Nonlinear Analysis 74 (2011) 171–181

Contents lists available at ScienceDirect

Nonlinear Analysis

journal homepage: www.elsevier.com/locate/na

The economics of forestry and a set-valued turnpike of the classical type ☆

M. Ali Khan [a], Adriana Piazza [b,*]

[a] Department of Economics, The Johns Hopkins University, Baltimore, MD 21218, United States

[b] Departamento de Matemática, Universidad Técnica Federico Santa María, Avda. España 1680, Casilla 110-V, Valparaíso, Chile

ARTICLE INFO

Article history:
Received 29 March 2010
Accepted 19 August 2010

Keywords:
Periodic turnpike
Set-valued turnpike
Non-interiority
Non-differentiability
Good program
Approximately optimal program
Large but finite time horizon
Asymptotic convergence
Forest management

ABSTRACT

In recent work, the authors set classical turnpike theory in the context of the economics of forestry, as developed by Mitra and Wan, and presented two far-reaching results. In this paper, we present a conceptual generalization that takes this theory and configures it around a set in the space of forest configurations rather than around the *golden-rule* forest configuration. Our set-valued analysis hinges on periodicity and yields the earlier results as corollaries under a non-interiority condition on the felicity function that shrinks the set to the point. The question that we pose, and answer, has obvious relevance to more general contexts and, in particular, to turnpike theory as developed by Samuelson, Gale, McKenzie, and their followers.

1. Introduction

In the first issue of the journal *Nonlinear Analysis*, Paul Samuelson [1] presented a periodic turnpike theorem. From the perspective of classical turnpike theory, as recently delineated in [2,3], it is exactly the type of theorem that one would expect and want. Given initial and final capital stocks over a sufficiently large but finite horizon, an optimal intertemporal resource allocation program stays arbitrary close, most of the time, to the solution of an infinite horizon optimal intertemporal resource allocation program, even when the felicity function is subject to periodic oscillations, which is to say, even when the relevant turnpike is periodic. As explained in [4], and subsequently in [5,2], given the differing time horizons, and therefore differing programming problems that are involved, the question is more subtle than that of the asymptotic convergence of the solution of an infinite horizon variational problem, or of the continuity of its solution with respect to initial stocks. Periodicity, as in [1], only adds to the complexity of the question, though the question is now being phrased in the non-stationary register rather than the classical stationary one.

It is of interest that in the same year that he published his periodic turnpike theorem, Samuelson [6] turned his attention to the economics of forestry, and asked whether a profit maximizing firm would be led by market conditions to produce maximal long-run sustainable timber yields. The registers of the two enquiries were different: the first was in the context of a Ramseyian planner [7] concerned with long-term societal interest whereas the second concerned competition and the price theory of the firm. A decade was to pass before Mitra and Wan [8,9] reconfigured Samuelson's forestry problem

☆ This work was initiated during Piazza's visit to the University of Illinois at Urbana-Champaign in April–May, 2009, and continued when Khan held the position of Visiting Professor at the University of Queensland, June–July 2009, and at the Nanyang Technological University. The authors thank Flavio Menezes, Euston Quah, Anne Villamil and Nicholas Yannelis, and the individual departments, for their hospitality. Adriana Piazza gratefully acknowledges the financial support of FONDECYT under Project 11090254, that of Programa Basal PFB 03, CMM, U. de Chile and Project ANILLO ACT-88.

* Corresponding author. Fax: +56 32 2654 563.
E-mail addresses: akhan@jhu.edu (M.A. Khan), adriana.piazza@usm.cl (A. Piazza).

0362-546X/$ – see front matter
doi:10.1016/j.na.2010.08.031

into a Ramseyian planning exercise, and recast it into what has subsequently come to be known as the Gale–McKenzie theory of intertemporal resource allocation [10,11]. In particular, they showed that in a setting with linear felicity functions, periodicity is the rule rather than the exception, and that optimal solutions to the undiscounted infinite horizon problem oscillate around the forest configuration producing maximal sustainable timber yield irrespective of the given initial configuration. It is only in the case of strictly concave undiscounted felicity functions that we obtain convergence to this maximal sustainable timber yield forest configuration. And so, as a first step, it is natural to limit oneself to the strictly concave case, and ask whether there is a viable and robust turnpike theory that can be constructed. This is to ask whether for arbitrary given initial and final forest configurations, optimal forest management would dictate staying arbitrarily near the golden-rule forest configuration for most of a large but finite planning horizon. This question has recently been answered in [12,3].

