On the Mitra-Wan Forestry Model: A Unified Analysis*

M. Ali Khan†
Adriana Piazza‡

First version: June 20, 2009. This version: August 10, 2009.

Abstract

We present a substantive and far-reaching generalization of the principal results in the economics of forestry, as formalized by Mitra-Wan (1986). Rather than a polarized dichotomy of linear and strictly concave, differentiable benefit (felicity) functions, we develop the theory in the context of functions that are supported at the golden-rule consumption and not necessarily concave. Through a non-interiority condition on the zeroes of a resulting "discrepancy function," we show the equivalence of finitely maximal, maximal, minimal value-loss and optimal programs, and thereby answer questions left open by Brock and Mitra. Our synthesizing criterion is new to the capital-theory literature, and in the concave setting, proves to be necessary and sufficient for the asymptotic convergence of good programs. (118 words)

Keywords Forest management, good programs, finitely maximal, maximal, optimal, value-loss, upper semicontinuity, discrepancy function, non-interiority, existence, asymptotic convergence.

JEL Classification Numbers C62, D90, Q23

1 Introduction

In 1986, Mitra-Wan [32] place the economics of forestry, developed by Faustmann, Wicksell, Ohlin and Samuelson, squarely within the modern theory of intertemporal allocation in discrete time, developed by Gale, McKenzie and Brock. In a retrospective reading, twenty years later, Mitra [28, p. 137] had the following summary evaluation.

> In forestry management, there has been a tradition which claims that "the goal of good policy is to have sustained forest yield, or even maximum sustained yield somehow defined" (Samuelson, 1976, p.146). In an attempt to understand this tradition, Mitra and Wan (1986) formulated ... forestry economics ... as a particular case of modern capital theory. Exploiting the optimization methods familiar from this general theory, they were able to show that, when the utility function is strictly concave, starting from any initial forestry configuration, the optimally managed forest converges over time to the forest with the maximum sustained yield, which corresponds to the "golden rule" of the forestry model. This demonstration provided a theoretical basis for the tradition in forestry management.

*This work was initiated during Piazza's visit to Johns Hopkins in May 2008 and completed during her visit to the University of Illinois at Urbana-Champaign in April-May, 2009. In addition to the hospitality of Nicholas Yannelis and that of the Economics Department at Illinois, the authors gratefully acknowledge invaluable discussion and correspondence with Tapan Mitra and Tom Cosimano. A previous version of the paper was presented at the *Fourth Workshop in Macroeconomic Dynamics* held at the National University of Singapore, July 31 to August 1, 2009.

†Department of Economics, The Johns Hopkins University, Baltimore, MD 21218 and Department of Economics, University of Illinois at Urbana-Champaign, Urbana, IL. 61801. **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

It is interesting that Mitra focusses only on the case of strictly concave benefit (felicity) functions, and does not refer to the linear case in which the periodicity of an optimally managed forest is the general rule. In fact, in this subsequent paper, he develops the analysis in the context of benefit functions that are concave and strictly mid-concave,[1] and in his subsequent notes on the dual-aged forest [29], and in the work of the authors themselves [18], the analysis is focussed entirely on the strictly concave case. While it appears that there is nothing further to be said for the linear case, what really bears emphasis is that the theoretical basis provided by this pioneering paper, and the analysis underlying it, has remained virtually untouched.[2] To be sure, there have been extensions of the model to the discounted setting; analyses of its many variations; and the effective exploitation of its basic methodology to other apparently unrelated issues such as the 'choice of technique' in development planning.[3] However, the prevailing view has been that there is little to add to the 1986 theorems themselves.

This is a little puzzling in view of the rather strict dichotomous nature of the Mitra-Wan analysis: one set of results for linearity and another for strict concavity. This polarity of conception has pervaded virtually all of subsequent investigations.[4] The question has remained open as to what one ought to expect for a concave function? When the question is posed in this way, there is perhaps the expectation that a synthetic result can be obtained simply by somehow putting together the separate analyses of the two cases. At one level this intuition is incontestable, but what it misses is the precise formulation of a result for the concave setting that recovers the strictly concave and the linear cases as corollaries. What is a surprise, and perhaps the principal finding of this paper, is that in the pursuit of such a result, the assumption of concave felicities can be dispensed with altogether, and a substantial part of the theory developed in the setting of upper semicontinuous benefit (felicity) functions.[5] In such a setting, we offer a non-interiority condition that is sufficient for the asymptotic convergence of good programs to the golden-rule forest configuration, and when the benefit function is concave, though not necessarily differentiable, our condition is both necessary and sufficient.

While such a unification is a primary motivation behind this essay, we are also intrigued by Mitra's recent axiomatic investigation of intergenerational equity, and his espousal of, what we call here, *finite maximality.* In particular, he shows, under concave and strictly mid-concave benefit functions, the equivalence of finitely maximal and maximal programs, and writes,

> In studying the long run properties of paths which are finitely maximal and those which are maximal according to the familiar overtaking criterion, the latter concept does not possess any discernible advantage. This leads us to conclude that in the context of the forestry model, one can completely dispense with the more restrictive overtaking criterion.[6]

For his equivalence theorem, Mitra not only had to assume strict concavity of benefit functions,[7] but also their continuous differentiability. This was necessitated by his invocation of the full duality theory associated with maximal programs and the consequent utilization of the Arrow-Hurwicz-Uzawa constraint qualification. He left it as an open question as to whether these hypotheses could be removed under an alternative method

[1] As we show in the Appendix, this is simply a rephrasing of strict concavity.

[2] In their search of the literature, the authors were somewhat surprised to learn that Mitra [28, 29], and their own [18], may be the only exceptions.

[3] For the discounted setting, see Mitra-Wan [31] and Salo-Tahvonen [39, 40]; for variants, see Mitra-Ray-Roy [33], Salo-Tahvonen [41], Cominetti-Piazza [7], Rapaport et al. [38] and Piazza [36, 37]; and for development planning, see Khan-Mitra [15, 16]. Commenting in their 2006 survey on the usefulness of the price-supported golden-rule in studying long run dynamic behavior of optimal programs in the undiscounted case, Mitra-Nishimura [30, Section 7] see the two papers, [32] and [16], together and note the "effective demonstration in applications of the theory to study the Faustmann solution in the forest management problem and in the analysis of the choice of technique in development planning."

[4] In addition to the references in the paragraph above and in Footnote 3, the reader is referred to [3, 8, 17] and their references.

[5] As the reader will see from the sequel, it is only in the analysis of finitely maximal programs that we need to assume continuity.

[6] See [28, p. 139]. In the quote, we have respectively substituted the terms *finitely maximal* and *maximal* for Mitra's *maximal* and *optimal.* We are aware that we add to the terminological confusion, but as referred to in Footnote 12 below, this is endemic to the subject.

[7] See Footnote 1 above, and also the Appendix below.

of proof.[8] Under an alternative proof, we show that Mitra's equivalence theorem holds under the larger class of benefit functions that we isolate here, and perhaps more importantly, the theorem dovetails to offer an even more satisfactory unified conception in which finitely maximal, maximal, minimal value-loss and optimal programs are all identical, and with the added payoff that the existence of one implies that of all the others. As such, our primary and secondary motivations coalesce in one theorem.[9]

The basic idea behind the analysis presented in this paper revolves around a non-interiority assumption, but one made on the set of timber yields rather than on the set of forest configurations of "today and tomorrow", which is to say, on the space of consumption levels rather than that of capital stocks. As such, there is a slight retreat in our analysis from the Gale-McKenzie-Brock reduced form formulation, with its exclusive focus on state variables, to a primitive assumption involving the control variables. Briefly stated, and leaving the details to the sequel, we formulate a natural "discrepancy function" f of the timber yields, and on isolating the set S_f of its "zeros," require the maximal yield to be an interior point of the convex hull of this set. In Figure 1, based on a situation where a tree of age i ($i = 1, 2, \cdots, n$) yields b_i units of timber when chopped down (the biomass coefficient), b_σ/σ is the golden-rule timber yield, $\sigma \in \{1, 2, \cdots, n\}$, and $w(\cdot)$ the felicity function on yields, it is only in Figure 1b that the condition we propose is not fulfilled. Figure 1a depicts a strictly concave function, Figure 1c an upper semicontinuous function that can be supported at the golden-rule stock, and Figure 1d a concave function. The basic idea is to show the asymptotic convergence of good programs, a fundamental notion due to Gale [11], if and only if this non-interiority assumption holds.[10] With this result in hand, the corresponding results in [32] can be generalized, and the basic theory fully outlined. In summary, a subtext underlying our entire analysis is the question as to what we know today that Mitra-Wan did not know in 1986? and accordingly, we conceive of the unified analysis presented here as much a contribution to the economics of forestry as to the general theory of intertemporal allocation of resources.[11]

In this connection, it is worth underscoring that we also consider optimal, as opposed to maximal, programs, and show their existence when the condition we offer on the discrepancy function holds.[12] This is in keeping with the recent emphasis in [44, 45, 46], and is somewhat opposed to Brock's initial intuition in [2] based on his example on the "von Neumann economy" which does not have an optimal program.[13] This example has dominated subsequent treatments in ensuring the exclusion of the optimality criterion. In fact our results answer, albeit for the Mitra-Wan tree farm, a question that Brock left open in 1970.[14] Finally, we can characterize the set of periodic programs with zero accumulated loss, and through them the optimal Faustmann policy, when this non-interiority condition does not hold. We again underscore the fact that these results, and the analyses underlying them, are set in a context of functions that are not assumed to be concave or even continuous,[15] leave alone differentiable, provided they are supported at (b_σ/σ), the consumption at the golden rule configuration.

---

[8]See [28, Footnote 25] where Mitra writes in the context of his continuous differentiability assumption on the benefit function, "While this assumption is crucial to our method of investigation (duality theory), it is not clear whether it is indispensable for the results of the next section on the relation between maximal and optimal paths. It would seem that a "primal route" to those results should be possible."

[9]See Theorem 7.2 below. This theorem can perhaps be regarded as the principal contribution of this essay.

[10]See Theorem 5.2 below. This emphasis on asymptotic convergence of good programs is also evident in [29], [44], and of course in [46]; also see [47].

[11]The authors thank Steve Lugauer, Mike Pries and Jim Sullivan for orienting them to this question.

