaido; see [10] for references and further discussion in terms of the RSS model.

[24] The reader could check her understanding of the basics of the model by trying to figure out the proofs for herself before looking at the ones presented in the Appendix.

[25] A version of this result was first presented in [17].

There is no presumption that the program whose existence is asserted in the above claims is optimal in any sense. We turn to finite optimal programs in the next two results. Proposition 4.2 is phrased in terms only of an initial forest configuration, and claims that the average benefit of a large but finite program can get arbitrarily close to that obtained from the maximally sustainable stationary timber yield as the time horizon becomes large enough. Proposition 4.3 makes a similar claim when the terminal forest configuration is also specified.

Proposition 4.2 *For each $z \in \Delta$, and each $T \in I\!\!N$,*

$$U(z,T) \geq Tw(b\hat{c}) - \sigma w(b\hat{c}).$$

Proposition 4.3 *Given $z_0, z_f \in \Delta$ and $T \geq n$, we have*

$$U(z_0, z_f, 0, T) \geq Tw(b\hat{c}) - (n+\sigma)w(b\hat{c}). \tag{8}$$

If $T < n$ and there is a program $\{x(t)\}_{t=0}^{T}$ satisfying that $x(0) = z_0$ and $x(T) = z_f$ then inequality (8) *also holds.*

Next, we turn to a simple inequality that follows from the fact that the value-loss of any production plan is non-negative.

Proposition 4.4 *For every T and every program $\{x(t)\}_{t=0}^{T}$ the following inequality is satisfied:*

$$\textstyle\sum_{t=0}^{T-1}[w(bc(t)) - w(b\hat{c})] \leq n(b_\sigma/\sigma).z \tag{9}$$

Our final result asserts that any finite program that is optimal with a particular level of approximation, has its sub-programs also optimal with respect to the same level of approximation. It is analogous to [10, Proposition 6.6].

Proposition 4.5 *Let $T \in I\!\!N$, $M > 0$ and $\{x(t)\}_{t=0}^{T}$ be a program such that $\sum_{t=0}^{T-1} w\big(bc(t)\big) \geq U(x(0), x(T), 0, T) - M$. Then for all S_1 and S_2, $0 \leq S_1 < S_2 < T$, we have*

$$\textstyle\sum_{t=S_1}^{S_2-1} w(bc(t)) \geq U\big(x(S_1), x(S_2), S_1, S_2\big) - M$$

5 Four Substantive Lemmas

With these preliminary results out of the way, we can turn to the deeper substance of the argumentation. As in [10], it revolves around the four footholds presented below as Lemmata 5.1 to 5.4: the *visiting* lemma, the *stability* lemma, the *value-loss* lemma and the *aggregate value-loss* lemma. However, before we take each in turn, it is worth elaborating on what was already emphasized in the introduction: that even though these results are inspired by Lemmata 7.1 to 7.4 for the RSS model, the particularities of the MW model bring in analytical difficulties of their own that need to be surmounted. Briefly put, in the RSS model a unit amount of labor is to be allocated to the production of a single consumption good and to the production of n types of machines; whereas in the MW model, a unit amount of land is to be parcelled out between the cultivation of trees of n possible ages. Thus in one, the stock variable is an element of $I\!\!R^n$ with a well-defined order on it; whereas in the other, it is a probability measure with a finite support that can be represented as a point in the simplex in $I\!\!R^n$, with no clear order, and therefore little possibility of formalizing notions either of *inaction* or of *free disposal.*[26] All this adds renders the proof of a result in one inapplicable to that of the other, and adds to their considerable complication. This is especially true of the Lemmata 5.3 and 5.4 below, the Radner [19] *value-loss* and *aggregate value-loss* lemmata. Note also the considerable sharpening of the conclusion of Lemma 5.2 and Corollary 5.1 relative to their RSS counterparts in [10].

[26] This point of view as regards the Mitra-Wan tree-farm is original to [8].

Lemma 5.1 *Given $M > 0$ and $\epsilon > 0$, there exists $\tau \in I\!N$ such that for each program $\{x(t)\}_{t=0}^{\tau}$ satisfying*

$$\textstyle\sum_{t=0}^{\tau-1} w(bc(t)) \geq \tau w(b\hat{c}) - M,$$

there exists $t \in [0, \tau]$ such that $\|x(t) - \hat{x}\| \leq \epsilon$.

Lemma 5.2 *Let the program $\{x(t)\}_{t=0}^{n}$ be such that*

$$\delta\big(x(t), x(t{+}1)\big) = 0 \quad \textit{for } t = 0, \ldots, n{-}1. \tag{10}$$

then $x(t) = \hat{x}$ for all $t \in [0, \sigma]$.

Corollary 5.1 *Given $\epsilon > 0$, there exist $\gamma > 0$ such that for each program $\{x(t)\}_{t=0}^{n}$ satisfying $\delta\big(x(t), x(t{+}1)\big) < \gamma$ for $t = 0, \ldots, n{-}1$, we have $\|x(t) - \hat{x}\| < \epsilon$ for all $t = 0, \ldots, \sigma$.*

Lemma 5.3 *Given $\epsilon > 0$, there exists $\gamma > 0$ such that for each $T \in I\!N$, and each program $\{x(t)\}_{t=0}^{T}$ satisfying $\|x(0) - \hat{x}\| < \epsilon/n$, $\|x(T) - \hat{x}\| < \epsilon/n$ and $\delta\big(x(t), x(t{+}1)\big) < \gamma$ for all $t = 0, \ldots, T{-}1$, we have*

$$\|x(t) - \hat{x}\| < \epsilon \quad \textit{for all} \quad t = 0, \ldots, T.$$

Lemma 5.4 *Given $\epsilon > 0$, there exist $\gamma > 0$ and $M > 0$ such that for each $T \geq n{+}\sigma$, and each program $\{x(t)\}_{t=0}^{T}$ satisfying $\|x(0) - \hat{x}\| \leq \gamma$, $\|x(T) - \hat{x}\| \leq \gamma$ and $\sum_{t=0}^{T-1} w(bc(t)) \geq U(x(0), x(T), 0, T) - M$, we have*

$$\textstyle\sum_{t=0}^{T-1} \delta\big(x(t), x(t{+}1)\big) < \epsilon.$$

For a detailed discussion and interpretation of these results, the reader is referred to [10, Section 4].

6 Concluding Remarks

We now conclude the non-technical part of this work by delineating three directions in which the results demand extension and further investigation.

The first of these is the rather immediate question as to how much of the theory can be salvaged when the non-interiority condition 2.1 does *not* hold? Since this condition is necessary *and* sufficient for asymptotic convergence of good programs, and it is easy to provide examples of periodic optimal programs when it does not hold, perhaps the obvious answer is simply: none of it. However, the question clearly deserves another less-facile look. The fact that infinite-horizon optimal optimal programs converge to the von-Neumann facet is a basic result of the subject, and surely what is true of asymptotic convergence could also possibly be true of classical turnpike theory.