However, given Samuelson's work, it is surely worthwhile to take another look at the case of linear felicities. Here, there is a continuum of periodic optimal paths, one for each initial forest configuration, and there is no question of a turnpike theory of the classical type. Indeed, there is no single candidate for possible service as a turnpike for large but finite optimal programs starting from an arbitrary initial condition. However, one can proceed by taking the given initial configuration x_0 as a datum of the analysis, focus on the infinite horizon periodic optimal program that starts from it, and ask whether this program is a turnpike for all programs starting from x_0 and ending at $x_T \neq x_0$. In other words, does the T-period program starting from x_0 and ending at x_T for arbitrarily large T get onto the infinite horizon program also starting from x_0, and stay on it until the last few periods to end at x_T? To be sure, this is not a question within the ambit of what is seen as classical turnpike theory in [2]; it is rather a version of what McKenzie [13] terms the *early* turnpike. (He refers to asymptotic convergence of optimal programs as the *late* turnpike.) However, there is an antecedent question that can be asked in the light of the fact that the two theorems presented in [3] do somewhat more than simply provide a classical turnpike theory for the Mitra–Wan tree farm in the strictly concave case; they bypass the restricting bipolar dichotomy of linear and strictly concave functions, and consider instead a setting where the felicity functions are assumed only to be concave.

The point of departure of this paper lies in the unification theorem of [14]. In the context of non-differentiable, upper semi-continuous functions that are supportable at the golden-rule stock, they present a *condition on the primitives of the model* which is necessary and sufficient for all good programs to converge to this stationary stock, and thereby allow a robust turnpike theory to be constructed. This condition can be simply stated: it is the requirement that the solution to a finite dimensional optimization problem be the singleton golden-rule stock. An easier to check *non-interiority condition* proves to be sufficient but not necessary: the golden-rule stock is not an interior point of the convex hull of the set of zeros of the *discrepancy function* corresponding to the (stationary) felicity function, which is to say, the function of the difference between the felicity and its linear affine supporting function at the golden-rule stock. In the particular case of a concave felicity function, the *non-interiority condition* becomes also necessary. This then leads immediately to the situation where the necessary and sufficient condition does *not* hold and to the observation that in such a setting, the solution to the optimization problem mentioned above, say V, is exactly the set of initial configurations with periodic optimal programs—and indeed, depending on the felicity function, can become arbitrarily small. The question then naturally arises as to the characteristics of the optimal programs that do not start from, or end at, configurations in this set V. This question has not been posed in the literature before, and we answer it here.

The conceptual innovation pursued in this paper then needs re-emphasis. It is to construct a set-valued turnpike theory of the classical type. Rather than an investigation of a particular periodic program as a possible candidate for a turnpike, as in [1], it is to bunch the entire continuum of periodic programs together, and investigate the entire set V as a possible candidate for a turnpike. Rather than the metaphor of a freeway or turnpike, as originally described in [4] and quoted in [13], the relevant metaphor here is that of an air-lane or sea-lane that takes on journeys that are long enough, even though it may not be the most direct route; and somewhat more relevantly here, even though within the lane, there are many possible routes and which particular route is taken on one occasion is not the most relevant consideration. From a technical point of view, it is then to substitute a set for a point, and to obtain a non-trivial generalization of the theory that reduces to the standard one when the set V shrinks to a point and the non-interiority condition is automatically activated.

We have already clarified that even though inspired by Samuelson's periodic turnpike theorem, ours is not the same as his. Our set-valued turnpike, in keeping with the stationary assumptions of our model, is not periodic, even though individual optimal paths within it are. It is also worth noting how our *set-valued* turnpike theorem differs from Keeler's *twisted* turnpike theorem, as in [15,16], and from McKenzie's *neighborhood* turnpike theorem, as in [17,18]. Keeler's results can best be seen as antecedent to those of Samuelson's: he works with the original turnpike conception without any consumption and in which labor is a produced input, one rooted in the von Neumann model [19], and allows for non-stationarity in the technology rather than in the felicities, and a non-stationarity of a more general type than periodicity. As emphasized above, the primary interest in the results that we report possibly hinges on the fact that they pertain to a stationary environment. McKenzie's neighborhood turnpike theorem also pertains to a stationary environment, and for the canonical model that goes beyond that of [19] or [4], but it also has a nature totally different to that of ours. His result lies on the interface between the discounted and undiscounted settings, rather than in the undiscounted setting, as for the case that we consider here. For any arbitrarily small neighborhood of the golden-rule stock, his theorem shows the existence of a threshold discount factor such that for all discount factors greater than it, and for any initial capital stock, the program optimal for the chosen discount factor eventually lies in the neighborhood. If one adheres to the formulation of classical turnpike theory as delineated in [2], this is really a type of uniform asymptotic convergence theorem, a late turnpike type of result, with the uniformity, and the resulting set of turnpikes, revolving on the different discount factors that lie within a specified range. But leaving issues of classification and semantics aside, in delineating how our result differs substantively and technically from others presented in the literature, we simultaneously highlight directions in which it can find further elaboration and extension.

The plan of the remainder of the paper is now straightforward. In Section 2, we present the basic outlines of the forestry model, and the benchmark results. Section 3 presents the main result of the paper, whose proof is given in Section 5, while Section 4 provides some subsidiary results with the corresponding proofs. Also, in Section 4, with the notation and concepts at hand, we also indicate the principal technical contribution of the paper.