[12]There is an important issue of rather unfortunate terminology here; optimal programs in the work of [15, 16] are referred to in this work as *maximal* programs, and strongly optimal programs as simply *optimal programs.* We take our terminology from McKenzie [24, p. 256], and it is also used in [44]. It is perhaps also worth mentioning that optimal programs are referred to as *overtakingly optimal* in [46], and the terminology of maximal programs is given a different meaning in [28]. This terminological proliferation is already evident in Brock [2]. Also see Footnote 6 above.

[13]Brock introduces the example at the very beginning of his paper (Example 2.1), and then returns to it at the end.

[14]Brock writes "One might ask if it is possible to strengthen Theorem 1 [the theorem on the existence of maximal programs] by showing existence of an optimal programme. We do not know the answer to this question but we can find an example" where the answer is negative. This example is the one referred to in Footnote 13.

[15]It is of course the lack of concavity that leads us to emphasize the absence of an explicit continuity assumption. In this connection, also see Footnote 8 above. In the context of piecewise linearity, threshold effects arising from non-differentiable felicity (benefit) functions have recently been re-emphasized by [12].

One final observation regarding the relevance of the analysis presented here to the general theory of intertemporal allocation of resources. In departing from the notational framework of [32], and treating a forest configuration as simply a point in a $(n-1)$-dimensional simplex,[16] and the model as the pair $(w, b),\ b = (b_1, b_2, \cdots, b_n)$, we see that the two basic assumptions of the general theory, "inaction" and "free disposal" are not fulfilled.[17] The absence of these assumptions is of course circumvented in the literature,[18] but by making it explicit, we can dispense with all assumptions on the biomass coefficients[19] other than the Brock-Mitra-Wan assumption of a unique golden-rule configuration. This is the assumption of the existence of $\sigma \in \{1, 2, \cdots, n\}$ such that $(b_\sigma/\sigma) > (b_i/i)$ for all $i \neq \sigma$. In this connection, Zaslavski's recent rewriting [46] of the general theory in the context of compact metric spaces, and without any linear and ordered structures, is also of interest, and serves as a useful point of introduction to our work. Despite its apparent generality (no convexity assumptions on the technology or concavity assumptions on the benefit function), the sufficient conditions isolated by Zaslavski, and in particular his emphasis on the interiority assumption, do not translate to the Mitra-Wan forestry model. In particular, what he takes as one of the hypothesis of his results, the asymptotic convergence of good programs, is precisely what we need to prove as a consequence of our non-interiority condition. The fact that it also implies this condition is an added and satisfying bonus.

The remainder of the paper is as follows. In Section 2, we present a brief recapitulation of the general theory, and apply its basic conceptual vocabulary in Section 3 to the Mitra-Wan forestry model and its golden-rule configuration. In keeping with the emphasis given to good programs in [29] and in [46, 47], Sections 4 and 5 concern good programs: the first deals with existence and characterization, and the second with the question of asymptotic convergence. Section 5 presents the non-interiority condition on the "discrepancy function" and alternative characterizations and translations of the von-Neumann facet. Section 6 concerns optimal and maximal programs, and is devoted to showing the equivalence between maximal, optimal and minimal value loss programs when the non-interiority condition holds, and to spelling out the relations between these concepts that are valid in general. Section 7 considers finitely maximal programs, and presents the unification theorem. Section 8 concludes the paper with remarks pertaining to future work that we contemplate for both the deterministic and stochastic settings. The fact that strict concavity is identical to concavity *and* strict mid-concavity is relegated to an Appendix.

2 The General Theory: A Recapitulation

The Gale-McKenzie reduced form model [11, 22] has by now been comprehensively surveyed, and received handbook and textbook treatment; see [23, 24, 27, 9]. For a reader wanting only a brief and basic introductory outline, Brock [2] remains the relevant reference.[20] In the setting of a finite-dimensional Euclidean space, Brock defined maximal and optimal programs of an infinite period optimization problem, and under a uniqueness assumption of the solution of a static version of this problem, showed the existence of maximal programs.[21] Through an example, he also showed the non-existence of an optimal program. Brock's existence proof is interesting on at least two counts: (i) his reliance on a uniqueness assumption, in addition to compactness of the constraint set, and continuity on it of the objective function, (ii) his simultaneous characterization of maximal programs as necessarily implying suitably-defined minimum aggregate value-loss. The methodology of his proof, as well as his counterexample, have proved influential, and especially in bringing out the intertwined nature of the issues of existence and characterization.[22] Thus, Mitra-Wan in the context of their forestry model, and Khan-Mitra in the context of the RSS model, limit themselves to maximal (as opposed to optimal) programs, and use their existence proofs to characterize Faustmann policies, on the one hand, and Stiglitz policies, on the other; see [32, 16].

[16]This notational simplification is proposed in [39] and pursued in [14, 18].

[17]In addition to [23, 24], see the more recent statements of this theory in Mitra [27] and Dana et al. [9].

[18]In addition to [32] and [18], this is also implicit in treatments of the multi-specie extensions of the Mitra-Wan forest in [7], and in [36, 37].

[19]As will emerge in the sequel, we do not use assumptions A.1, A.2, A.4 and A.7 in [32] at all.

[20]See, for example, the reference to Brock's 1970 paper in [28, Section 7].

[21]Footnotes 12 and 13 above are relevant here.

[22]The authors are grateful to Alex Himonas and Luciano de Castro for discussion of these issues. Also see Footnote 13 above.

In this brief recapitulation of the general theory, we begin with some preliminary notation. Let $I\!N$ be the set of non-negative integers and $I\!R$ ($I\!R_+$) be the set of real (non-negative) numbers. In this section, we shall work in a compact metric space (X, ρ). Let $\Omega \subseteq X \times X$ be the (stationary) technology and $u : X \times X \longrightarrow I\!R$ the (stationary) felicity or benefit function. We shall assume that Ω is nonempty and closed, and u a bounded upper semicontinuous function. This pair of objects (Ω, u) represents the model, and we present two of the concepts that are basic to the theory.

Definition 2.1 *A sequence $\{x(t)\}_{t=0}^{\infty} \subset X$ is called a* program *if $(x(t), x(t+1)) \in \Omega$ for all integers $t \geq 0$. For any natural number T, a sequence $\{x(t)\}_{t=0}^{T} \subset X$ is called a T*-program *if $(x(t), x(t+1)) \in \Omega$ for all integers $0 \leq t \leq T-1$. For any $x \in X$, we shall say that a program $\{x(t)\}_{t=0}^{\infty}$, or a T-program $\{x(t)\}_{t=0}^{T}$, is a* program, *or a T*-program *from $x \in X$, if $x(0) = x$.*

We suppose, as in the literature taking its lead from Ramsey (1928), that future welfare levels are treated like current ones in the planner's objective function.

Definition 2.2 *A program $\{x^*(t)\}_{t=0}^{\infty}$, is* optimal *if for any program $\{x(t)\}_{t=0}^{\infty}$, such that $x(0) = x^*(0)$ we have*

$$\limsup_{T\to\infty} \sum_{t=0}^{T} u(x(t), x(t+1)) - u(x^*(t), x^*(t+1))) \leq 0.$$

In a recent paper, Zaslavski [46] exercises Occam's razor, and presents results which do not rely on any convexity assumptions on the technology or concavity assumptions on the felicity function. In particular, they do not rely on any "free disposal" or "inaction" assumptions. We present two of these results in an attempt to bring out how far the theory can proceed only with a metric structure, though by the assumption of hypotheses that are either not fulfilled in the Mitra-Wan model or whose verification represents precisely the analysis that needs to be carried out. Despite the direct inapplicability of these theorems to our context, a point more fully established in the next section, they are useful as a parsimonious introduction to the general theory, and in highlighting three substantive points: (i) an emphasis on interiority,[23] and on the implicit retention of Brock's uniqueness assumption, (ii) an emphasis on good programs and their convergence, as opposed to their Cesàro summability, as in [2],[24] and, (iii) a return to the consideration of optimal (as opposed to maximal) programs.

For any natural number T, let

$$\sigma(u, \Omega, T) = \sup \left\{ \sum_{t=0}^{T-1} u(x(t), x(t+1)) : \{x(t)\}_{t=0}^{T} \text{ is a } T\text{-program} \right\},$$

where we follow the convention that the supremum of an empty set is negative infinity.

We now reproduce the first substantive assumption on the pair (Ω, u) in [46].

Assumption 2.1 *There exists $(\hat{x}, \hat{x}) \in \Omega$ and a constant $c > 0$ such that (i) $(\hat{x}, \hat{x})$ is an interior point of Ω, (ii) $u(\cdot, \cdot)$ is continuous at $(\hat{x}, \hat{x})$, (iii) $Tu(\hat{x}, \hat{x}) + c \geq \sum_{i=1}^{T} \sigma(u, \Omega, T)$ for all natural numbers $T \geq 1$.*

It is easy to see that under Assumption 2.1, for each natural number T and each T-program,

$$\sum_{i=1}^{T-1} u(x(t), x(t+1)) \leq \sigma(u, \Omega, T) \leq Tu(\hat{x}, \hat{x}) + c. \tag{1}$$

[23] Zaslavski's interiority assumption is made on the technology, and as such is very different from that our assumption developed below.

[24] For Cesàro summability, the reader is referred to the Wikipedia entry on **Cesàro mean**. Following [2], this property has subsequently referred to in the literature as the "average turnpike property;" see, for example, [32] and [16] and also Lemma 6.2 and Footnote 33 below. Other than in this sentence, and in keeping with the discussion in [19], we avoid the term *turnpike* or *average turnpike* in this paper.

This implies that the sequence $\{g_x(T)\}_{T=1}^{\infty}$ is bounded or diverges to negative infinity, where

$$g_x(T) \equiv \left\{ \sum_{i=1}^{T-1} u(x(t), x(t+1)) - Tu(\hat{x}, \hat{x}) \right\}.$$

For the next assumption, we need the notion of good programs originally due to Gale [11].

Definition 2.3 *A program $\{x(t)\}_{t=0}^{\infty}$ is* good *if the corresponding sequence $\{g_x(T)\}_{T=1}^{\infty}$ is bounded.*

We now reproduce the second substantive assumption on the pair (Ω, u) in [46]. Note that this assumption implicitly requires Brock's uniqueness assumption on the golden-rule stock, and is of course not an assumption on the primitives on (Ω, w) of the model.[25]

Assumption 2.2 *Any good program converges to $\hat{x}$ as defined in Assumption 2.1.*

We can now present the basic existence result from [46, Theorem 2.2]. The invocation of the hypothesis of asymptotic convergence of good programs for a result on the existence of optimal programs should be particularly noticed.