This work, along with that of [10], has taken classical turnpike theory (circumscribed as it is by assumptions of uniform strict concavity, and on occasion, differentiability, of benefit (felicity) functions) and extended it to concave functions that are not necessarily differentiable. In terms of a second direction, one is then naturally led to ask whether one can relax the concavity (and continuity) assumption itself? This question has not been posed so far in the capital-theory literature, but if mathematical economics is to justify itself as an intellectually worthwhile activity, surely that justification must revolve in part on the acceptance of each new result leading to the pursuit of questions that would not have been considered even remotely feasible before it.

These two directions stems directly from the two theorems reported here; the third is rather more overarching. A subtext of this entire work is the (somewhat uneasy) relationship between the RSS and MW models, with the relative difficulties of one being matched by those of another, and presenting inevitable analytical trade-offs. This tension clearly asks for a move towards a synthesis that obtains both models as special cases, and provides a non-trivial extension to what is now associated with the names of Gale and McKenzie and is frequently referred to as the general theory of intertemporal allocations of resources.

7 Appendix: Technicalities of Proofs

Proof of Proposition 4.1: We propose the following program $\{x(t)\}_{t=0}^{n}$,

$$\begin{array}{rcl} x(0) &=& z_0 \\ x(1) &=& \big(1,0,\dots,0\big) \\ x(2) &=& \big(\sum_{i=1}^{n-1} z_{f,i}\,,\, z_{f,n}\,,\, 0,\dots,0\big) \\ \vdots & & \vdots \\ x(j) &=& \big(\sum_{i=1}^{n-j+1} z_{f,i}\,,\, z_{f,n-j+2}\,,\dots, z_{f,n-1}\,,\; z_{f,n}\,,\, \underbrace{0,\dots,0}_{n-j}\big) \\ \vdots & & \vdots \\ x(n) &=& z_f \end{array} \qquad \begin{array}{rcl} c(0) &=& z_0 \\ c(1) &=& \big(\sum_{i=1}^{n-1} z_{f,i}\,,\, 0,\dots,0\big) \\ c(2) &=& \big(\sum_{i=1}^{n-2} z_{f,i}\,,\, 0,\dots,0\big) \\ \vdots & & \vdots \\ c(j) &=& \big(\sum_{i=1}^{n-j} z_{f,i}\,,\, 0,\dots,0\big) \\ \vdots & & \vdots \end{array}$$

■

Proof of Corollary 4.1: *i.* First consider *any* program $\{x(t)\}_{t=0}^{T-n}$ from z_0. The proposition above tells us that there is a program $\big\{\bar{x}(t)\big\}_{t=0}^{n}$ where $\bar{x}(0) = x(T-n)$ and $\bar{x}(n) = z_f$. The proof follows by defining $x(t+T-n) = \bar{x}(t)$ for $t = 1,\dots,n$.
ii. It follows easily from the definition of program. ■

Proof of Proposition 4.2: If $T \le \sigma$ then the proposition follows directly because $U(z,T) \ge 0 \ge Tw(b\hat{c}) - \sigma w(b\hat{c})$.

If $T > n$, consider the following program $\{x(t)\}_{t=0}^{\infty}$ from z

$$\begin{array}{rcl} x(0) &=& z \\ x(1) &=& \big(1,0,\dots,0\big) \\ x(2) &=& \big(\frac{\sigma-1}{\sigma}\,,\frac{1}{\sigma}\,,0,\dots,0\big) \\ \vdots & & \vdots \\ x(j) &=& \big(\frac{\sigma-j+1}{\sigma}\,,\underbrace{\frac{1}{\sigma},\dots,\frac{1}{\sigma}}_{j-1},\underbrace{0,\dots,0}_{n-j}\big) \\ \vdots & & \vdots \\ x(t) &=& \hat{x} \qquad \text{for all } t \ge \sigma \end{array} \qquad \begin{array}{rcl} c(0) &=& z \\ c(1) &=& \big(\frac{\sigma-1}{\sigma}\,,0,\dots,0\big) \\ c(2) &=& \big(\frac{\sigma-2}{\sigma}\,,0,\dots,0\big) \\ \vdots & & \vdots \\ c(j) &=& \big(\frac{\sigma-j}{\sigma}\,,0,\dots,0\big) \\ \vdots & & \vdots \\ c(t) &=& \hat{c} \qquad \text{for all } \sigma \le t \le T \end{array}$$

And we deduce

$$U(z,T) \ge \sum_{t=0}^{T-1} w\big(bc(t)\big) \ge \sum_{t=\sigma}^{T-1} w\big(bc(t)\big) = (T-\sigma)w\big(b\hat{c}\big)$$

■

Proof of Proposition 4.3: If $T \ge n$, by Proposition 4.2 we know that there is a program $\{x(t)\}_{t=0}^{T-n}$ from z_0 such that

$$\sum_{t=0}^{T-n-1} w\big(bc(t)\big) \ge (T-n)w(b\hat{c}) - \sigma w(b\hat{c})$$

and by Proposition 4.1, there is a program $\big\{x(t)\big\}_{t=T-n}^{T}$ from $x(T-n)$ such that $x(T) = z_f$. Concatenating the two, we obtain the program $\{x(t)\}_{t=0}^{T}$ from z_0 such that $x(T) = z_f$ and we can deduce

$$U(z_0, z_f, 0, T) \ge \sum_{t=0}^{T-1} w\big(bc(t)\big) \ge \sum_{t=0}^{T-n-1} w\big(bc(t)\big) \ge (T-n)w(b\hat{c}) - \sigma w(b\hat{c}) = Tw(b\hat{c}) - (n+\sigma)w(b\hat{c}).$$

If $T < n$, then the existence of the program $\{x(t)\}_{t=0}^{T}$ assures that $U(z_0, z_f, 0, T) \geq 0$ and observing that $(T-n-\sigma)w\big(bc(t)\big) < 0$ the proof follows. ■

Proof of Proposition 4.4: We know that $\delta\big(x(t), x(t+1)\big) \geq 0$, hence

$$0 \leq \sum_{t=0}^{T-1} \delta\big(x(t), x(t+1)\big) = \sum_{t=0}^{T-1} [w(\tfrac{b_\sigma}{\sigma}) - w\big(bc(t)\big)] + \hat{p}\ (x(0) - x(T)).$$

Using the definition of Δ we deduce that

$$\sum_{t=0}^{T-1} w\big(bc(t)\big) - w(\tfrac{b_\sigma}{\sigma}) \leq \|\hat{p}\|\ \|x(0) - x(T)\| \leq \|\hat{p}\| = n\tfrac{b_\sigma}{\sigma}z. \tag{11}$$

■

Proof of Proposition 4.5: Suppose by contradiction that there exist S_1 and S_2, $(0 \leq S_1 < S_2 < T)$ such that $\sum_{t=S_1}^{S_2-1} w\big(bc(t)\big) < U\big(x(S_1), x(S_2), S_1, S_2\big) - M$. Then, there is a program $\{x'(t)\}_{t=S_1}^{S_2}$ such that $x'(S_1) = x(S_1)$, $x'(S_2) = x(S_2)$ and

$$\sum_{t=S_1}^{S_2-1} w\big(bc(t)\big) < \sum_{t=S_1}^{S_2-1} w(bc'(t)) - M.$$

Extend the program to the time interval $[0, \ldots, T]$ as follows

$$x'(t) = \begin{cases} x(t) & t = 0, \ldots, S_1 - 1 \\ x(t) & t = S_2 + 1, \ldots, T \end{cases}$$