2. The model and benchmark results

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 $\Delta = \left\{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\right\}$. 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 supremum norm of x by $\|x\|_\infty$.

In addition to its original formulation [8,9], an outline of the Mitra–Wan forestry model is also available in [20]. Here we depart from the original specification and work with the reformulation presented in [21] and pursued in [22,12,14,3,23]. Under this specification, the model consists simply of the pair (b, w), where b is a non-negative vector of timber 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 1 to n cover completely a homogeneous plot of land of normalized unit size.

Note that we do not use their timber-content function $f(\cdot)$ introduced in [8,9], and make no assumptions on the timber coefficients other than the following Brock–Mitra–Wan uniqueness condition.

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

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 loses its economic value. However, one difference should be noted. In their treatment, Mitra and Wan take N to be the age at which the timber 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".[1] 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 assuming n to be the age at which a tree dies, and point out that the technicalities of the analysis do not change with this augmentation of the state's dimension.

For each period $t \in \mathbb{N}$ we denote by $x_i(t) \geq 0,\ i = 1, \ldots, n$, 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 from 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

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

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

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

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,

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

[1] See [9, p. 232]. The same point is made in [20, Section 4, Paragraph 5].

Definition 2.3. The vector of harvests needed to perform this transition is given by the function $\lambda : \Omega \to \mathbb{R}^n_+$,

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

In addition, it is easy to see that

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

The preferences of the planner are represented by a felicity function, $w : [0, \infty) \to \mathbb{R}$, which is assumed to be non-decreasing and upper semi-continuous. Define for any $(x, x') \in \Omega$ the function $u(x, x')$ as

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

We also assume that the function w is supported at the point (b_σ/σ) by a strictly increasing affine function, which is to say that there is at least one $z > 0$ satisfying

$$w(y) \leq w\left(\frac{b_\sigma}{\sigma}\right) + z\left(y - \frac{b_\sigma}{\sigma}\right) \quad \text{for all } y \in \mathbb{R}_+. \tag{2}$$

We now turn to the specification of the golden-rule forest configuration:

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

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

We now present some basic antecedent results; except those indicated, they are all taken from [14].

Theorem 2.1. *There exists a unique golden-rule stock* $\hat{x} = \left(\underbrace{\frac{1}{\sigma}, \ldots, c, \frac{1}{\sigma}}_{\sigma}, 0, \ldots, 0\right)$.

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 if the forest remains at the golden-rule stock is $b\hat{c} = \frac{b_\sigma}{\sigma}$.

Set $\hat{p} \in \mathbb{R}^n_+$, $\hat{p} = z\frac{b_\sigma}{\sigma}(1, 2, \ldots, n)$, where z is the slope of the linear affine function defined in (2). Next, we define the function $\delta : \Omega \longrightarrow \mathbb{R}$.

Definition 2.5. The *value loss* associated with any $(x, x') \in \Omega$ is given by

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

It is easy to see that the function $\delta(\cdot, \cdot)$ is lower semi-continuous and the following lemma proves that $\delta(x, x') \geq 0$ for any $(x, x') \in \Omega$. The non-negativity of the value loss function was first established in [9, Lemma 3.1] when w is differentiable and concave.

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

$$\delta(x, x') \geq z\left[\sum_{i=1}^{n-1}\left(\frac{b_\sigma}{\sigma} - \frac{b_i}{i}\right)i(x_i - x'_{i+1}) + \left(\frac{b_\sigma}{\sigma} - \frac{b_n}{n}\right)nx_n\right] \geq 0.$$

We use the following notion of *good* and *bad* programs introduced by Gale [10].

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

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

In the sequel, and when no confusion can arise, we will write simply $\delta(t)$ instead of $\delta\left(x(t), x(t+1)\right)$.

Let $x_0 \in \Delta$. Set $\mu(x_0) = \inf\left\{\sum_{t=0}^{\infty} \delta(t) : \{x(t)\} \text{ is a program from } x_0\right\}$.
It is shown in [12, Remark 2.2] 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 exists a good program $\{x(t)\}$ such that*

$$\sum_{t=0}^{\infty} \delta(t) = \mu(x_0).$$

The fact that every good program converges to the golden-rule stock in the case where w is strictly concave or differentiable was established in [9, Lemma 6.4]. In [14], Khan and Piazza provide a necessary and sufficient condition for assuring the convergence of every good program to the golden-rule stock for any upper semi-continuous utility function w supported at the golden-rule consumption that is not necessarily either concave or differentiable. We describe this characterization in the following terms.

Let the *discrepancy function* f be

$$f(\xi) = w\left(\frac{b_\sigma}{\sigma}\right) - w(b_\sigma \xi) + z\left(b_\sigma \xi - \frac{b_\sigma}{\sigma}\right).$$

We can appeal to (2) to assert that $f(\xi) \geq 0$ for all ξ and due the upper semi-continuity of w, f attains its minimum in a closed set S_f, which contains $1/\sigma$.