Theorem 2.1 *Under Assumptions 2.1 and 2.2, for any z in X, if there exists a good program $\{x(t)\}_{t=0}^{\infty}$ with $x(0) = z$, there exists an optimal program $\{x^*(t)\}_{t=0}^{\infty}$ with $x^*(0) = z$.*

Next, we present a necessary and sufficient condition from [46, Theorem 2.4] for the characterization of optimal programs. But for this, we first need to consider, for any given $M > 0$, the set X_M of initial stocks $x \in X$ for which there exists a program $\{x(t)\}_{t=0}^{\infty}$, $x(0) = x$, and

$$\sum_{t=0}^{T-1} (u(x(t), x(t+1)) - u(\bar{x}, \bar{x})) \geq -M \text{ for all } T \geq 1.$$

These are stocks that lead to programs that are good "of the order" M.

We can now present

Theorem 2.2 *Under Assumptions 2.1 and 2.2, any program $\{x(t)\}_{t=0}^{\infty}$, with $x(0) \in \cup\{X_M : M \in (0, \infty)\}$, is optimal if and only if (i) $\lim_{t\to\infty} \rho(x(t), \hat{x}) = 0$, and (ii) for each natural number T, and for any T-program with $\{y(t)\}_{t=0}^{T}$, with $y(0) = x(0)$, $y(T) = x(T)$,*

$$\sum_{t=0}^{T-1} u(y(t), y(t+1)) \leq \sum_{t=0}^{T-1} u(x(t), x(t+1)).$$

We now turn to the Mitra-Wan tree farm and show how these theorems, inspite of providing useful relevant benchmarks, are not directly applicable.

3 The Mitra-Wan Tree Farm and its Golden-Rule Configuration

From now on, we move from a compact metric space to a n-dimensional Euclidean space $\mathbb{R}^n$. We shall work in the $(n-1)$-dimensional simplex $\Delta = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}$. 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 $\|x\|_\infty$. We will consider the distance induced by the supremum norm $\text{dist}(x, x') = \|x - x'\|_\infty$.

In addition to its original formulation [31, 32], an outline of the Mitra-Wan forestry model is also available in [28], and of the special dual-aged case, in [27, 29].[26] Here we depart from the original specification and

[25] The reader will have to wait till Section 3 for a formal definition of the golden-rule stock to see that the implicit definition of $\hat{x}$ in [46] is identical to it.

[26] However, note Footnote 20 in [28].

work with the reformulation presented in [39] and pursued in [14, 18, 40]. 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 $w : [0, \infty) \to I\!R$ 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. We invite the reader to compare this parsimonious conception with that of [32], and to 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.[27]

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."[28] 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.[29]

In summary, for each period $t \in I\!N$, we denote $x_i(t) \geq 0,\ i = 1, \ldots, n$, the surface area occupied by trees of age i at time t. We represent the state of the forest by the vector $x(t) = \big(x_1(t), \ldots, x_n(t)\big) \in \Delta$. At every stage the planner 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 the age 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

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

This leads us to rewrite Definition 2.1 as

Definition 3.1 *A sequence $\{x(t)\}_{t=0}^\infty$ is called a program if for each $t \geq 0$*

$$\begin{cases} x(t) \in \Delta, \\ x_{i+1}(t+1) \leq x_i(t) \quad i = 1, \ldots, n{-}1 \end{cases} \tag{2}$$

We can now 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 (2). Namely,

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

Note, in passing, that the transition set Ω is convex, closed and stationary, and as such complications arising from non-convexity and non-stationarity can only originate from the benefit function.

Definition 3.2 *The vector of harvests to perform this transition is given by the function $\lambda : \Omega \to I\!R^n_+$,*

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

In addition, it is easy to see that $(x, x') \in \Omega \Leftrightarrow (x, x') \in \Delta \times \Delta$ and $\lambda(x, x') \geq 0$.

[27] As was pointed out in Footnote 19, we do not use Assumption A.7 in [32] at all.
[28] See [32, p. 232]. The same point is made in [28, Section 4, Paragraph 5].
[29] This is simply a somewhat subtle point of interpretation; the technicalities of the two analyses remain the same.

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

$$u(x, x') = w(bc) \text{ where } c = \lambda(x, x'). \tag{3}$$

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

$$w(y) \leq w(\tfrac{b_\sigma}{\sigma}) + z(y - \tfrac{b_\sigma}{\sigma}) \qquad \text{for all } y \in I\!R_+ \tag{4}$$

We now turn to the specification of the golden-rule forest configuration, and this is a good point to relate the analysis presented below to the antecedent literature. Note that the specification of the technology Ω precludes Brock's "inaction" and "free disposal" assumptions, as it does Zaslavski's interiority condition. The set Ω has no interior point in $I\!R^{2n}$ and the natural coordinate pre-order in this space does not apply to it.[30] Furthermore, Mitra-Wan appeal to the differentiability of the benefit function to provide directly the golden-rule forest configuration and the golden-rule prices associated with it.[31] We provide a self-contained argument to characterize the golden-rule forest configuration in Theorem 3.1.

Definition 3.3 *A* golden-rule stock $\hat{x} \in I\!R^n_+$ *is such that* $(\hat{x}, \hat{x})$ *is a solution to the problem:*

$$maximize_{(x,x)\in\Omega}\ u(x, x) = maximize_{(x,x)\in\Omega}\ w(b\lambda(x, x)).$$

Definition 3.3 coincides with the definition provided by Mitra and Wan [32]. Set $\hat{p} \in I\!R^n_+$, $\hat{p} = z\frac{b_\sigma}{\sigma}(1, 2, \ldots, n)$.

Next, we define the function[32] $\delta : \Omega \longrightarrow I\!R$.

Definition 3.4 *The* value loss *associated with any* $(x, x') \in \Omega$ *is given by*

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

It is easy to see that the function $\delta(\cdot, \cdot)$ is lower semicontinuous and the following lemma proves that $\delta(x, x') \geq 0$ for any $(x, x') \in \Omega$, and also determines a disaggregated lower bound of the value loss function that will be used afterwards in the characterization of the von Neumann facet. The non-negativity of the value loss function is already established in [32, Lemma 3.1] when w is differentiable and concave.

Lemma 3.1 *For any* $(x, x') \in \Omega$ *we have*

$$\delta(x, x') \geq z\big[\textstyle\sum_{i=1}^{n-1}(\frac{b_\sigma}{\sigma} - \frac{b_i}{i})\, i\, (x_i - x'_{i+1}) + (\frac{b_\sigma}{\sigma} - \frac{b_n}{n})nx_n\big] \geq 0 \tag{5}$$

Proof. Due to (4), we know that

$$w(y) \leq w(\tfrac{b_\sigma}{\sigma}) + z(y - \tfrac{b_\sigma}{\sigma}) \quad \text{for all } y \in I\!R.$$

In particular, taking $y = b\lambda(x, x')$ we get

$$\begin{aligned}
\delta(x, x') &\geq w(\tfrac{b_\sigma}{\sigma}) - w(\tfrac{b_\sigma}{\sigma}) - z\big[\textstyle\sum_{i=1}^{n-1} b_i(x_i - x'_{i+1}) + b_n x_n - \frac{b_\sigma}{\sigma}\big] - \sum_{i=1}^{n} z\frac{b_\sigma}{\sigma} i(x'_i - x_i)\\
&= z\big[-\textstyle\sum_{i=1}^{n-1} b_i(x_i - x'_{i+1}) - b_n x_n + \frac{b_\sigma}{\sigma} - \frac{b_\sigma}{\sigma} \underbrace{\textstyle\sum_{i=1}^{n} i(x'_i - x_i)}_{(A)}\big]
\end{aligned}$$

[30] To endow Δ with a pre-order, the natural way to proceed would be to consider a *reduced*, $(n-1)$*-dimensional state* of the forest, like for example: $z = (x_2, \ldots, x_n)$, using the area constraint to deduce the area occupied by trees of age 1, $x_1 = 1 - \sum_{i=2}^n x_i$. We can now work with the pre-order defined by $z \geq z'$ if $z_i \geq z'_i$ for all $i = 2, \ldots, n$. To see that the new formulation of the model does not fulfill the free disposal assumption, consider the state $z = (0, \ldots, 0)$ representing the case where all the trees are of age 1. It is possible to go from the state z today to the state $y = (1, \ldots, 0)$ tomorrow (all the trees are of age 2). Any state $z' \neq z$, will satisfy $z' \geq z$, but the transition from z' to y will be not possible.

[31] In addition to Assumption 2.1 above, see [2, Assumption 1] and [32, Lemma 3.1]. Note that Brock's "sufficiency" assumption [2, Assumption 2] is automatically fulfilled in our context.

[32] A warning to the reader that the function $\delta(\cdot, \cdot)$ is to be distinguished from the real number δ, typically assumed to be positive.

Rearranging the last summation we get:

$$\begin{aligned}(A) &= \sum_{i=1}^{n} ix'_i - \sum_{i=1}^{n} ix_i = \sum_{i=0}^{n-1}(i+1)x'_{i+1} - \sum_{i=1}^{n} ix_i \\ &= \sum_{i=0}^{n-1} x'_{i+1} + \sum_{i=1}^{n-1} i(x'_{i+1} - x_i) - nx_n = 1 - nx_n + \sum_{i=1}^{n-1} i(x'_{i+1} - x_i)\end{aligned} \tag{6}$$

We substitute this expression in the previous inequality,

$$\begin{aligned}\delta(x,x') &\geq z\big\{-\sum_{i=1}^{n-1} b_i(x_i - x'_{i+1}) - b_n x_n + \tfrac{b_\sigma}{\sigma} - \tfrac{b_\sigma}{\sigma}\big[1 - nx_n + \sum_{i=1}^{n-1} i(x'_{i+1} - x_i)\big]\big\} \\ &= z\big[\sum_{i=1}^{n-1}(\tfrac{b_\sigma}{\sigma} - \tfrac{b_i}{i})\, i\, (x_i - x'_{i+1}) + (\tfrac{b_\sigma}{\sigma} - \tfrac{b_n}{n})nx_n\big] \geq 0\end{aligned}$$

where the last inequality follows easily by observing that $(\frac{b_\sigma}{\sigma} - \frac{b_i}{i}) \geq 0$ and $(x,x') \in \Omega$ and $z > 0$. ■

The following theorem is basic to the subject. A proof is provided by [32, Theorem 3.1] using slightly stronger hypothesis on the biomass coefficients. We provide a proof suitable to our framework, one that does without the already mentioned differentiability assumptions on the felicity function w.