And the following contradiction arises, $\sum_{t=0}^{T} w\big(bc(t)\big) < \sum_{t=0}^{T} w(bc'(t)) - M \leq U(z_0, z_f, 0, T) - M.$ ■

Proof of Lemma 5.1: Let us assume the contrary: for each $k \in I\!N$ there exists a program $\{x^k(t)\}_{t=0}^{k}$ such that

$$\|x^k(t) - \hat{x}\| > \epsilon \ \text{ and } \ \sum_{t=0}^{k-1} w\big(bc^k(t)\big) \geq kw(\tfrac{b_\sigma}{\sigma}) - M \tag{12}$$

Let $M' = n\frac{b_\sigma}{\sigma}z > 0$, by (9) we know that every program fulfills $\sum_{t=0}^{T-1} w\big(bc(t)\big) - w(\frac{b_\sigma}{\sigma}) \leq M'$. Given any $s < k$, by combining the two previous inequalities we deduce

$$\sum_{t=0}^{s-1}[w\big(bc^k(t)\big) - w(\tfrac{b_\sigma}{\sigma})] = \sum_{t=0}^{k-1}[w\big(bc^k(t)\big) - w(\tfrac{b_\sigma}{\sigma})] - \sum_{t=s}^{k-1}[w\big(bc^k(t)\big) - w(\tfrac{b_\sigma}{\sigma})] \geq -(M + M') \tag{13}$$

By extracting a subsequence and a diagonalization process we obtain that there exist a strictly increasing sequence of natural numbers $\{k_j\}_{j=1}^{\infty}$ and a sequence $\{x^*(t)\}_{t \in I\!N}$ such that

$$x^{k_j}(t) \to x^*(t) \ \text{ when } \ j \to \infty \ \text{ for all } t \geq 0$$

It is not difficult to see that $\{x^*(t)\}_{t \in I\!N}$ is a program. From (13) we deduce that for every natural number s, $\sum_{t=0}^{s-1} w\big(bc^*(t)\big) - sw(\frac{b_\sigma}{\sigma}) \geq -M - M'$, meaning that $\{x^*(t)\}_{t \in I\!N}$ is a good program. Then Lemma 2.2 implies that

$$\|x^*(t) - \hat{x}\| \to 0 \ \text{ when } \ t \to \infty.$$

On the other hand, it follows from (12) and the definition of $x^*(t)$ that

$$\|x^*(t) - \hat{x}\| > \epsilon \ \text{ for all } \ t$$

and the contradiction proves the lemma. ■

Proof of Lemma 5.2: By Lemma 2.3, we know that $c(t) \in S_c$ for $0 \le t < n$ which implies that $x_i(\sigma) = 0$ for all $i > \sigma$. Indeed, if there was $j > \sigma$ such that $x_j(\sigma) > 0$, then we would have $x_n(n+\sigma-j) > 0$ and $c(n+\sigma-j) \notin S_c$.

From the above we know that

$$x(\sigma) = \big(c_\sigma(\sigma-1), c_\sigma(\sigma-2), \dots, c_\sigma(0), 0 \dots, 0\big)$$

where $c_\sigma(i) \in S_f$. The area balance together with condition (2.1) implies $x(\sigma) = \hat{x}$.

Finally, it is easy to see that $x(t+1) = \hat{x}$ and $\delta\big(x(t), x(t+1)\big) = 0$ imply $x(t) = \hat{x}$ and then the proposition follows by backwards induction. ■

Proof of Corollary 5.1: Suppose, contrary to our claim, that for every k there is $\{x^k(t)\}_{t=0}^n$ such that $\delta\big(x^k(t), x^k(t+1)\big) \le \frac{1}{k}$ and there is $t_k \in [0, \dots, \sigma]$ satisfying $\|x^k(t_k) - \hat{x}\| \ge \epsilon$. As $\{t_k\} \subseteq [0, \dots, \sigma]$ there must be at least one value $t_0 \in [0, \dots, \sigma]$ such that $t_k = t_0$ infinitely many times. We extract a subsequence $\{k_j\}$ such that $t_{k_j} = t_0$ and hence $\|x^{k_j}(t_0) - \hat{x}\| \ge \epsilon$. By extracting a subsequence from $\{k_j\}$ (to simplify the notation, we denote this subsequence also by $\{k_j\}$), we obtain that there is a sequence $\{x^*(t)\}$ such that

$$x^{k_j}(t) \to x^*(t) \quad \text{for all } t = 0, \dots, n.$$

It is easy to see that $\{x^*(t)\}$ is a program and by the continuity of $\delta(\cdot,\cdot)$ we have that $\delta\big(x^*(t), x^*(t+1)\big) = 0$ for all $t = 0, \dots, n$ and the lemma above implies that $x^*(t) = \hat{x}$ for all $t = 0, \dots, \sigma$.

On the other hand, $\|x^{k_j}(t_0) - \hat{x}\| \ge \epsilon$ for all k_j hence $\|x^*(t_0) - \hat{x}\| \ge \epsilon$ and a contradiction arises proving the corollary. ■

Proof of Lemma 5.3: We divide the proof into two parts: $T < n$ and $T \ge n$.

Case $T < n$**:** A procedure similar to the one on the corollary above allows to affirm that given ϵ there is γ_1 such that: $\delta(x, x') < \gamma_1$ implies $\operatorname{dist}(\lambda(x, x'), S_c) < \epsilon_1 = \epsilon/(n2^n)$. [27]

Although the computation is quite cumbersome, the argument of the proof is based in a simple idea: if the distances $\|x(0) - \hat{x}\|$ and $\|x(T) - \hat{x}\|$ are small and the harvesting policy is similar to the periodic program harvesting, then the state of the forest cannot go far from $\hat{x}$ (in less than n steps) without making a large value loss in at least one step.

To see that $\|x(t) - \hat{x}\| < \epsilon$ for all $t = 1, \dots, T-1$ we start bounding the value of $x_i(t)$ for all $i = \sigma+1, \dots, n$ and after that we bound $\|x_i(t) - \frac{1}{\sigma}\|$ for all $i = 1, \dots, \sigma$.