Next, we define the following subsets of $\mathbb{R}^n_+$:

$$S_c = \{c \in \mathbb{R}^n_+ : c_\sigma \in S_f \text{ and } c_i = 0 \text{ for all } i \neq \sigma\}$$
$$V = \{x \in \Delta : x_i \in S_f \text{ for all } i \leq \sigma \text{ and } x_i = 0 \text{ for all } i > \sigma\}.$$

Khan and Piazza [14] provide some properties of the set V: (i) $\hat{x} \in V$, (ii) every good program converges asymptotically to V (*cf.* Lemma 2.2), (iii) every initial configuration inside V yields a σ-periodic, zero-value loss program and there are no zero-value loss programs originating outside of this set (*cf.* Remark 2.1) and (iv) it is exactly the solution set of the finite dimensional optimization problem

$$(P)\quad \begin{cases} \alpha = \min \displaystyle\sum_{i=1}^{\sigma} w(b_\sigma/\sigma) - w(b_\sigma x_i) \\ \qquad \text{s.t. } x \in \Delta. \end{cases}$$

Hence, it is of interest to know when

Condition 2.1 (*Outstanding Condition*)**.** $V = \{\hat{x}\}$.

This can of course be determined by the explicit resolution of (P). An easier to check sufficient condition, also provided in [14], is

Condition 2.2 (*Non-interiority*)**.** $(1/\sigma) \notin$ *int co* S_f.

Of course, the non-interiority Condition 2.2 is assured if w is strictly concave, but there is a broader family of functions satisfying it. In the particular case of a concave felicity function, we get that co $S_f = S_f$ and Condition 2.2 becomes also necessary.

As discussed in [14], the following results are obtained without Condition 2.1.

Proposition 2.3. *The von Neumann facet is*

$$\left\{(x, x') \in \Omega : \delta(x, x') = 0\right\} = \left\{(x, x') \in \Omega : \lambda(x, x') \in S_c\right\}.$$

Remark 2.1. Given $x \in V$, consider the σ-periodic program from x where the harvest consists 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. *Every good program $\{x(t)\}$ is such that* $\text{dist}(x(t), V) \to 0$.

Next, we present the optimality criteria that 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

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

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

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

If 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 $\left(\sum_t^{\delta}(t) = \mu(x_0)\right)$, 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.*

Let us spell out two preliminary results that we have when Condition 2.1 holds. First, we present a stronger version of Lemma 2.2.

Lemma 2.3. *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 Condition 2.1 holds, the following conditions are equivalent:*

i. $\{x(t)\}$ *is optimal.*
ii. $\sum_{t=0}^{\infty} \delta(t) = \mu(x_0)$.
iii. $\{x(t)\}$ *is maximal.*

3. Main result

We introduce notation for the aggregate value of finite optimal programs. Let $z_0, z_f \in \Delta$ and $T \geq 1$;

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

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

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 convention that $U(z_0, z_f, 0, T) = -\infty$.

Theorem 3.1. *Given $M > 0$ and $\epsilon > 0$ there exists $L \in \mathbb{N}$ such that for all $T > L$ and each program $\{x(t)\}_{t=0}^{T}$ satisfying*

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

we have that

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

Furthermore, if Condition 2.1 holds,

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

4. Substantive previous results

In this section we present the technical arguments needed to prove the principal result of this work. They were first presented in [3] in the context of a concave felicity function. We only present those proofs where a major revision was needed to adapt them to the weaker assumptions of this framework.

The three propositions presented here develop intuition into the basic dynamics underlying the MW model, and even though the statements are notationally somewhat complex, the essential ideas are simple. The concavity of the felicity function does not play a fundamental role in their proofs; hence the reader is referred to [3] for the proofs.

Proposition 4.1. *Given $z_0, z_f \in \Delta$ and $T \geq n$ we have*

$$U(z_0, z_f, 0, T) \geq Tw(b\hat{c}) - (n+\sigma)w(b\hat{c}). \tag{3}$$

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 (3) also holds.

The following is a simple inequality due to the fact that the value loss of any production plan is non-negative.

Proposition 4.2. *For every T and every program $\{x(t)\}_{t=0}^{T}$ the following inequality is satisfied:*

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

Our final proposition 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 [2, Proposition 6.6].

Proposition 4.3. *Let $T \in \mathbb{N}, M > 0, z_0, z_f \in \Delta$ and $\{x(t)\}_{t=0}^{T}$ be a program such that $x(0) = z_0, x(T) = z_f$ and $\sum_{t=0}^{T-1} w(bc(t)) \geq U(z_0, z_f, 0, T) - M$. Then for all S_1 and S_2, $0 \leq S_1 < S_2 < T$, we have*

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

Lemma 4.1. *Given $M > 0$ and $\epsilon > 0$, there exists $\tau \in \mathbb{N}$ such that for each program $\{x(t)\}_{t=0}^{\tau}$ satisfying*

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

there exists $t \in [0, \tau]$ such that $\mathrm{dist}(V, x(t)) \leq \epsilon$. Furthermore, if Condition 2.1 holds then $\|x(t) - \hat{x}\| \leq \epsilon$.