Theorem 3.1 *There exists a unique golden-rule stock* $\hat{x} = (\underbrace{\tfrac{1}{\sigma}, \ldots, \tfrac{1}{\sigma}}_{\sigma}, 0, \ldots, 0)$

Proof. First observe that $\hat{c} = \lambda(\hat{x},\hat{x})$ is such that $\hat{c}_\sigma = \frac{1}{\sigma}$ and $\hat{c}_i = 0$ for all $i \neq \sigma$, hence $b\hat{c} = \frac{b_\sigma}{\sigma}$. The proposition above yields

$$\delta(x,x) = w(\tfrac{b_\sigma}{\sigma}) - w(b\lambda(x,x)) - \hat{p}(x-x) \geq 0 \quad \text{for all } (x,x) \in \Omega$$

implying immediately that $u(\hat{x},\hat{x}) = w(\frac{b_\sigma}{\sigma}) \geq w(b\lambda(x,x))$ for all $(x,x) \in \Omega$.

It is left to see that $\hat{x}$ is the unique golden-rule stock. Let us suppose that there is $x \neq \hat{x}$ solution to the problem stated in Definition 3.3. Using that w is strictly increasing at $\frac{b_\sigma}{\sigma}$ we get

$$\tfrac{b_\sigma}{\sigma} = b\lambda(x,x) = \sum_{i=1}^{n-1} b_i(x_i - x_{i+1}) + b_n x_n \tag{7}$$

On the other hand, condition $(x,x) \in \Omega$ forces $(x_i - x_{i+1}) \geq 0$, which together with (3) yields

$$b\lambda(x,x) = \sum_{i=1}^{n-1} \tfrac{b_i}{i} i(x_i - x_{i+1}) + \tfrac{b_n}{n} nx_n \leq \sum_{i=1}^{n-1} \tfrac{b_\sigma}{\sigma} i(x_i - x_{i+1}) + \tfrac{b_\sigma}{\sigma} nx_n$$

with strict inequality unless $(x_i - x_{i+1}) = 0$ for all $i \neq \sigma$ and $x_n = 0$. Setting $x = x'$ in equation (6) we can rewrite the right hand side above to get

$$b\lambda(x,x) \leq \tfrac{b_\sigma}{\sigma}\big[\sum_{i=1}^{n-1} i(x_i - x_{i+1}) + nx_n\big] = \tfrac{b_\sigma}{\sigma}[1 - nx_n + nx_n] = \tfrac{b_\sigma}{\sigma}$$

It is easy to see that to have (7), we need $x_i = x_{i+1}$ for all $i \neq \sigma$, $x_n = 0$ and $x_\sigma - x_{\sigma+1} = \frac{1}{\sigma}$, namely, $\lambda(x,x) = \hat{c}$. From this and $(x,x) \in \Omega$ it follows that $x = \hat{x}$. ■

4 On Good Programs: Existence and Characterization

As emphasized in the introduction, the principal thrust of the analysis presented in this paper revolves around the identification of a necessary and sufficient condition for the asymptotic convergence of good programs. Towards this end, we develop in this section some preliminary results concerning the existence and characterization of good programs.

For the remainder of the paper, we shall abbreviate $\{x(t)\}_{t=0}^{\infty}$ by $\{x(t)\}$ and rewrite Definition 2.3 in terms of the parameters of the forestry model.

Definition 4.1 *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* $\lim_{T\to\infty} \sum_{t=0}^{T}[w(bc(t)) - w(\frac{b_\sigma}{\sigma})] = -\infty$.

As shown in [32, Lemma 4.3], the following general result of Gale applies to the Mitra-Wan forestry model. A proof for our notational framework is available as the proof of [18, Proposition 2.1].

Proposition 4.1 *The space of programs is partitioned into good and bad programs.*

The proposition below shows an equivalent characterization of good and bad programs and its proof is available as the proof of [18, Proposition 2.2] with the usual care in regarding z as an element of the subdifferential.

Proposition 4.2 *A program* $\{x(t)\}$ *is good if and only if* $\sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big)$ *is finite, and bad if and only if* $\sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big)$ *is infinite.*

Next we record the existence of at least one good program from any initial state as constructed in [18, Remark 2.2].

Lemma 4.1 *There is a good program from every* $x_0 \in \Delta$.

Let $x_0 \in \Delta$. Set $\mu(x_0) = \inf\big\{\sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big) : \{x(t)\} \text{ is a program from } x_0\big\}$.

The lemma above implies that $\mu(x_0) < \infty$. We can now establish the existence of a program that attains minimum aggregate value-loss. This is a benchmark result in the literature, but the proof we present adapts an argument in [9, Proposition 1.4.2] and circumvents Cantor's diagonalization argument in [2], and following him, in [32] and [16].

Proposition 4.3 *From any* $x_0 \in \Delta$ *there exists a program* $\{x(t)\}$ *such that*

$$\sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big) = \mu(x_0). \tag{8}$$

Proof. Let us define the functions

$$\bar{\delta} : \Delta \times \Delta \to \overline{I\!R}_+, \qquad \bar{\delta}(x, x') = \begin{cases} \delta(x, x') & \text{if } (x, x') \in \Omega \\ \infty & \text{else} \end{cases}$$

and

$$\gamma : \Pi \to \overline{I\!R}_+, \qquad \gamma(\{x(t)\}) = \sum_{t \in I\!N} \bar{\delta}\big(x(t), x(t+1)\big)$$

where $\Pi = \prod_{t=0}^{\infty} \Delta$. Let $\gamma_T(\{x(t)\}_{t\in I\!N}) = \sum_{t=0}^{T-1} \delta\big(x(t), x(t+1)\big)$. The function $\bar{\delta}$ is lower semi continuous and so it is γ_T with respect to the product topology, for every T. In addition $\gamma_T \leq \gamma_{T+1}$, hence γ is the increasing limit of l.s.c. functions, it is therefore lower semi continuous. We know that Π is compact in the product topology and that there is at least one good program, so the domain of γ is non-empty. Then, there is a minimizer $\{x^*(t)\}$ such that $\gamma(\{x^*(t)\}) = \mu(x_0) < \infty$. ■

5 On Good Programs: Asymptotic Properties

It is now well-understood from the general theory that with strictly concave felicity functions, the von-Neumann facet comprises only the pairs of states whose harvest corresponds to the harvest associated to the golden-rule forest configuration, and as a result, any good program asymptotically converges to the golden-rule stock. This result is established for the forestry model in [32, Lemma 6.4]. We prove the convergence of good programs towards the golden rule stock, $\hat{x}$, not only when w is strictly concave but for a broader family of benefit functions, and towards that end, present a condition concerning the support function at the point (b_σ/σ) with slope z. Let the discrepancy function, f, be

$$f(\xi) = w(\tfrac{b_\sigma}{\sigma}) - w(b_\sigma \xi) + z(b_\sigma \xi - \tfrac{b_\sigma}{\sigma}). \tag{9}$$

Thanks to (4), we know that $f(\xi) \geq 0$ for all $\xi \in I\!R_+$. Let $S_f \subseteq I\!R_+$ be the set where f attains zero, its global minimum. Of course $(1/\sigma) \in S_f$. In the particular cases where w is linear, concave or strictly concave, the set S_f is respectively $I\!R_+$, a closed interval or $\{1/\sigma\}$. In fact, the identity $S_f = \{1/\sigma\}$ corresponds to the family of strictly supported felicity functions,

$$S_f = \{1/\sigma\} \iff w(y) > w(\tfrac{b_\sigma}{\sigma}) + z(y - \tfrac{b_\sigma}{\sigma}) \qquad \text{for all } y \neq \tfrac{b_\sigma}{\sigma}$$

Let $S_c = \{c \in I\!R^n_+ : c_i = 0 \;\text{ for all }\; i \neq \sigma \;\text{ and }\; c_\sigma \in S_f\}$, the set of harvests of σ-aged trees that belong to the zeroes of the discrepancy function, and which allows us to give an alternative characterization of the von Neumann facet. This characterization will play a crucial analytical role in the sequel.

Proposition 5.1 *The von Neumann facet is*

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

Proof. Recall Lemma 3.1, and in particular (5) proved there:

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

Observing the sign of the coefficients $(\frac{b_\sigma}{\sigma} - \frac{b_i}{i})$ it is easy to conclude that

$$\delta(x, x') = 0 \text{ implies } x_i = x'_{i+1} \text{ for all } i \neq \sigma, i < n \text{ and } x_n = 0. \tag{10}$$

It remains to prove that $c_\sigma \in S_f$. Using (6) and (10) we can find a much simpler expression for

$$\textstyle\sum_{i=1}^{n} i(x'_i - x_i) = 1 - n x_n + \sum_{i=0}^{n-1} i(x'_{i+1} - x_i) = 1 - \sigma(x_\sigma - x'_{\sigma+1}),$$

and so,

$$\begin{aligned} \delta(x, x') = 0 \quad &\Rightarrow \quad \delta(x, x') = w(\tfrac{b_\sigma}{\sigma}) - w(b_\sigma c_\sigma) - z b_\sigma[\tfrac{1}{\sigma} - (x_\sigma - x'_{\sigma+1})] = 0 \\ &\Rightarrow \quad f(c_\sigma) = 0 \qquad\qquad\qquad \text{(Using (9)).} \end{aligned}$$

■

The following set V is essential for the rest of the exposition.

$$V = \{x \in \Delta : x_i \in S_f \;\text{ for all }\; i \leq \sigma \;\text{ and }\; x_i = 0 \text{ for all } \; i > \sigma\}. \tag{11}$$

Evidently, $\hat{x} \in V$. Proposition 5.1 implies that V is the set of initial conditions from where zero value loss programs originate. Such programs must be σ-periodic. We claim that this behavior is typical in the sense that a good program converges to V. Hence, it will be of use to know when

Condition 5.1 (Standing Condition) $V = \{\hat{x}\}$

We will show that the condition above is necessary and sufficient to assure the asymptotic convergence of every good program to the golden rule stock, and that the following, easier-to-check, condition is sufficient to assure Condition 5.1.