First consider $i > \sigma$. In this case, we can express $x_i(t)$ as a linear combination of $x_{i+t-T}(T)$ or $x_n(t+n-i)$ and the harvests between t and T or $t+n-i$ (that are controlled by ϵ_1) to deduce that

$$x_i(t) < \tfrac{\epsilon}{2^n} + n\,\epsilon_1 \quad \text{for all } i > \sigma. \tag{14}$$

We need to divide the study into two cases:

1. Case $i-t+T \le n$,

$$x_i(t) = x_{i+T-t}(T) + \textstyle\sum_{j=0}^{T-t-1} c_{i+j}(t+j) < \tfrac{\epsilon}{2^n} + (T-t)\epsilon_1 < \tfrac{\epsilon}{2^n} + n\,\epsilon_1 \;<\; \epsilon$$

2. Case $i-t+T > n$,

$$x_i(t) = x_n(t+n-i) + \textstyle\sum_{j=0}^{n-i-1} c_{i+j}(t+j) = \sum_{j=0}^{n-i} c_{i+j}(t+j) \;<\; n\,\epsilon_1 \;<\; \epsilon$$

[27] $S_c = \{c \in I\!R_+^n :\ c_\sigma \in S_f \ \text{ and } \ c_i = 0 \ \text{ for all } \ i \ne \sigma\}$

To deal with the i-th age class when $i \le \sigma$ we start by proving that

$$\text{if} \quad \|x(t)-\hat{x}\| < \epsilon_2 \quad \text{then} \quad |x_i(t{+}1) - \tfrac{1}{\sigma}| < n\,\epsilon_1 + 2\,\epsilon_2 \quad \text{for all } i = 1,\dots,\sigma. \tag{15}$$

Case $2 \le i \le \sigma$,

$$|x_i(t+1) - \tfrac{1}{\sigma}| = |x_{i-1}(t) - c_{i-1}(t) - \tfrac{1}{\sigma}| \le |c_{i-1}(t)| + |x_{i-1}(t) - \tfrac{1}{\sigma}| < \epsilon_1 + \epsilon_2$$

Case $i = 1$,

$$x_1(t+1) \;=\; \textstyle\sum_{i=1}^n c_i(t) \quad\Longrightarrow\quad c_\sigma(t) \le x_1(t{+}1) \le (n-1)\,\epsilon_1 + c_\sigma(t)$$

$$\begin{aligned}
\Longrightarrow |x_1(t+1) - \tfrac{1}{\sigma}| \;&\le\; (n-1)\,\epsilon_1 + |x_\sigma(t) - x_{\sigma+1}(t{+}1) - \tfrac{1}{\sigma}| \\
&\le\; (n-1)\,\epsilon_1 + |x_\sigma(t) - \tfrac{1}{\sigma}| + |x_{\sigma+1}(t{+}1)| \;<\; (n-1)\,\epsilon_1 + 2\,\epsilon_2
\end{aligned}$$

Repeated application of (14) and (15) yields,

$$\begin{aligned}
\|x(0)-\hat{x}\| < \tfrac{\epsilon}{2^n} \quad &\Longrightarrow \quad \|x(1)-\hat{x}\| < 2\tfrac{\epsilon}{2^n} + n\,\epsilon_1 \\
&\Longrightarrow \quad \|x(2)-\hat{x}\| < 2(2\tfrac{\epsilon}{2^n} + n\,\epsilon_1) + n\,\epsilon_1 \\
&\qquad \vdots \\
&\Longrightarrow \quad \|x(T{-}1)-\hat{x}\| < 2^{T-1}\tfrac{\epsilon}{2^n} + (2^{T-1}-1)n\,\epsilon_1
\end{aligned}$$

and thus $\|x(t)-\hat{x}\| < \epsilon$ for all $t = 1,\dots,T{-}1$.

Case $T \ge n$: Corollary 5.1 states that there is γ_2 such that for every program $\{x(t)\}_{t=0}^n$ satisfying $\delta\big(x(t), x(t+1)\big) < \gamma_2$ for all $t < n$, we have $\|x(t)-\hat{x}\| < \epsilon/2^n$ for all $t = 0,\dots,\sigma$. Apply this result to the programs $\{x(t{+}i)\}_{t=0}^n$ with $i = 0,\dots,T{-}\sigma$ to conclude that $\|x(t)-\hat{x}\| < \epsilon/2^n < \epsilon$ for $t = 0,\dots,T-n+\sigma$. Afterwards, apply part 1, to conclude that during the last $n-\sigma$ steps the state also fulfills: $\|x(t)-\hat{x}\| < \epsilon$ for $t = T-n+\sigma,\dots,n$.

Take $\gamma = \min\{\gamma_1,\gamma_2\}$ and the lemma follows. ∎

Proof of Lemma 5.4: First observe that the hypothesis $\sum_{t=0}^{T-1} w(bc(t)) \ge U(x(0), x(T), 0, T) - M$ implies

$$M \ge \textstyle\sum_{t=0}^{T-1} w(bc'(t)) - w\big(bc(t)\big)$$

for any program $\{x'(t)\}$ such that $x'(0) = x(0)$ and $x'(T) = x(T)$. We can then find a bound of the accumulated value loss of $\{x(t)\}$ related to the accumulated value loss of $\{x'(t)\}$,

$$\begin{aligned}
\textstyle\sum_{t=0}^{T-1} \delta\big(x(t), x(t{+}1)\big) \;&=\; \textstyle\sum_{t=0}^{T-1} [w(\tfrac{b_\sigma}{\sigma}) - w(bc(t))] - \hat{p}\,\big(x(T) - x(0)\big) \\
&\le\; M + \textstyle\sum_{t=0}^{T-1} [w(\tfrac{b_\sigma}{\sigma}) - w(bc'(t))] - \hat{p}\,\big(x'(T) - x'(0)\big) \\
&=\; M + \textstyle\sum_{t=0}^{T-1} \delta\big(x'(t), x'(t{+}1)\big)
\end{aligned}$$

Now, taking $M \le \epsilon/2$ it suffices to prove that there is a program $\{x'(t)\}$ as above, yielding an accumulated value loss smaller than $\epsilon/2$. We build such a program to prove its existence. Given $T \ge n+\sigma$, we look for a program such that:

$$\begin{cases}
\delta\big(x'(t), x'(t{+}1)\big) < \frac{\epsilon}{2(n+\sigma)} & t = 0,\dots,\sigma{-}1 \\
x'(t) = \hat{x} & t = \sigma,\dots,T{-}n \\
\delta\big(x'(t), x'(t{+}1)\big) < \frac{\epsilon}{2(n+\sigma)} & t = T{-}n,\dots,T
\end{cases}$$

Of course, we have $\delta\big(x'(t), x'(t{+}1)\big) = 0$ for all $t = \sigma, \ldots, T{-}n{-}1$.

Let γ_1 be such that $\delta\big(x'(t), x'(t{+}1)\big) < \frac{\epsilon}{2(n+\sigma)}$ if $\|\big(x'(t), x'(t{+}1)\big) - (\hat{x}, \hat{x})\| \le \gamma_1$. Take $\gamma = \min\big\{\frac{\gamma_1}{n}, \frac{1}{n\sigma}\big\}$.

We start by building the first σ elements of the program. [28] We will see that for any $x(0)$ satisfying $\|x(0) - \hat{x}\| \le \gamma$ there is a program $\{x(t)\}_{t=0}^{\sigma}$ from $x(0)$ such that

$$\begin{cases} \|x'(t) - \hat{x}\| < \gamma_1, & t = 1, \ldots, \sigma{-}1 \\ x(\sigma) = \hat{x} \end{cases} \tag{16}$$

implying that $\delta\big(x'(t), x'(t{+}1)\big) \le \frac{\epsilon}{2(n+\sigma)}$ for all $t < \sigma$.