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

$$\mathrm{dist}(x^k(t), V) > \epsilon \quad \text{and} \quad \sum_{t=0}^{k-1} w\left(bc^k(t)\right) \geq kw\left(\frac{b_\sigma}{\sigma}\right) - M. \tag{4}$$

Let $M' = n\frac{b_\sigma}{\sigma}z$; by Proposition 4.2 we know that every program fulfills $\sum_{t=0}^{T-1} w(bc(t)) - w\left(\frac{b_\sigma}{\sigma}\right) \leq M'$.
Given any $s < k$, by combining the two previous inequalities we deduce

$$\sum_{t=0}^{s-1}\left[w\left(bc^k(t)\right) - w\left(\frac{b_\sigma}{\sigma}\right)\right] = \sum_{t=0}^{k-1}\left[w\left(bc^k(t)\right) - w\left(\frac{b_\sigma}{\sigma}\right)\right] - \sum_{t=s}^{k-1}\left[w\left(bc^k(t)\right) - w\left(\frac{b_\sigma}{\sigma}\right)\right] \geq -(M+M'). \tag{5}$$

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 $\{x^*(t)\}_{t\in\mathbb{N}}$ such that

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

It is not difficult to see that $\{x^*(t)\}_{t\in\mathbb{N}}$ is a program. From (5) we deduce that for every natural number s, $\sum_{t=0}^{s-1} w(bc^*(t)) - sw\left(\frac{b_\sigma}{\sigma}\right) \geq -M - M'$, meaning that $\{x^*(t)\}_{t\in\mathbb{N}}$ is a good program. Then Lemma 2.2 implies that

$$\mathrm{dist}(x^*(t), V) \to 0 \quad \text{when } t \to \infty.$$

On the other hand, it follows from (4) and the definition of $x^*(t)$ that

$$\mathrm{dist}(x^*(t), V) > \epsilon \quad \text{for all } t$$

and the contradiction proves the first assertion of the lemma. If Condition 2.1 holds, the conclusion follows directly. □

Lemma 4.2. *Let $\{x(t)\}_{t=0}^{n}$ be such that*

$$(x(t), x(t+1)) \in \Omega \quad \text{and} \quad \delta(t) = 0 \quad \text{for } t = 0, \ldots, n-1.$$

Then

$$x(t) \in V \quad \text{for all } t \in [0, \sigma].$$

Proof. By Proposition 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 and the fact that $c(t) \in S_c$ for all $t = 0, \ldots, c, \sigma - 1$ we know that

$$x(\sigma) = (c_\sigma(\sigma - 1), c_\sigma(\sigma - 2), \ldots, c_\sigma(0), 0, \ldots, 0) \in V.$$

Finally, it is easy to see that $x(t + 1) \in V$ and $\delta(t) = 0$ imply $x(t) \in V$ and then the proposition follows by backwards induction. □

Corollary 4.1. *Given $\epsilon > 0$, there exists a $\gamma > 0$ such that for each program $\{x(t)\}_{t=0}^n$ satisfying $\delta(t) < \gamma$ for $t = 0, \ldots, n - 1$ we have* $\operatorname{dist}(x(t), V) < \epsilon$ *for all $t = 0, \ldots, \sigma$.*

Proof. Consider the extended value loss function $\bar{\delta} : \Delta \times \Delta \to \mathbb{R}_+ \cup \{+\infty\}$,

$$\bar{\delta}(x, x') = \begin{cases} \delta(x, x') & \text{if } (x, x') \in \Omega \\ +\infty & \text{otherwise} \end{cases}$$

and the following optimization problem:

$$(P) \quad \begin{cases} \alpha = \min \displaystyle\sum_{t=0}^{n-1} \bar{\delta}(t) \\ \quad\quad \text{s.t. } \max\{\operatorname{dist}(x(t), V) : t = 0, \ldots, \sigma\} \ge \epsilon'. \end{cases}$$

The domain of the objective function, Δ^n, is a compact set and so it is the closed subset defined by the constraint of (P). Take $\epsilon' \in (0, \epsilon]$ small enough to assure that the feasible set is non-empty, i.e., ϵ' small enough to have $\Delta \cap \{x : \operatorname{dist}(x, V) \ge \epsilon'\} \ne \phi$, and that there is a program fulfilling the constraint (to make sure that the objective function is not identically equal to ∞ when restricted to a feasible set).

Hence, a straightforward application of the Weierstrass theorem yields that the optimal value α is finite and the solution set is not empty. Furthermore, α is strictly positive as we have deliberately left out of the feasible set the points where the sum of the value losses is zero.