Condition 5.2 (Non-interiority) $(1/\sigma) \notin int\ co(S_f)$.

In the particular case where the felicity function is concave, the condition above turns out to be also necessary, and can be simplified to

Condition 5.3 (Non-Interiority, Concave case) $(1/\sigma) \notin int\ S_f$.

Of course, this last condition is assured when w is strictly concave.

Before getting into the proofs of these assertions, we present an alternative characterization of V as the solution of the following optimization problem.

Proposition 5.2 *Consider the optimization problem,*

$$(P) \begin{cases} \alpha = \min \sum_{i=1}^{\sigma} w(b_\sigma/\sigma) - w(b_\sigma x_i) \\ \quad s.t.\ x \in \Delta \end{cases} \tag{12}$$

Then, $\alpha = 0$ and the solution to P, $S(P)$, is given by the set V.

Proof. We can easily see that P has a non-empty solution because its objective function is lower semi-continuous and the feasible set is non-empty and compact. Consider $x \in S(P)$. Only the first σ coordinates are involved in the minimization and w is increasing, hence every $x \in S(P)$ must be of the form $x = (x_1, x_2, \ldots, x_\sigma, 0, \ldots, 0)$.

Take such x and consider the $\sigma-$periodic program that is obtained when harvesting only and completely the $\sigma-$age class at each step. The value loss accumulated during the first σ steps is of course non-negative and can be expressed as

$$\begin{aligned} 0 \leq \sum_{t=0}^{\sigma-1} \delta\big(x(t), x(t+1)\big) &= \sum_{t=0}^{\sigma-1} [w(\tfrac{b_\sigma}{\sigma}) - w(b_\sigma x_\sigma(t)) - \hat{p}\ (x(t+1) - x(t))] \\ &= \sum_{t=0}^{\sigma-1} [w(\tfrac{b_\sigma}{\sigma}) - w(b_\sigma x_{\sigma-t})] - \hat{p}\ (x(\sigma) - x(0)) \\ &= \sum_{i=1}^{\sigma} [w(\tfrac{b_\sigma}{\sigma}) - w(b_\sigma x_i)] \end{aligned}$$

with equality when $x = \hat{x}$. From this, we know that $\alpha = 0$ and $\hat{x} \in S(P)$.

From the above, we also know that given any $x \in S(P)$, we have that the aggregate value loss along σ steps of a periodic program is zero. Proposition 5.1 implies that $x_i \in S_f$ for all $i = 1, \ldots, \sigma$. Hence $x \in V$, and we have proved that $S(P) \subseteq V$.

To prove the converse consider any state $x \in V$. Consider the $\sigma-$periodic program that is obtained when harvesting only and completely the $\sigma-$age class at each step. Evidently $\lambda\big(x(t), x(t+1)\big) \in S_c$, which implies $\delta\big(x(t), x(t+1)\big) = 0$ for all t. Therefore

$$\begin{aligned} 0 &= \sum_{t=0}^{\sigma-1} \delta\big(x(t), x(t+1)\big) \\ &= \sum_{t=0}^{\sigma-1} [w(\tfrac{b_\sigma}{\sigma}) - w(b_\sigma x_{\sigma-t})] + \hat{p}\,\underbrace{(x(\sigma) - x(0))}_{=0} \implies x \in S(P) \end{aligned}$$

■

Remark 5.1 *It is possible to have zero value loss during σ steps even if the initial state does not belong to V. In fact, requiring $\sum_{i=1}^{\sigma} w(\frac{b_\sigma}{\sigma}) - w(b_\sigma x_i) = 0$ is stronger than the condition $\sum_{t=0}^{\sigma-1} \delta\big(x(t), x(t+1)\big) = 0$.*

We illustrate this with a toy example. Consider a forest where $n = 4$, $\sigma = 2$ and $S_f = [\frac{1}{2} - \phi, \frac{1}{2}]$ for some $\phi > 0$. Consider the first three stages of the following program:

$$\begin{aligned} x(0) &= (\tfrac{1}{2} - \phi, \tfrac{1}{2} + \phi, 0, 0) \\ x(1) &= (\tfrac{1}{2}, \tfrac{1}{2} - \phi, \phi, 0) \\ x(2) &= (\tfrac{1}{2} - \phi, \tfrac{1}{2}, 0, \phi) \end{aligned}$$

Observe that $x(0) \notin V$ and that on the first two stages the value loss is zero. Furthermore, observe that no matter what harvest is imposed on the third stage, there will be a positive value loss due to the fact that $x_4(2) > 0$. This suggest a similar but stronger condition that implies $x \in V$.

Lemma 5.1 *If* $\sum_{t=0}^{n-1}\delta\big(x(t),x(t{+}1)\big)=0 \Longrightarrow x(t)\in V$ *for all* $t=0,\dots,\sigma$.

Proof. $\delta\big(x(t),x(t{+}1)\big)=0$ implies that $\lambda\big(x(t),x(t{+}1)\big)\in S_c$ which in particular means

$$c_i(t)=0 \text{ for all } i>\sigma \quad \text{and} \quad x_n(t)=0. \tag{13}$$

We claim that this implies that $x_i(t)=0, i>\sigma, t\le\sigma$. Indeed, if there is $x_i(t)>0$ with $i>\sigma$ and $t\le\sigma \Longrightarrow x_{i+1}(t+1)>0 \Longrightarrow x_n(n+t-i)>0$ which contradicts (13).

We know then that $x(\sigma)=\big(x_1(\sigma),\dots,x_\sigma(\sigma),0,\dots,0\big)$ where $x_i(\sigma)=\sum_{j=1}^{n}c_j(\sigma-i)=c_\sigma(\sigma-i)\in S_f$ which gives $x(\sigma)\in V$.

Finally, it is easy to see that $x(t+1)\in V$ and $\delta\big(x(t),x(t{+}1)\big)=0$ implies $x(t)\in V$. Then the lemma follows by backwards induction. ■

Consider the lower semicontinuous function $\gamma_n(\{x(t)\}_{t=0}^{n})=\sum_{t=0}^{n-1}\delta\big(x(t),x(t{+}1)\big)$, defined in Proposition 4.3. By the lemma above, we know $\gamma_n(\{x(t)\}_{t=0}^{n})=0 \Longrightarrow x(t)\in V, t=0,\dots,\sigma$. Furthermore,

Lemma 5.2 *Given* $\epsilon>0$, *there is* $\delta>0$ *such that if dist*$\big(x,V\big)\ge\epsilon$ *then we have* $\gamma_n(x,x_1,\dots,x_n)\ge\delta$ *for all* $(x_1,\dots,x_n)\in\Delta^n$.

Proof. Suppose, contrary to our claim, that for every k there is $x^k(0)$ such that $\operatorname{dist}(x^k(0),V)\ge\epsilon$ and $\gamma_n(x^k(0),x^k(1),\dots,x^k(n))\le\frac{1}{k}$ for at least one $(n{-}1)$-tuple $(x^k(1),\dots,x^k(n))$. Of course, $(x^k(i),x^k(i{+}1))$ belongs to the compact set Ω, otherwise $\gamma_n(x^k(0),x^k(1),\dots,x^k(n))=\infty$. We know that there must be at least one converging subsequence: $(x^{k_j}(0),x^{k_j}(1),\dots,x^{k_j}(n))$. Let $(\bar{x}(0),\bar{x}(1),\dots,\bar{x}(n))$ be its limit, by the lower semicontinuity of γ_n we have that

$$\gamma_n(\bar{x}(0),\bar{x}(1),\dots,\bar{x}(n))\le\liminf_j \gamma_n(x^{k_j}(0),x^{k_j}(1),\dots,x^{k_j}(n))=0$$

and the proposition above implies that $\bar{x}(0)\in V$.

On the other hand, $\operatorname{dist}\big(x^k(0),V\big)\ge\epsilon$ for all n then $\operatorname{dist}(\bar{x}(0),V)\ge\epsilon$ and a contradiction arises proving the lemma. ■

We are now in position of proving the important convergence result previously announced,

Lemma 5.3 *Every good program* $\{x(t)\}$ *is such that dist*$(x(t),V)\to 0$.

Proof. It is evident that if $\{x(t)\}$ is a good program, then $\gamma_n(\{x(T+t)\}_{t=0}^{n})\to 0$ when $T\to\infty$. This convergence together with the lemma above implies that $\operatorname{dist}(x(T),V)\to 0$.

Corollary 5.1 *If* $\{x(t)\}$ *is a good program, then dist*$(c(t),S_c)\to 0$.

Finally, we are led to the principal result of this section: the following strengthening of [32, Lemma 6.4] to our non-concave, non-differentiable context.

Theorem 5.1 $\{\hat{x}\}=V$ *iff any good program satisfies* $\lim_t x(t)=\hat{x}$.

Proof. The first implication follows directly from the lemma above. To prove the converse, it suffices to observe that if there is $x\ne\hat{x}$ such that $x\in V$, then there is zero value loss $\sigma-$periodic program originating from x. Evidently, this program is good and does not converge to $\hat{x}$. ■

Given the last theorem it is of interest to know when Condition 5.1: $V=\{\hat{x}\}$ holds. We look for conditions on the primitives of the model assuring that $V=\{\hat{x}\}$. Evidently, $S_f=\{1/\sigma\}$ is sufficient to assure it. Even more, due to the area balance, conditions

$$1/\sigma=\min\{\xi:\xi\in S_f\} \quad \text{and} \quad 1/\sigma=\max\{\xi:\xi\in S_f\}$$

are also sufficient. All these can be expressed as the Non-interiority Condition 5.2

$$\tfrac{1}{\sigma}\notin \operatorname{int}\operatorname{co}(S_f).$$

Lemma 5.4 *If the Non-interiority Condition 5.2 holds then* $\{\hat{x}\} = V$*.*

Remark 5.2 *The condition above is not necessary as the following toy example shows. Consider a forest where* $n = 3$*,* $\sigma = 2$ *and* $S_f = \{\frac{1}{3}, \frac{1}{2}, \frac{3}{4}\}$*. Of course* $\frac{1}{\sigma} \in (\frac{1}{3}, \frac{3}{4}) = int\ co(S_f)$*, so the condition above is not fulfilled. But the set* V *is*

$$V = \{x \in \Delta : x_1, x_2 \in S_f \text{ and } x_3 = 0\} = \{(x_1, x_2, 0) : x_1 + x_2 = 1, x_1, x_2 \in S_f\} = \{\hat{x}\}.$$

If the felicity function is concave, then the non-interiority Condition 5.2 turns into the simpler Condition 5.3

$$\tfrac{1}{\sigma} \notin \text{int } (S_f)$$

which turns out to be necessary and sufficient.