Given any $x(0)$ such that $\|x(0) - \hat{x}\| < \gamma$, $x(0)$ is of the form

$$x(0) = \big(\tfrac{1}{\sigma} + \phi_1, \ldots, \tfrac{1}{\sigma} + \phi_\sigma, \phi_{\sigma+1} \ldots, \phi_n\big) \quad \text{where } \textstyle\sum_i \phi_i = 0 \text{ and } |\phi_i| < \gamma \text{ for all } i \ \text{ and } \phi_i > 0 \text{ for } i > \sigma.$$

We claim that the following sequence fulfills relations (1) and (16):

$$\begin{array}{rcl} x'(0) & = & x(0) \\ & \vdots & \\ x'(t) & = & \Big(\underbrace{\tfrac{1}{\sigma}, \ldots, \tfrac{1}{\sigma}}_{t}, \underbrace{\tfrac{1}{\sigma} + \phi_1, \ldots, \tfrac{1}{\sigma} + \phi_{\sigma-t-1}}_{\sigma-t-1}, x_\sigma(t), x_{\sigma+1}(t), \underbrace{0, \ldots, 0}_{n-\sigma-1}\Big), \quad 0 < t < \sigma \\ & \vdots & \\ x'(\sigma) & = & \hat{x} \end{array}$$

where $x'_\sigma(t) = \frac{1}{\sigma} + \phi_{\sigma-t} + \min\big(0, \sum_{i=\sigma-t+1}^{n} \phi_i\big)$ and $x'_{\sigma+1}(t) = \max\big(0, \sum_{i=\sigma-t+1}^{n} \phi_i\big)$

We check easily that $\|x'(t) - \hat{x}\| < n\gamma \le \gamma_1$. It is left to see that the proposed sequence is in fact a program, to this end we prove first that $x'(t) \in \Delta$ for $t = 1, \ldots, \sigma - 1$. ¿From $\gamma \le \frac{1}{n\sigma}$ it is evident that $x'(t) \in I\!R^n_+$ for $t = 1, \ldots, \sigma - 1$, let us see that $\sum_i x'_i(t) = 1$

$$\begin{array}{rcl} \sum_i x'_i(t) & = & t\frac{1}{\sigma} + \sum_{i=1}^{\sigma-t-1}(\frac{1}{\sigma} + \phi_i) + \frac{1}{\sigma} + \phi_{\sigma-t} + \min\big(0, \sum_{i=\sigma-t+1}^{n} \phi_i\big) + \max\big(0, \sum_{i=\sigma-t+1}^{n} \phi_i\big) \\ & = & \frac{1}{\sigma}\big[t + (\sigma - t - 1) + 1\big] + \sum_{i=1}^{\sigma-t-1} \phi_i + \phi_{\sigma-t} + \sum_{i=\sigma-t+1}^{n} \phi_i \\ & = & 1 + \sum_{i=1}^{n} \phi_i = 1 \end{array}$$

The proof is completed by showing that $\lambda\big(x'(t), x'(t{+}1)\big) \ge 0$ for $t = 0, \ldots, \sigma - 1$ (see Remark 2.3).

Case: $\mathbf{t = 0}$

$$\lambda\big(x(0), x'(1)\big) = \Big(\underbrace{0, \ldots, 0}_{\sigma-2}\ , -\min\big(0, \textstyle\sum_{i=\sigma}^{n} \phi_i\big)\ ,\ \frac{1}{\sigma} + \phi_\sigma - \max\big(0, \sum_{i=\sigma}^{n} \phi_i\big)\ ,\ \underbrace{\phi_{\sigma+1} \ldots, \phi_n}_{n-\sigma}\Big)$$

Only the non-negativity of the σ-th coordinate is not evident, but it follows from the fact that $|\phi_\sigma - \max\big(0, \sum_{i=\sigma}^{n} \phi_i\big)| = |\phi_\sigma - \max\big(0, \phi_\sigma + \sum_{i=\sigma+1}^{n} \phi_i\big)| \le |\phi_\sigma| + |\sum_{i=\sigma+1}^{n} \phi_i| \le \sum_{i=\sigma}^{n} |\phi_i| \le (n{-}\sigma{+}1)\gamma \le \frac{n-\sigma+1}{n}\frac{1}{\sigma} \le \frac{1}{\sigma}$.

Case: $\mathbf{t = 1, \ldots, \sigma{-}2}$

$$\lambda\big(x'(t), x'(t{+}1)\big) = \Big(\underbrace{0, \ldots, 0}_{\sigma-2}\ ,\ c_{\sigma-1}(t)\ ,\ c'_\sigma(t)\ ,\ c'_{\sigma+1}(t)\ ,\ \underbrace{0, \ldots, 0}_{n-\sigma-1}\Big)$$

where

[28] A direct adaptation of [8, Proposition 5.5].

$$\begin{array}{rcl} c'_{\sigma-1}(t) & = & -\min\big(0, \sum_{i=\sigma-t}^{n}\phi_i\big) \ \geq \ 0 \\ c'_{\sigma}(t) & = & \frac{1}{\sigma} + \phi_{\sigma-t} + \min\big(0, \sum_{i=\sigma-t+1}^{n}\phi_i\big) - \max\big(0, \sum_{i=\sigma-t}^{n}\phi_i\big) \\ c'_{\sigma+1}(t) & = & x_{\sigma+1}(t) \ \geq \ 0 \end{array}$$

After some computations we can see that $c'_\sigma(t) \geq \frac{1}{\sigma} - |\phi_{\sigma-t}| - |\sum_{i=\sigma-t+1}^{n}\phi_i| \geq \frac{1}{\sigma} - (n-\sigma+t)\gamma > 0$. And finally,

Case: $\mathbf{t = \sigma - 1}$

$$\lambda\big(x'(\sigma-1), \hat{x}\big) = \big(\underbrace{0,\ldots,0}_{\sigma-1} \ , \ x_\sigma(\sigma-1) \ , \ x_{\sigma+1}(\sigma-1) \ , \ \underbrace{0,\ldots,0}_{n-\sigma-1}\big) \geq 0$$

To finish the proof we propose the last n elements of the sequence. [29] We look for a program satisfying

$$\begin{cases} x'(T-n) = \hat{x} \\ \|x'(t) - \hat{x}\| < \gamma_1, \quad t = T-n+1, \ldots, T-1 \\ x'(T) = x(T) \end{cases} \tag{17}$$

The point $x(T)$ can be handled in the same way as $x(0)$ to be written as $x(T) = \big(\frac{1}{\sigma} + \phi_1, \ldots, \frac{1}{\sigma} + \phi_\sigma, \phi_{\sigma+1} \ldots, \phi_n\big)$ where $\sum_i \phi_i = 0$ and $|\phi_i| < \gamma < \frac{1}{\sigma}$ for all i and $\phi_i > 0$ for $i > \sigma$. Let $\phi = \max_i\{\phi_i\} \geq 0$.