Let $\gamma = \alpha/n$. If $\{x(t)\}_{t=0}^n$ is a program such that $\delta(t) < \gamma$, then $\sum_{t=0}^n \bar{\delta}(t) < \alpha$ and hence $\{x(t)\}$ does not fulfill the constraint of (P) and finally $\operatorname{dist}(x(t), V) < \epsilon' \le \epsilon$ for all $t = 0, \ldots, \sigma$. □

Lemma 4.3. *Given $\epsilon > 0$, there exists $\gamma > 0$ such that for each $T \in \mathbb{N}$ and each program $\{x(t)\}_{t=0}^T$ satisfying* $\operatorname{dist}(x(0), V) < \epsilon/2$, $\operatorname{dist}(x(T), V) < \epsilon/2$ *and $\delta(t) < \gamma$ for all $t = 0, \ldots, c, T - 1$, we have*

$$\operatorname{dist}(x(t), V) < \epsilon \quad \text{for all } t = 0, \ldots, T.$$

Furthermore, if Condition 2.1 holds,

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

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

1. *Case $T < n$.* Although the computation is quite onerous, the argument of the proof is based on a simple idea: if the distances $\operatorname{dist}(x(0), V)$ and $\operatorname{dist}(x(T), V)$ are small and the harvesting policy is similar to a periodic program harvesting, then the state of the forest cannot go far from V (in less than n steps) without making a large value loss at least once.
   The lower semi-continuity of $\delta(x, x')$ allows one to affirm that given ϵ, there is a γ_1 such that $\delta(x, x') < \gamma_1$ implies $\operatorname{dist}(\lambda(x, x'), S_c) < \epsilon_1 = \epsilon/(2n^2)$. [2]
   We start bounding the value of $x_i(t)$ for all $i = \sigma + 1, \ldots, n$, and after that we bound $\operatorname{dist}(x_i(t), S_f)$ for all $i = 1, \ldots, \sigma$. In the case $i > \sigma$, we can express $x_i(t)$ as a linear combination of the $x_{i+t-T}(T)$ or $x_n(t + n - i)$ and the harvests between t and T or between t and $t + n - i$ (that are controlled by ϵ_1) to deduce that $x_i(t) < \epsilon$ for all $i > \sigma$. We need to divide the study into two cases:
   (a) Case $i + T - t \le n$;
   $$x_i(t) = x_{i+T-t}(T) + \sum_{j=0}^{T-t-1} c_{i+j}(t + j) < \frac{\epsilon}{2} + (T - t)\epsilon_1 < \frac{\epsilon}{2} + n\frac{\epsilon}{2n^2} < \epsilon.$$
   (b) Case $i + T - t > n$;
   $$x_i(t) = x_n(t + n - i) + \sum_{j=0}^{n-i-1} c_{i+j}(t + j) = \sum_{j=0}^{n-i} c_{i+j}(t + j) < n\epsilon_1 = n\frac{\epsilon}{2n^2} < \epsilon.$$

[2] Indeed, the proof of this property follows by the same method as in Corollary 4.1.

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

$$\text{if } \operatorname{dist}(x(t), V) < \epsilon_2 \text{ then } \operatorname{dist}(x_i(t+1), S_f) < n\,\epsilon_1 + \epsilon_2 \quad \text{for all } i = 1, \ldots, \sigma. \tag{6}$$

Case $2 \le i \le \sigma$[3];

$$\operatorname{dist}(x_i(t+1), S_f) \le |x_i(t+1) - x_{i-1}(t)| + \operatorname{dist}(x_{i-1}(t), S_f) = |c_{i-1}(t)| + \operatorname{dist}(x_{i-1}(t), S_f) < \epsilon_1 + \epsilon_2.$$

Case $i = 1$;

$$\begin{aligned} x_1(t+1) = \sum_{i=1}^{n} c_i(t) &\implies c_\sigma(t) \le x_1(t+1) \le (n-1)\,\epsilon_1 + c_\sigma(t) \\ &\implies |x_1(t+1) - c_\sigma(t)| \le (n-1)\,\epsilon_1 \\ \implies \operatorname{dist}(x_1(t+1), S_f) &\le (n-1)\,\epsilon_1 + \operatorname{dist}(c_\sigma(t), S_f) < n\epsilon_1. \end{aligned}$$

Repeated application of and (6) yields

$$\begin{aligned} \operatorname{dist}(x(0), V) < \frac{\epsilon}{2} &\implies \operatorname{dist}(x(1), V) < n\epsilon_1 + \epsilon/2 \\ &\implies \operatorname{dist}(x(2), V) < n\epsilon_1 + n\epsilon_1 + \epsilon/2 = 2n\epsilon_1 + \epsilon/2 \\ &\ \ \vdots \\ &\implies \operatorname{dist}(x(T-1), V) < (T-1)n\epsilon_1 + \epsilon/2 \end{aligned}$$

and thus $\operatorname{dist}(x(t), V) < n^2\epsilon_1 + \epsilon/2 = \epsilon$ for all $t = 1, \ldots, T-1$.