Lemma 5.5 *Let the felicity function* w *be concave. Then, the Non-interiority Condition 5.3 holds if and only if* $\{\hat{x}\} = V$*.*

Proof. The sufficiency of the non-interiority Condition 5.3 follows directly from Lemma 5.4.

To see the necessity, suppose that it does not hold. Then there is $\phi > 0$ such that $[\frac{1}{\sigma} - \phi, \frac{1}{\sigma} + \phi] \subset S_f$ and

$$x = \left(\tfrac{1}{\sigma} - \phi, \tfrac{1}{\sigma} + \phi, \tfrac{1}{\sigma}, \ldots, \tfrac{1}{\sigma}, 0, \ldots, 0\right) \in V.$$

■

Theorem 5.2 *Let* w *be concave non-differentiable. Then the Non-interiority Condition 5.3 holds if and only if any good program* $\{x(t)\}$ *satisfies* $\lim_{t\to\infty} x(t) = \hat{x}$*.*

Proof. If Non-interiority Condition 5.3 holds, the lemma above yields $\{\hat{x}\} = V$ and then Lemma 5.3 implies that any good program satisfies $x(t) \to \hat{x}$. If it does not hold, then there exists $x \in V$, $x \neq \hat{x}$, and the σ-periodic program from x is a non-converging zero value loss program. ■

6 On Maximal and Optimal Programs

Next, we turn to the existence results for optimal programs, and also present an equivalence theorem that shows the equivalence of optimal, maximal and minimal value-loss programs. We distinguish carefully how far one can proceed without the Condition 5.1, and indicate, as in the sections above, versions of our results that are already available in the literature.

First, we need a formal definition of maximal programs. In Section 2, the notion of an optimal program is defined for the general model, and on using Equation (3), Definition 2.2 can be routinely transferred to the forestry model. We can then present the two basic concepts considered in this section.

Definition 6.1 *A program* $\{x^*(t)\}$ *is* optimal *if for any program* $\{x(t)\}$ *such that* $x(0) = x^*(0)$ *we have*

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

Definition 6.2 *A program* $\{x^*(t)\}$ *is* maximal *if for any program* $\{x(t)\}$ *such that* $x(0) = x^*(0)$ *we have*

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

Let us first see an easy technical lemma.

Lemma 6.1 *Every maximal program is good.*

Proof. Let $\{x^*(t)\}$ be a maximal program from x_0 and $\{x(t)\}$ *any* good program from x_0, i.e., there is $M \in I\!R$ such that for all $T \geq 0$, $\lim_T \sum_{t=0}^{T-1} w\big(bc(t)\big) - w(\frac{b_\sigma}{\sigma}) \geq M$. Hence, we get

$$\begin{aligned} \textstyle\sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) &= \textstyle\sum_{t=0}^{T-1} w(bc(t)) - w(\frac{b_\sigma}{\sigma}) - \sum_{t=0}^{T-1} w(bc^*(t)) - w(\frac{b_\sigma}{\sigma}) \\ \textstyle\sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) &\geq M - \textstyle\sum_{t=0}^{T-1} w(bc^*(t)) - w(\frac{b_\sigma}{\sigma}) \end{aligned}$$

To obtain a contradiction suppose that $\{x^*(t)\}$ is bad, letting $T \to \infty$ we get $0 \geq \liminf_T \sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) \geq M - (-\infty)$ which is obviously absurd. ∎

Next, we present a well-known result in a more general framework.[33] Brock's proof is not directly applicable since he utilizes the fact that convergence of the felicity function implies convergence of their arguments, and the benefit function used here is defined on the harvested timber levels rather than on the forest configurations. The proof presented by Mitra-Wan is based on the concavity and differentiability of the felicity function w and cannot be used in our context, either.[34] We present an alternative proof that relays heavily on the linearity of the function $\lambda\big(x(t), x(t+1)\big)$.

Lemma 6.2 *The Cesàro means of every good program $\{x(t)\}$ converge to the golden-rule forest configuration; namely,*

$$\bar{x}(t) = \tfrac{x(0)+\ldots+x(t-1)}{t} \to \hat{x} \quad \textit{when } t \to \infty. \tag{14}$$

Proof. We first observe that the convexity of Ω implies $(\bar{x}(t), \bar{x}'(t)) = \big(\frac{x(0)+\ldots+x(t-1)}{t}, \frac{x(1)+\ldots+x(t)}{t}\big) \in \Omega$. Let $\bar{x}$ be any accumulation point of $\{\bar{x}(t)\}$. It is easy to see that if $\bar{x}(t_k) \to \bar{x}$ then

$$x'(t_k) = \tfrac{x(1)+\ldots+x(t_k)}{t_k} = \bar{x}(t_k) + \tfrac{x(t_k)-x(0)}{t_k} \to \bar{x} + 0$$

hence, $(\bar{x}, \bar{x})$ is an accumulation point of $\big(\frac{x(0)+\ldots+x(t-1)}{t}, \frac{x(1)+\ldots+x(t)}{t}\big)$ and in consequence

$$(\bar{x}, \bar{x}) \text{ belongs to the closed set } \Omega. \tag{15}$$

Furthermore, and thanks to the linearity of $\lambda(x, x')$, we get that $\lambda\big(\bar{x}(t), \bar{x}'(t)\big) = \frac{1}{t}\sum_{l=0}^{t-1} \lambda\big(x(t), x(t+1)\big)$.

Let $\{x(t)\}$ be a good program and $\bar{x}(t_k)$ be a subsequence converging to $\bar{x}$, we know that

$$\textstyle\sum_{l=0}^{t-1} w\big(bc(l)\big) - w(\frac{b_\sigma}{\sigma}) \geq M \qquad \text{for all } t \implies \lim_k \frac{1}{t_k} \sum_{l=0}^{t_k-1} w\big(bc(l)\big) \geq w(\frac{b_\sigma}{\sigma}).$$

Hence, we have that $(\bar{x}, \bar{x}) \in \Omega$ and $w(\lambda(\bar{x}, \bar{x})) = \lim_k w(\lambda(\bar{x}(t_k), \bar{x}'(t_k))) \geq w(\frac{b_\sigma}{\sigma})$, implying that $\bar{x} \in S(P) = \{\hat{x}\}$. ∎

The existence of a maximal program is furnished by the following result.

Theorem 6.1 *Any program $\{x^*(t)\}$ from $x_0 \in \Delta$ that minimizes accumulated value loss is maximal. Consequently, there exists a maximal program from each initial state $x_0 \in \Delta$.*

Proof. First observe that

$$\begin{aligned} \textstyle\sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) &= -\textstyle\sum_{t=0}^{T-1} \delta\big(x(t), x(t+1)\big) - \hat{p}\ \big(x(T) - x_0\big) \\ &\quad + \textstyle\sum_{t=0}^{T-1} \delta\big(x^*(t), x^*(t+1)\big) + \hat{p}\ \big(x^*(T) - x_0\big) \\ &= -\textstyle\sum_{t=0}^{T-1} \delta\big(x(t), x(t+1)\big) + \sum_{t=0}^{T-1} \delta\big(x^*(t), x^*(t+1)\big) + \hat{p}\ \big(x^*(T) - x(T)\big) \end{aligned} \tag{16}$$

[33] This result is given in [2, Lemma 4] and in [32, Lemma 4.3]. Also see Footnote 24 in this connection.

[34] Furthermore, the proof of Lemma 4.3 in [32] relies on their Lemma 3.1, whose proof is phrased in terms of the differentiability of the benefit function; see their equation (3.7).

Suppose, contrary to our claim, that $\{x^*(t)\}$ is not maximal. Then, there are $\epsilon_0 > 0$, T_0 and a program $\{x(t)\}$ from x_0 such that $\sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) > \epsilon_0$ for all $T \geq T_0$. Since $\{x^*(t)\}$ is good, the alternative program $\{x(t)\}$ must also be good (otherwise, the former inequality could not be hold for all $T \geq T_0$).

By (16) and the minimality of $\sum \delta^*$ there is $\epsilon \in (0, \epsilon_0)$ and $T_1 \geq T_0$ such that for all $T \geq T_1$,

$$\epsilon < \hat{p}\ (x^*(T) - x(T)).$$

Now, we may use Lemma 6.2 to obtain

$$\epsilon < \liminf_T \hat{p}\ (\bar{x}^*(T) - \bar{x}(T)) = 0,$$

and a contradiction arises. Thanks to Proposition 4.3, we know that given x_0, there is always a program such that $\mu(x(0)) = \sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big)$, and hence this program is maximal. ■

We see next a result that complements the theorem above.

Theorem 6.2 *If $\{x^*(t)\}$ is an optimal program from x_0 then it minimizes accumulated value loss, i.e., $\sum_{t=0}^{\infty} \delta\big(x^*(t), x^*(t+1)\big) = \mu(x_0)$.*

Proof. Suppose, contrary to our claim, that the program $\{x^*(t)\}$ does not minimize the accumulated value loss and let $\{x(t)\}$ be a minimizer. Hence, there exists $\epsilon_0 > 0$ such that

$$\sum_{t=0}^{\infty} \delta\big(x^*(t), x^*(t+1)\big) - \sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big) > \epsilon_0.$$

And given $\epsilon \in (0, \epsilon_0)$ there is T_0 such that

$$\sum_{t=0}^{T-1} \delta\big(x^*(t), x^*(t+1)\big) - \sum_{t=0}^{T-1} \delta\big(x(t), x(t+1)\big) > \epsilon \quad \text{for all } T \geq T_0.$$

By (16) we deduce

$$\sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) > \epsilon + \hat{p}\ (x^*(T) - x(T)) \qquad \text{for all } T \geq T_0.$$

We know as well that there is T_1 such that

$$\sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) < \limsup_T \sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) + \epsilon/2 \qquad \text{for all } T \geq T_1.$$

From the last two inequalities we get

$$\limsup_T \sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) > \epsilon/2 + \hat{p}\ (x^*(T) - x(T)) \qquad \text{for all } T \geq \max\{T_0, T_1\}$$

which readily implies that

$$\limsup_T \sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) > \epsilon/2 + \lim_T \hat{p}\ (\bar{x}^*(T) - \bar{x}(T)) = \epsilon/2$$

contradicting the optimality of $\{x(t)\}$. ■

Our next set of results are a testimony to the power of the non-interiority condition 5.2. When it holds then maximal and optimal programs coincide and in consequence, a program is optimal if and only if it minimizes accumulated value-loss.