We give again a constructive proof which is somewhat involved. Given σ and n, let $k \in \mathbb{N}$ and $j \in [0, \ldots, \sigma-1]$ be such that $j = n(\sigma)$ and $n = k\sigma + j$. We know that $x(T-n) = \hat{x}$. During $n - \sigma$ stages we apply the harvesting policy given by

$$c\big(x'(t)\big) = \big(0, \ldots, 0, \underbrace{x'_\sigma(t) - \phi}_{\sigma-th\ pos}, 0, \ldots, 0, \underbrace{x'_n(t)}_{=0}\big) \qquad\qquad T-n \leq t < T-\sigma$$

generating the sequence $\{x'(t)\}_{t=T-n+1}^{T-\sigma}$,

$$\begin{array}{rcl} x'(T-n+1) & = & \big(\frac{1}{\sigma} - \phi, \frac{1}{\sigma}, \ldots, \frac{1}{\sigma}, \phi, 0, \ldots, 0\big) \\ x'(T-n+2) & = & \big(\frac{1}{\sigma} - \phi, \frac{1}{\sigma} - \phi, \frac{1}{\sigma}, \ldots, \frac{1}{\sigma}, \phi, \phi, 0, \ldots, 0\big) \\ \vdots & & \vdots \\ x'(T-\sigma) & = & \big(\underbrace{\frac{1}{\sigma} - k\phi, \ldots, \frac{1}{\sigma} - k\phi}_{j}, \underbrace{\frac{1}{\sigma} - k\phi + \phi, \ldots, \frac{1}{\sigma} - k\phi + \phi}_{\sigma-j}, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\big). \end{array}$$

Due to Remark 2.3 it is not difficult to check that the generated sequence is such that $\big(x'(t), x'(t+1)\big) \in \Omega$ for $T-n \leq t < T-\sigma$. Furthermore, $\|\hat{x} - x'(t)\| < n\gamma$ for $T-n \leq t \leq T-\sigma$.

For the next $\sigma - 1$ stages we harvest as follows,

$$\text{for } T-\sigma \leq t < T-j\,, \quad c'_i(t) = \begin{cases} k\phi & i = t + \sigma \\ x'_\sigma(t) - \phi & i = \sigma \\ x'_n(t) & i = n \qquad (x'_n(t) = \phi) \\ 0 & \text{otherwise} \end{cases}$$

$$\text{and for } T-j \leq t < T-1\,, \quad c'_i(t) = \begin{cases} k\phi + \phi & i = t + \sigma \\ x'_\sigma(t) - \phi & i = \sigma \\ x'_n(t) & i = n \qquad (x'_n(t) = \phi) \\ 0 & \text{otherwise} \end{cases}$$

[29] An adaptation of [8, Lemma 6.2].

generating the sequence $\{x'(t)\}_{t=T-\sigma+1}^{T-1}$,

$$\begin{aligned}
x'(T-\sigma+1) &= \big(\tfrac{1}{\sigma}+\phi, \tfrac{1}{\sigma}-2k\phi, \underbrace{\tfrac{1}{\sigma}-k\phi, \ldots, \tfrac{1}{\sigma}-k\phi}_{j-1}, \underbrace{\tfrac{1}{\sigma}-k\phi+\phi, \ldots, \tfrac{1}{\sigma}-k\phi+\phi}_{\sigma-j-1}, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\big) \\
x'(T-\sigma+2) &= \big(\underbrace{\tfrac{1}{\sigma}+\phi, \tfrac{1}{\sigma}+\phi}_{2}, \tfrac{1}{\sigma}-3k\phi, \underbrace{\tfrac{1}{\sigma}-k\phi, \ldots, \tfrac{1}{\sigma}-k\phi}_{j-1}, \underbrace{\tfrac{1}{\sigma}-k\phi+\phi, \ldots, \tfrac{1}{\sigma}-k\phi+\phi}_{\sigma-j-2}, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\big) \\
\vdots \quad & \quad \vdots \\
x'(T-j) &= \big(\underbrace{\tfrac{1}{\sigma}+\phi, \ldots, \tfrac{1}{\sigma}+\phi}_{\sigma-j}, \tfrac{1}{\sigma}-(\sigma-j+1)k\phi, \underbrace{\tfrac{1}{\sigma}-k\phi, \ldots, \tfrac{1}{\sigma}-k\phi}_{j-1}, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\big) \\
x'(T-j+1) &= \big(\underbrace{\tfrac{1}{\sigma}+\phi, \ldots, \tfrac{1}{\sigma}+\phi}_{\sigma-j+1}, \tfrac{1}{\sigma}-(\sigma-j+2)k\phi-\phi, \underbrace{\tfrac{1}{\sigma}-k\phi, \ldots, \tfrac{1}{\sigma}-k\phi}_{j-2}, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\big) \\
\vdots \quad & \quad \vdots \\
x'(T-1) &= \big(\underbrace{\tfrac{1}{\sigma}+\phi, \ldots, \tfrac{1}{\sigma}+\phi}_{\sigma-1}, \tfrac{1}{\sigma}+(1-n)\phi, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\big).
\end{aligned}$$

And finally harvesting

$$c'(T-1) = \big(\phi-\phi_2, \ldots, \phi-\phi_\sigma, \underbrace{\tfrac{1}{\sigma}+(1-n)\phi-\phi_{\sigma+1}}_{\geq 0\ (\sigma-th\ pos.)}, \phi-\phi_{\sigma+2}, \ldots, \phi\big).$$

we obtain $x'(T) = x(T)$. Again the fact that $c'(t) \geq 0$ for $t = T-\sigma, \ldots, T-1$ implies that the proposed sequence satisfies $\big(x'(t), x'(t+1)\big) \in \Omega$. ■

We now turn to the proof of Theorem A. The idea of the proof is to use Lemma 5.3 to bound the difference $\|x(t) - \hat{x}\|$. In general, it will not be possible to apply this lemma to the whole interval $[0, T]$. To overcome this difficulty we divide $[0, T]$ in conveniently chosen subintervals of bounded lengths, so that the lemma will be valid in all but a finite number of subintervals, where this finite number depends only on M and ϵ.