2. *Case* $T \ge n$; Corollary 4.1 states that there is γ_2 such that for every program $\{x(t)\}_{t=0}^{n}$ satisfying $\delta(t) < \gamma_2$ for all $t < n$, we have $\operatorname{dist}(x(t), V) < \epsilon/2$ for all $t = 0, \ldots, \sigma$. We apply this result to the programs $\{x(t+i)\}_{t=0}^{n}$ with $i = 0, \ldots, T-n$ to conclude that $\operatorname{dist}(x(t), V) < \epsilon/2$ for $t = 0, \ldots, T-n+\sigma$.
   Afterwards, we apply the result of the first part of the lemma to conclude that the state also fulfills $\operatorname{dist}(x(t), V) < \epsilon$ for $t = T-n+\sigma, \ldots, n$.

On taking $\gamma = \min\{\gamma_1, \gamma_2\}$, the lemma follows. □

5. Proof of Theorem 3.1

Theorem 3.1. *Given $M > 0$ and $\epsilon > 0$ there exists $L \in \mathbb{N}$ such that for all $T > L$ and each program $\{x(t)\}_{t=0}^{T}$ satisfying*

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

we have that

$$\operatorname{Card}\{i \in [0, \ldots, T-1] : \operatorname{dist}(x(t), V) > \epsilon\} \le L.$$

Furthermore, if Condition 2.1 holds,

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

The idea of the proof is to use Lemma 4.3 to bound the distance from $x(t)$ to the set V. In general, it will not be possible to apply this lemma to the whole interval $[0, T]$. To overcome this difficulty we divide $[0, T]$ into 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. Given ϵ, by Lemma 4.3 there is γ such that for each $\{x(t)\}_{t=0}^{T}$ satisfying

$$\operatorname{dist}(x(0), V) < \epsilon/2, \qquad \operatorname{dist}(x(T), V) < \epsilon/2 \quad \text{and} \quad \delta(t) < \gamma \quad \text{for all } t \in [0, \ldots, T-1]$$

we have

$$\operatorname{dist}(x(t), V) < \epsilon \quad \text{for all } t \in [0, \ldots, T].$$

[3] We use a triangular inequality involving distances from a point to a set. To prove that $\operatorname{dist}(p, A) \le \operatorname{dist}(q, A) + \|p - q\|$, first assume that $\bar{r} \in A$ is the point such that $\operatorname{dist}(q, A) = \|q - \bar{r}\|$; then

$$\operatorname{dist}(q, A) + \|p - q\| = \|q - \bar{r}\| + \|p - q\| \ge \|p - \bar{r}\| \ge \operatorname{dist}(p, A).$$

If $\bar{r} \in \operatorname{closure}(A) \setminus A$ take $\{r_n\} \subset A$, $r_n \to \bar{r}$ and repeat the process above.

Given a program $\{x(t)\}$ satisfying the hypothesis and taking S, τ such that $0 \le S \le S+\tau \le T$ we can use Proposition 4.3 to obtain

$$\sum_{t=S}^{S+\tau-1} w\,(bc(t)) \ge U\,(x(S), x(S+\tau), S, S+\tau) - M$$

and Proposition 4.1 to deduce

$$U\,(x(S), x(S+\tau), S, S+\tau) - M \ge \tau w(b\hat{c}) - (n+\sigma)w(b\hat{c}) - M.$$

From the above and Lemma 4.1 it follows that there is a $\bar{\tau}$ such that

$$\text{for any } S \in [0, T-\bar{\tau}] \text{ there is } t \in [S, S+\bar{\tau}] \text{ such that } \operatorname{dist}(x(t), V) < \epsilon/2. \tag{7}$$

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

$$\bar{\tau} \le (t_i - t_{i-1}) \le 2\bar{\tau} \quad \text{and} \quad \operatorname{dist}(x(t_i), V) < \epsilon/2 \quad \text{for all } i = 1, \dots, K-1,$$

using the following algorithm: by (7) there is a $t_1 \in [\bar{\tau}, 2\bar{\tau}]$ such that $\operatorname{dist}(x(t_1), V) < \epsilon/2$. Using (7) again we know that there exists $t_2 \in [t_1+\bar{\tau}, \dots, t_1+2\bar{\tau}]$ such that $\operatorname{dist}(x(t_2), V) < \epsilon/2$. We proceed inductively defining

$$t_{i+1} \in [t_i+\bar{\tau}, t_i+2\bar{\tau}] \quad \text{with } \operatorname{dist}(x(t_{i+1}), V) < \epsilon/2.$$