Proposition 6.1 *Assume that $V = \{\hat{x}\}$. Then every maximal program is optimal.*

Proof. Let $\{x^*(t)\}$ be a maximal program from x_0 and $\{x(t)\}$ any other program from x_0.

Consider first the case where $\{x(t)\}$ is bad: by (16) and using that $\{x^*(t)\}$ is good, we deduce that

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

If the alternative program is good then we know that $\lim_T \sum_{t=0}^{T} \delta\big(x(t), x(t+1)\big)$ is well defined, as well as $\lim_T \sum_{t=0}^{T} \delta\big(x^*(t), x^*(t+1)\big)$ and also that $\lim_T x^*(T) = \lim_T x(T) = \hat{x}$, because (5.2) holds. Then, considering (16) again and letting $T \to \infty$ we get that the limit of the right hand side is defined and hence it is the limit of the left hand side:

$$\limsup_T \sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) = \liminf_T \sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) \leq 0.$$

■

Corollary 6.1 *Assume that $V = \{\hat{x}\}$. Then there exists an optimal program $\{x(t)\}$ from any initial state $x_0 \in \Delta$.*

From Theorem 6.1 and 6.2 and Proposition 6.1, we directly obtain the following equivalence result. The equivalence of (ii) and (iii) for the dual-aged farm, and under the hypothesis of strict concavity of the benefit function, is available in [29, Sections 2.1 and 2.2].[35]

Theorem 6.3 *Assume that $V = \{\hat{x}\}$ and that $\{x(t)\}$ is program from x_0. Then the following are equivalent: (i) $\{x(t)\}$ is optimal, (ii) $\sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big) = \mu(x(0))$, (iii) $\{x(t)\}$ is maximal.*

7 On Finitely Maximal Programs

In [28], Mitra has given an axiomatic underpinning to the following notion of an optimal program. The reader should note that for the two theorems in this section, the upper semicontinuity hypothesis on the benefit function w is strengthened to continuity.

Definition 7.1 *A program $\{x^*(t)\}$ is* finitely maximal *if for any natural number $T \geq 1$, and any program $\{x(t)\}$ such that $x(0) = x^*(0)$, $x(t) = x^*(t)$ for all $t \geq T$,*

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

Let us first see an easy technical lemma that is an analogue of Lemma 6.1 and is original to [28, Proposition 2]. The proof we present follows the same lines as in [28], but adapted to fit our notation and to rely on programs constructed in [18]. These programs allow the planner to move from an arbitrary forest configuration to a golden-rule configuration in σ periods, and also from a golden-rule forest configuration to an arbitrary forest configuration in n periods.

Lemma 7.1 *Every finitely maximal program is good.*

Proof. Let $G = -(\sigma + n + 1)w(\frac{b_\sigma}{\sigma})$. Let $\{x^*(t)\}$ be a finitely maximal program. We can assert that for all $T \geq 1$ $\sum_{t=0}^{T}(w(bc^*(t)) - w(\frac{b_\sigma}{\sigma})) \geq G$. Suppose instead that there exists $T \geq 1$ such that

$$\sum_{t=0}^{T}(w(bc^*(t)) - w(\tfrac{b_\sigma}{\sigma})) < G. \quad (17)$$

[35] Note, in keeping with Footnotes 12 and 6, optimality in [28] is our notion of maximality. Such an equivalence result is also available in [44] for the RSS model under alternative hypotheses pertaining to that model.

Since $w(\cdot)$ is a non-negative function, $T > (\sigma + n + 1)$. Now consider the following $(T+1)$ tuple:

$$\bar{x}(t) = \begin{cases} x^*(0) & \text{if } t = 0 \\ x(t) & \text{if } t = 1, \cdots, \sigma - 1, \\ \hat{x} & \text{if } t = \sigma, \sigma + 1, \cdots, T - n \\ x(t) & \text{if } t = T - (n+1), \cdots, T-1, \\ x^*(t) & \text{if } t = T, \end{cases}$$

where $(x(1), \cdots, x(\sigma - 1))$ is as constructed in the proof of [18, Proposition 5.5], and $\big(x(T-n), \cdots, x(T)\big)$ is as constructed in the proof of [18, Lemma 6.2].

Now define the infinite sequence $\{x(t)\}$ such that

$$x(t) = \begin{cases} \bar{x}(t) & \text{if } t = 0, \cdots, T-1, \\ x^*(t) & \text{if } t > T. \end{cases}$$

Certainly $\{x(t)\}$ is a program that starts from the same initial stock as the given finitely maximal program, and differs from it only for the subsequent $(T-1)$ periods. But we now obtain

$$\textstyle\sum_{t=0}^{T} w(bc(t)) - w(bc^*(t)) = \sum_{t=0}^{T}(w(bc(t)) - w(\frac{b_\sigma}{\sigma})) - \sum_{t=0}^{T}(w(bc^*(t)) - w(\frac{b_\sigma}{\sigma})).$$

The first term on the right hand side is greater than G by virtue of the fact that $w(\cdot)$ is a non-negative function, and the second term is greater than $-G$ by virtue of (17). We thus obtain the fact that the left hand side is positive, and thereby contradict the fact that $\{x^*(t)\}$ is a finitely maximal program. ■

We see next a result that complements Theorem 6.2 above, but under the standing assumption 5.1 which simply guarantees, by virtue of Theorem 5.1, the asymptotic convergence of good programs. The essence of the argument revolves around the programs utilized in the proof of Lemma 7.1, but with the additional refinement that they make arbitrary small value-losses if the initial and terminal forest configurations in question are arbitrarily close to the golden-rule configuration.[36]

Theorem 7.1 *Assume that $V = \{\hat{x}\}$ and that w is continuous. Then a finitely maximal program $\{x^*(t)\}$ from x_0 minimizes accumulated value loss, i.e., $\sum_{t=0}^{\infty}\delta\big(x^*(t), x^*(t+1)\big) = \mu(x_0)$.*

Proof. Suppose, contrary to our claim, that the finitely maximal program $\{x^*(t)\}$ does not minimize the accumulated value loss and let $\{x(t)\}$ be a minimizer. Hence, there exists $\epsilon_0 > 0$ such that

$$\textstyle\sum_{t=0}^{\infty}\delta\big(x^*(t), x^*(t+1)\big) - \sum_{t=0}^{\infty}\delta\big(x(t), x(t+1)\big) > \epsilon_0.$$

And given $\epsilon \in (0, \epsilon_0)$ there is T_0 such that

$$\textstyle\sum_{t=0}^{T-1}\delta\big(x^*(t), x^*(t+1)\big) - \sum_{t=0}^{T-1}\delta\big(x(t), x(t+1)\big) > \epsilon \quad \text{for all } T \geq T_0. \tag{18}$$

Now, for $(\epsilon/4)$, by [18, Proposition 5.5], there exists δ_1 such that for all $x \in \Delta$ satisfying $\|x - \hat{x}\|_\infty < \delta_1$, there exists a program $\{x^i(t)\}$ from x, and differing from $\hat{x}$ only for the first σ periods, for which $\sum_{t=0}^{\sigma-1}\delta\big(x^i(t), x^i(t+1)\big) < \epsilon/4$.

Furthermore, for $(\epsilon/4)$, by [18, Lemma 6.2], there exists δ_2 such that for all $x \in \Delta$ satisfying $\|x - \hat{x}\|_\infty < \delta_2$, there exists a $(n+1)$-program from $\hat{x}$, $\{\bar{x}(t)\}_{t=0}^{n+1} = (\hat{x}, \cdots, x)$, such that $\sum_{t=0}^{n}\delta\big(\bar{x}(t), \bar{x}(t+1)\big) < (\epsilon/4)$.

For our next two claims, we shall use the non-interiority condition 5.2 to ensure, by virtue of Theorem 5.2, the asymptotic convergence of all good programs.

Since every minimal value-loss program is good by virtue of Proposition 4.2, there exists T_1 such that $\|x(t) - \hat{x}\|_\infty < \delta_1$ for all $T \geq T_1$.

[36] This construction can be seen as one of the principal contributions of [18].

Since the finitely maximal program $\{x^*(t)\}$ is a good program by virtue of Lemma 7.1, there exists T_2 such that $\|x^*(t) - \hat{x}\|_\infty < \delta_2$ for all $T \geq T_2$.

Now let $T_3 = \max\{T_0, T_1, T_2\}$, and define the infinite sequence $\{x^m(t)\}$ such that

$$x^m(t) = \begin{cases} x(t) & \text{if } t = 0, \cdots, T_3, \\ x^i(t) & \text{for } t = T_3 + 1, \cdots, T_3 + \sigma, \\ \bar{x}(t) & \text{for } t = T_3 + \sigma + 1, \cdots, T_3 + \sigma + n - 1 \\ x^*(t) & \text{for all } t \geq T_3 + \sigma + n \end{cases}$$

Now, on re-writing (18) with $T = (T_3 + 1)$, we obtain

$$\textstyle\sum_{t=0}^{T_3} \delta\big(x^*(t), x^*(t+1)\big) - \sum_{t=0}^{T_3} \delta\big(x(t), x(t+1)\big) > \epsilon.$$

Since $\delta\big(x^*(t), x^*(t+1)\big)$ is non-negative for all t, we obtain

$$\textstyle\sum_{t=0}^{T_4} \delta\big(x^*(t), x^*(t+1)\big) - \sum_{t=0}^{T_3} \delta\big(x(t), x(t+1)\big) > \epsilon \text{ where } T_4 = (T_3 + \sigma + n). \tag{19}$$

Finally, since

$$\textstyle\sum_{t=T_3+1}^{T_4} \delta\big(x^m(t), x^m(t+1)\big) < (\epsilon/2),$$

we obtain from (19),

$$\textstyle\sum_{t=0}^{T_4} \delta\big(x^*(t), x^*(t+1)\big) - \sum_{t=0}^{T_3} \delta\big(x(t), x(t+1)\big) - \sum_{t=T_3+1}^{T_4} \delta\big(x^m(t), x^m(t+1)\big) > (\epsilon/2).$$

Since the program $\{x^m(t)\}$ starts from x_0 and is identical to the program $\{x^*(t)\}$ for all $t \geq T_4$, we obtain

$$\textstyle\sum_{t=0}^{T_4} w(bc^m(t)) - w(bc^*(t)) > (\epsilon/2).$$

But this contradicts the assertion that $\{x^*(t)\}$ is a finitely maximal program. ■

We can now turn to the unification that has motivated this essay. Theorem 7.1 allows us to strengthen Theorem 6.3 into the following equivalence result.