Proof of Theorem A: Given ϵ, by Lemma 5.3 there is γ such that for each $\{x(t)\}_{t=0}^{T}$ satisfying

$$\|x(0) - \hat{x}\| < \epsilon/2^n,\ \|x(T) - \hat{x}\| < \epsilon/2^n \ \text{ and } \delta\big(x(t), x(t+1)\big) < \gamma \text{ for all } \ t \in [0, \ldots, T-1] \tag{18}$$

we have

$$\|x(t) - \hat{x}\| < \epsilon \text{ for all } \ t \in [0, \ldots, T] \tag{19}$$

Given a program $\{x(t)\}$ satisfying the hypothesis and taking S, τ such that $0 \leq S \leq S + \tau \leq T$ we can use Proposition 4.5 to obtain

$$\textstyle\sum_{t=S}^{S+\tau-1} w\big(bc(t)\big) \geq U\big(x(S), x(S+\tau), S, S+\tau\big) - M \tag{20}$$

and Proposition 4.3 to deduce

$$U\big(x(S), x(S+\tau), S, S+\tau\big) - M \geq \tau w(b\hat{c}) - (n+\sigma)w(b\hat{c}) - M \tag{21}$$

From the above and Lemma 5.1 it follows that there is $\bar{\tau}$ such that

$$\text{for any } \ S \in [0, T-\bar{\tau}] \ \text{ there is } \ t \in [S, S+\bar{\tau}] \ \text{ such that } \|x(t) - \hat{x}\| < \epsilon/2^n \tag{22}$$

We next divide the interval $[0, T]$ in subintervals $[t_i, t_{i+1}]$ with $i = 0, \ldots, K$ where $t_0 = 0$, $t_K = T$ and

$$\bar{\tau} \leq (t_i - t_{i-1}) \leq 2\bar{\tau} \ \text{ and } \|x(t_i) - \hat{x}\| < \epsilon/2^n \ \text{ for all } \ i = 1, \ldots, K-1,$$

using the following algorithm: by (22) there is $t_1 \in [\bar{\tau}, 2\bar{\tau}]$ such that $\|x(t_1) - \hat{x}\| < \epsilon/2^n$. Using (22) again we know that there exists $t_2 \in [t_1 + \bar{\tau}, \dots, t_1 + 2\bar{\tau}]$ such that $\|x(t_2) - \hat{x}\| < \epsilon/2^n$. We proceed inductively defining

$$t_{i+1} \in [t_i + \bar{\tau}, t_i + 2\bar{\tau}] \text{ with } \|x(t_{i+1}) - \hat{x}\| < \epsilon/2^n$$

We repeat this step until we obtain $(t_{K-1}+2\bar{\tau}) \geq T$, then we set $t_K = T$ and the construction of the sequence is finished.

For every $i = 1, \dots, K-2$, we can apply Lemma 5.3 whenever

$$\textstyle\sum_{t=t_i}^{t_{i+1}-1} \delta\big(x(t), x(t+1)\big) < \gamma \tag{23}$$

to affirm that $\|x(t) - \hat{x}\| < \epsilon$ for all $t \in [t_i, t_{i+1} - 1]$. We claim that there are $k \leq 2 + \gamma^{-1}\big[(n+\sigma)w(b\hat{c}) + n\frac{b_\sigma}{\sigma}\, z + M\big]$ subintervals *not* fulfilling (23). Indeed, denote by $\mathcal{K} \subseteq [1, \dots, K-2]$ the set of indexes such that $\sum_{t=t_i}^{t_{i+1}-1} \delta\big(x(t), x(t+1)\big) \geq \gamma$, it is easily seen that

$$\begin{aligned} \textstyle\sum_{t=0}^{T-1} \delta\big(x(t), x(t+1)\big) &= \textstyle\sum_{k=0}^{K-1} \sum_{t=t_i}^{t_{i+1}-1} \delta\big(x(t), x(t+1)\big) \\ &\geq \textstyle\sum_{k\in\mathcal{K}} \sum_{t=t_i}^{t_{i+1}-1} \delta\big(x(t), x(t+1)\big) \geq \gamma \ \mathrm{Card}\{\mathcal{K}\}. \end{aligned}$$

On the other hand we know that

$$\begin{aligned} \textstyle\sum_{t=0}^{T-1} \delta\big(x(t), x(t+1)\big) &= \textstyle\sum_{t=0}^{T-1} [w(b\hat{c}) - w\big(bc(t)\big)] + \hat{p}(x(0) - x(T)) \\ &\leq Tw(b\hat{c}) - U(x(0), x(T), 0, T) + \hat{p}(x(0) - x(T)) + M \\ &\leq (n + \sigma)w(b\hat{c}) + n\tfrac{b_\sigma}{\sigma} z + M \end{aligned}$$

Combining the last two inequalities we get $\mathrm{Card}\{\mathcal{K}\} \leq \gamma^{-1}[(n + \sigma)w(b\hat{c}) + n\frac{b_\sigma}{\sigma} z]$ and it follows that

$$\mathrm{Card}\big\{t = [0, \dots, T] \text{ such that } \|x(t) - \hat{x}\| > \epsilon\big\} \leq 2\bar{\tau}\left\{2 + \gamma^{-1}\big[(n+\sigma)w(b\hat{c}) + n\tfrac{b_\sigma}{\sigma}\, z + M\big]\right\}.$$

Set $L = 2\bar{\tau}\left\{2 + \gamma^{-1}\big[(n+\sigma)w(b\hat{c}) + n\frac{b_\sigma}{\sigma}\, z + M\big]\right\}$ and the theorem follows. ∎

Theorem B aims for a stronger condition than Theorem A: not only $\|x(t) - \hat{x}\| < \epsilon$ must hold for most of the time stages, but these time stages must be consecutive, i.e., violations to the condition $\|x(t) - \hat{x}\| < \epsilon$ (if any) can only occur during the initial time stages or the last ones. We use again Lemma 5.3 to bound the difference $\|x(t) - \hat{x}\|$, but to apply it to an interval almost as large as $[0, T]$ we have to pay the price that we cannot choose the parameter M but that its value will be determined in keeping with the needs of the proof.

Proof of Theorem B: By Lemma 5.3 we know that given $\epsilon > 0$ there is $\epsilon_1 > 0$ such that for any program $\{x(t)\}_{t=\tau_1}^{\tau_2}$ satisfying

$$\|x(\tau_1) - \hat{x}\| \leq \epsilon/2^n, \ \|x(\tau_2) - \hat{x}\| \leq \epsilon/2^n, \ \text{ and } \ \delta\big(x(t), x(t+1)\big) \leq \epsilon_1 \ \text{ for all } \ t = \tau_1, \dots, \tau_2 - 1 \tag{24}$$

the following inequality holds

$$\|x(t) - \hat{x}\| \leq \epsilon \ \text{ for all } \ t = \tau_1, \dots, \tau_2 \tag{25}$$

In order to bound $\delta\big(x(t), x(t+1)\big)$ we use Lemma 5.4 which states that given $\epsilon_1 > 0$ there are $\gamma > 0$ and $M > 0$ such that for every τ_1 and τ_2 (satisfying $\tau_2 - \tau_1 \geq n+\sigma$) and every program fulfilling

$$\|x(\tau_1) - \hat{x}\| \leq \gamma, \ \|x(\tau_2) - \hat{x}\| \leq \gamma, \ \text{ and } \ \textstyle\sum_{t=\tau_1}^{\tau_2-1} w\big(bc(t)\big) \geq U\big(x(\tau_1), x(\tau_2), \tau_1, \tau_2\big) - M \tag{26}$$

we have that $\sum_{t=\tau_1}^{\tau_2-1}\delta\big(x(t),x(t+1)\big) \leq \epsilon_1$, implying directly that $\delta\big(x(t),x(t+1)\big) \leq \epsilon_1$ for all $t=\tau_1,\dots,\tau_2-1$.