We repeat this step until we obtain $(t_{K-1}+2\bar{\tau}) \ge T$, 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 4.3 whenever

$$\sum_{t=t_i}^{t_{i+1}-1} \delta(t) < \gamma \tag{8}$$

to affirm that $\operatorname{dist}(x(t), V) < \epsilon$ for all $t \in [t_i, t_{i+1}-1]$. We claim that there are $k \le 2 + \gamma^{-1}\left[(n+\sigma)w(b\hat{c}) + n\frac{b_\sigma}{\sigma}z + M\right]$ subintervals *not* fulfilling (8). Indeed, denote by $\mathcal{K} \subseteq [1, \dots, K-2]$ the set of indexes such that $\sum_{t=t_i}^{t_{i+1}-1} \delta(t) \ge \gamma$; it is easily seen that

$$\begin{aligned} \sum_{t=0}^{T-1} \delta(t) &= \sum_{k=0}^{K-1} \sum_{t=t_i}^{t_{i+1}-1} \delta(t) \\ &\ge \sum_{k\in\mathcal{K}} \sum_{t=t_i}^{t_{i+1}-1} \delta(t) \ge \gamma \operatorname{Card}\{\mathcal{K}\}. \end{aligned}$$

On the other hand we know that

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

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

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

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

References

[1] P.A. Samuelson, The periodic turnpike theorem, Nonlinear Analysis, Theory, Methods and Applications 1 (1976) 1–13.
[2] M. Ali Khan, A.J. Zaslavski, On two classical turnpike results for the Robinson–Solow–Srinivasan (RSS) model, Advances in Mathematical Economics 13 (Part 1) (2010) 47–97. doi:10.1007/978-4-431-99490-9-3.
[3] M. Ali Khan, A. Piazza, Classical turnpike theory and the economics of forestry, Publicación Técnica CMM. No. 228, 2009.
[4] R. Dorfman, P.A. Samuelson, R.M. Solow, Linear Programming and Economic Analysis, McGraw-Hill, New York, 1958.
[5] P.A. Samuelson, Maximum principles in analytical economics. Les Prix Nobel en 1970. The Nobel Foundation, Stockholm, 1971, pp. 273–288.
[6] P.A. Samuelson, Economics of forestry in an evolving society, Economic Inquiry XIV (1976) 466–492.
[7] F. Ramsey, A mathematical theory of savings, Economic Journal 38 (1928) 543–559.
[8] T. Mitra, H.W. Wan, Some theoretical results on the economics of forestry, Review of Economics Studies LII (1985) 263–282.
[9] T. Mitra, H.W. Wan, On the Faustmann solution to the forest management problem, Journal of Economic Theory 40 (1986) 229–249.

[10] D. Gale, On optimal development in a multi-sector economy, Review of Economic Studies 34 (1967) 1–18.
[11] L.W. McKenzie, Accumulation programs of maximum utility and the von Neumann facet, in: J.N. Wolfe (Ed.), Value, Capital and Growth, Edinburgh University Press, Edinburgh, 1968, pp. 353–383.
[12] M. Ali Khan, A. Piazza, On uniform convergence of undiscounted optimal programs in the Mitra–Wan forestry model: the strictly concave case, International Journal of Economic Theory 6 (2010) 57–76.
[13] L.W. McKenzie, Turnpike theory, Econometrica 43 (1976) 841–865.
[14] M. Ali Khan, A. Piazza, On the Mitra–Wan forestry model: a unified analysis, Publicación Técnica CMM. No. 226, 2009.
[15] E.B. Keeler, A twisted turnpike theorem, Zeitschrift für Nationalökonomie (Suppl. 1) (1971) 91–98.
[16] E.B. Keeler, A twisted turnpike, International Economic Review 13 (1972) 160–166.
[17] L.W. McKenzie, A primal route to the turnpike and Lyapunov stability, Journal of Economic Theory 27 (1982) 194–209.
[18] L.W. McKenzie, Turnpike theory, discounted utility, and the von Neumann facet, Journal of Economic Theory 30 (1983) 330–352.
[19] J. von Neumann, Über ein ökonomisches gleichungs-system und eine verallgemeinerung des Brouwerschen fixpunktsatzes, in: K. Menger (Ed.), Ergebnisse eines Mathematischen Kolloquiums, 8, in: Review of Economic Studies, vol. 13, 1935–1936, pp. 1–9. Translated as a model of general economic equilibrium.
[20] T. Mitra, Intergenerational equity and the forest management problem, in: [24], 2005, pp. 137–173 (Chapter 5).
[21] S. Salo, O. Tahvonen, On equilibrium cycles and normal forests in optimal harvesting of tree vintages, Journal of Environmental Economics and Management 44 (2002) 1–22.
[22] M. Ali Khan, Intertemporal ethics, modern capital theory and the economics of forestry, in: [24], 2005, pp. 39–65 (Chapter 2).
[23] S. Salo, O. Tahvonen, On the economics of forest vintages, Journal of Economics Dynamics and Control 27 (2003) 1411–1435.
[24] S. Kant, R.A. Berry (Eds.), Economics, Sustainability, and Natural Resources: Economics of Sustainable Forest Management, vols. 1, 2, Springer, Berlin, 2005.