We proceed now to prove the existence of

$$\tau_1 \in [0,\tau] \text{ and } \tau_2 \in [T-\tau, T] \text{ such that } \|x(\tau_i)-\hat{x}\| < \gamma. \tag{27}$$

Let $M_1 = M+(n+\sigma)w(b\hat{c})$, Lemma 5.1 states that there is τ such that if the program $\{x(t)\}_{t=0}^{\tau}$ satisfies

$$\sum_{t=0}^{\tau-1} w\big(bc(t)\big) \geq \tau w(b\hat{c}) - M_1 \quad \text{then there is} \quad t\in[0,\tau] \quad \text{such that} \quad \|x(t)-\hat{x}\| < \gamma \tag{28}$$

By propositions 4.3 and 4.5 we know that

$$\sum_{t=0}^{\tau-1} w\big(bc(t)\big) \geq U\big(x(0),x(\tau),0,\tau\big) - M \geq \tau w(b\hat{c}) - (n+\sigma)w(b\hat{c}) - M$$
$$\sum_{t=T-\tau}^{T-1} w\big(bc(t)\big) \geq U\big(x(t-\tau),x(T),T-\tau,T\big) - M \geq \tau w(b\hat{c}) - (n+\sigma)w(b\hat{c}) - M$$

and the existence of $\tau_1 \in [0,\tau]$ and $\tau_2 \in [T-\tau,T]$ such that $\|x(\tau_i)-\hat{x}\| < \gamma$ is assured.

Now Proposition 4.5 tells us that for any program fulfilling the hypothesis, and for any $0 \leq \tau_1 < \tau_2 \leq T$, $\sum_{\tau_1}^{\tau_2-1} w\big(bc(t)\big) \geq U\big(x(\tau_1),x(\tau_2),\tau_1,\tau_2\big) - M$. This fact together with (27) proves that (26) holds if $T \geq 2\tau+n+\sigma$ and then $\delta\big(x(t),x(t+1)\big) \leq \epsilon_1$ for all $t=\tau_1,\dots,\tau_2-1$. Finally, this proves that (24) holds and we conclude as desired. ■

References

[1] Bewley, T. F., 1982. On the integration of equilibrium theory and turnpike theory. Journal of Mathematical Economics 10, 233-267.

[2] Dorfman, R., Samuelson, P. A., Solow, R. M., 1958. Linear Programming and Economic Analysis. McGraw Hill, New York.

[3] Gale, D., 1967. On optimal development in a multi-sector economy. Review of Economic Studies, 34, 1-18.

[4] Gale, D., 1970. Nonlinear duality and qualitative properties of optimal growth. In Abadie, J. (ed.), Integer and Nonlinear Prgramming. North Holland, Amsterdam, 309-319.

[5] Hicks, J. R. 1965. Capital and Growth. Oxford University Press, Oxford.

[6] Kant, S. and Berry, R. A. (eds.) 2005. Economics, Sustainability, and Natural Resources: Economics of Sustainable Forest Management, Volumes 1 and 2. Springer, Berlin.

[7] Khan, M. Ali, 2005. Intertemporal ethics, modern capital theory and the economics of forestry. Chapter 2, pp. 39-65 in [6].

[8] Khan M. Ali and Piazza A., 2008. On uniform convergence of undiscounted optimal programs in the Mitra-Wan forestry model: the strictly concave case. *Publicación Técnica CMM.* No. 222, 1-15.

[9] Khan M. Ali and Piazza A., 2009. On the Mitra-Wan Forestry Model: A Unified Analysis. *Publicación Técnica CMM.* No. 226. Paper presented at *Workshop in Economic Dynamics,* National University of Singapore, July 30 to August 1, 2009.

[10] Khan M. Ali and Zaslavski, A. J., 2008. On two classical turnpike results for the Robinson-Solow-Srinivasan (RSS) model. Advances in Mathematical Economics, forthcoming.

[11] Khan M. Ali and Zaslavski, A. J., 2009. On Locally Optimal Programs in the RSS (Robinson-Solow-Srinivasan) Model. Mimeographed, University of Queensland.

[12] McKenzie, L. W., 1982. A primal route to the turnpike and Lyapunov stability. Journal of Economic Theory 27, 194-209.

[13] McKenzie, L. W., 1998. Turnpikes. American Economic Review (Papers and Proceedings) 88, 1-14.

[14] Mitra T. and Wan H. W., 1985. Some theoretical results on the economics of forestry. Review of Economics Studies LII, 263-282.

[15] Mitra T. and Wan H. W., 1986. On the Faustmann solution to the forest management problem. Journal of Economic Theory 40, 229-249.

[16] Mitra, T., 2000. Introduction to dynamic optimization theory, in M. Majumdar et al. 2000, Optimization and Chaos, 31-108.

[17] Mitra, T., 2005. Intergenerational Equity and the forest management problem. Chapter 5, 137-173 in [6].

[18] Mitra, T., 2006. Notes on a Model of Forestry: the Undiscounted Case, mimeo, Cornell University.

[19] Radner, R., 1961. Paths of economic growth that are optimal only with respect to final states. Review of Economic Studies 28, 98-104.

[20] Rockafellar, R. T., 1970. Convex Analysis. Princeton University Press, Princeton. Reprint: 1997.

[21] Salo S. and Tahvonen O., 2002. On equilibrium cycles and normal forests in optimal harvesting of tree vintages. Journal of Environmental Economics and Management 44, 1-22.

[22] Salo S. and Tahvonen O., 2003. On the economics of forest vintages. Journal of Economics Dynamics and Control 27, 1411-1435.

[23] Samuelson, P. A., 1949. Market mechanisms and maximization: dynamics and linear programming. Rand Corporation, Santa Monica. Reprinted in Stiglitz, J. E. (ed.), The Collected Scientific Papers of Paul A. Samuelson. Vol. II. MIT Press, Cambridge.

[24] Samuelson, P. A., 1960. Efficient paths of capital accumulation in terms of the calculus of variations. In Arrow, K. J., Karlin, S., Suppes, P. (eds.), Mathematical Methods inn the Social Sciences. Stanford University Press, Palo Alto.

[25] Samuelson, P. A., 1965. A catenary turnpike theorem involving consumption and the golden rule. American Economic Review 55, 486-496.

[26] Samuelson P. A., 1971. Maximum principles in analytical economics. Les Prix Nobel en 1970. The Nobel Foundation, Stockholm, 273-288.

[27] Samuelson P. A., 1972. The Collected Scientific Papers of Paul A. Samuelson. Merton, R. C. (ed.), Volume III. MIT Press, Cambridge.

[28] Samuelson P. A., 1976. Economics of forestry in an evolving society. Economic Inquiry XIV, 466-492.

[29] Samuelson P. A., 1976. The periodic turnpike theorem. Nonlinear Analysis, Theory, Methods and Applications 1, 1-13.

[30] Samuelson P. A., 1977. The Collected Scientific Papers of Paul A. Samuelson. Crowley, K., Nagatani, H. (eds.), Volume IV. MIT Press, Cambridge.

[31] Samuelson, P. A., Solow, R. M., 1956. A complete capital model involving heterogeneous capital goods. Quarterly Journal of Economics LXX, 537-562.