Theorem 7.2 *Assume that $V = \{\hat{x}\}$, that w is continuous, and that $\{x(t)\}$ is program from x_0. Then the following are equivalent: (i) $\{x(t)\}$ is optimal, (ii) $\sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big) = \mu(x(0))$, (iii) $\{x(t)\}$ is maximal, (iv) $\{x(t)\}$ is finitely maximal.*

The equivalence of (iii) and (iv), under the hypothesis of strict concavity of the benefit function, is available in [28, Theorem 7].[37] In the introduction, we have already remarked under our different method of proof. The reader is also invited to compare Theorem 2.2 and the analogous results in [47].

8 Concluding Remarks

The non-interiority condition presented in this paper has served as a synthesizing criterion for the analysis of the Mitra-Wan tree farm by ensuring the asymptotic convergence of good programs, and by being both a necessary *and* sufficient condition for such a convergence when the benefit function is concave. Premised on such an asymptotic convergence, a theory of undiscounted dynamic programming is developed for a general intertemporal model in Dana-Le Van [8, 9], for the 2-sector RSS model in Khan-Mitra [17], and for the dual-aged forest in [29]. All of these results are set in the context of strictly concave felicity (benefit) functions, and there is little doubt that an analogue of the non-interiority condition reported here will also serve to move this theory towards completion by an extension to the larger class of functions isolated here. We also leave as an open problem the identification of a necessary and sufficient analogue that takes the general

[37] Note, in keeping with Footnotes 12 and 6, optimality in [28] is our notion of maximality. Note also Footnote 1.

theory of intertemporal resource allocation, developed and reported in [22, 23, 24] to such a larger class of felicity functions. A theory generalized along both of these lines should be of especial use in considerations that go beyond the deterministic to the stochastic context.[38]

It is worth reiterating that our focus in this paper, following the original conception of Mitra-Wan [32], has been on infinite horizon optimal programs without discounting, which is to say, on the undiscounted long-run. We already know from [31], [10] and [40] the difficulties of delineating optimal transition dynamics in the discounted setting, and whereas a full understanding of the model can hardly be had without a resolution of these difficulties, substantial analysis of the undiscounted case that is feasible remains to be done. More importantly, and given the recent emphasis on numerical computation,[39] the proximity of finite horizon optimal programs to their infinite-horizon counterparts (as in [19] for the RSS model and [18] for the Mitra-Wan model), and questions of their sensitivity to initial and terminal stocks (as in [3] and [26]), remain to be investigated in the setting considered in this paper. We leave this for future work.[40]

Indeed, the undiscounted setting has special reference to environmental economics: discounting by a planner makes even less sense when issues of climate change are in question or the value of a forest goes beyond its timber yield and takes environmental well-being into account.[41] This leads to a situation where the stock variables are arguments in the benefit functions, as in [34, 35]. Indeed, as emphasized in [4, p. 453] and [14], the subject naturally leads into intergenerational and intertemporal equity issues. Overlapping generations, with each succeeding generation having sole property rights to the forest, will lead to a complementary conception in which, unlike the perishability of Samuelson's chocolates, the commodity has durability over a finite number of generations,[42] and it is certainly of interest to see how the arguments developed in this paper fare for such a setting. This too, we leave for future work.

But moving beyond a general theory of undiscounted dynamic programming, and beyond a general statement regarding applications to deterministic and stochastic discrete-time dynamic systems in capital theory and environmental economics, it is clear that the non-interiority condition presented in this paper, and specifically Theorem 5.2, can be used to generalize the existence and asymptotic results for maximal and optimal programs in the RSS model presented in [16, 44], and for the dual-aged Mitra-Wan forest in [29]. It is clear that the non-interiority condition will also serve as a synthesizing criterion for contexts (other than the Mitra-Wan forestry model) catalogued in [27]. We leave these verifications as exercises for the interested reader.

9 Appendix

In [28], for any concave function $u : X \to I\!R$, with X a convex subset of a linear space, Mitra has defined a strict mid-concave function as one where for x_0, x_1 in X, and $x_0 \neq x_1$,

$$u(x_{\frac{1}{2}}) > \tfrac{1}{2}u(x_0) + \tfrac{1}{2}u(x_1). \tag{20}$$

We make the simple observation that this is tantamount to the assumption of strict concavity of u. Let $x_\lambda = \lambda x_0 + (1-\lambda)x_1$. We need to show that

$$u(x_\lambda) > \lambda u(x_0) + (1-\lambda)u(x_1) \qquad \text{for all } \lambda \in (0,1). \tag{21}$$

We consider first the case where $\lambda \in (0, \frac{1}{2})$. We can write x_λ as the convex combination of $x_{\frac{1}{2}}$ and x_1,

$$x_\lambda = \lambda x_0 + (1-\lambda)x_1 = \lambda(x_0 + x_1) + (1-2\lambda)x_1 = 2\lambda\ x_{\frac{1}{2}} + (1-2\lambda)x_1$$

[38] See [5] and [25] where the growth of a tree is modeled as a Weiner process. A generalization of the Mitra-Wan forestry model to more standard Brock-Mirman type models in the growth and uncertainty literature also remains to be accomplished.

[39] See the textbook of Judd [13] and his references.

[40] Arkin-Evstegneev [1] and Zaslavski [45] are comprehensive references to this set of issues.

[41] See Brock [4] and Chichilnisky et al. [6] for an emphasis on the undiscounted setting in the context of environmental and resource economics.

[42] For references and recent work on the OLG model, see [43] and [20].

Now, the concavity of u assures that $u(x_\lambda) \geq 2\lambda\ u(x_{\frac{1}{2}}) + (1-2\lambda)u(x_1)$ and thanks to (20) we know that

$$\begin{aligned} u(x_\lambda) &\geq 2\lambda\ u(x_{\frac{1}{2}}) + (1-2\lambda)u(x_1) \\ &> 2\lambda\left(\tfrac{1}{2}u(x_0) + \tfrac{1}{2}u(x_1)\right) + (1-2\lambda)u(x_1) \\ &= \lambda u(x_0) + (1-\lambda)u(x_1) \end{aligned}$$

proving (21) for $\lambda \in (0, \frac{1}{2})$.

Next, to deal with the case where $\lambda \in (\frac{1}{2}, 1)$, we write x_λ as the convex combination of x_0 and $x_{\frac{1}{2}}$: $x_\lambda = (2\lambda - 1)x_0 + \big(1 - (2\lambda - 1)\big)x_{\frac{1}{2}}$. And repeating the steps above, we prove (21) for $\lambda \in (\frac{1}{2}, 1)$.

References

[1] Arkin, V. I. and Evstigneev, I. V. 1987. Stochastic Models of Control and Economic Dynamics. Academic Press, New York.

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

[3] Brock, W. A., 1971. Sensitivity of optimal growth paths with respect to a change in target stock. Zeitschrift für Nationalökonomie, Supplement, 1, 73-89.

[4] Brock, W. A., 1977. A polluted golden age. Chapter 25 in Smith, V. L. (ed.), Economics of natural and environmental resources, Gordon and Breach, New York.

[5] Brock, W. A., Rothschild, M. and Stiglitz, J., 1989. Sochastic capital theory. In Feiwel, G. (ed)., Joan Robinson and Modern Economic Theory, New York University Press, New York, 591-622.

[6] Chichilnisky, G., Heal, G. and Beltratti, A., 1995. The green golden rule. Economics Letters 49, 175-179.

[7] Cominetti R. and Piazza A., 2006. Asymptotic convergence of optimal harvesting policies for a multiple species forest. Mathematics of Operations Research. Forthcoming. Also circulated as *Publicación Técnica CMM.* No. 158, 1-18.

[8] Dana, R. A. and C. Le Van, 1990. On the Bellman equation of the overtaking criterion. Journal of Optimization Theory and Applications 78, 605612. Addendun in the same journal, 67 (1993), 587-600.

[9] Dana, R. A., Le Van, C., Mitra, T. and Nishimura, K. (eds.) (2006). Handbook of Optimal Growth, Volume 1. Springer, Berlin.

[10] Dasgupta, S. and Mitra, T., 2002. "Optimal Policy Function in the Forestry Model: An Example", mimeo, Cornell University.

[11] Gale D., 1967. On optimal development in a multi-sector economy, Review of Economic Studies 37, 1-18.

[12] Hartl, R. F. and Kort, P., 2003. History dependence without unstable steady state: a non-differentiable framework. Journal of Mathematical Economics 39, 891-900.

[13] Judd K. L. 1998. Numerical Methods in Economics. The MIT Press, Cambridge.

[14] Khan, M. Ali, 2005. Intertemporal ethics, modern capital theory and the economics of forestry. Chapter 2 in Kant, S. and Berry, R. A. (eds.) Economics, Sustainability, and Natural Resources: Economics of Sustainable Forest Management, Springer, Berlin, 39-65.

[15] Khan, M. Ali and Mitra, T., 2002. Optimal growth under irreversible investment: the von Neumann and Mckenzie facets. Mimeo, Cornell University.

[16] Khan, M. Ali and Mitra, T., 2005. On choice of technique in the Robinson-Solow-Srinivasan model. International Journal of Economic Theory 1, 83-109.

[17] Khan, M. Ali and Mitra, T., 2006. Undiscounted optimal growth under irreversible investment: a synthesis of the value-loss approach and dynamic programming. Economic Theory 29, 341-362.

[18] Khan M. Ali and Piazza A., 2009. On uniform convergence of undiscounted optimal programs in the Mitra-Wan forestry model: the strictly concave case. *Publicación Técnica CMM.* No. 222, 1-15.

[19] Khan M. Ali and Zaslavski, A. J., 2008. On two classical turnpike results for the Robinson-Solow-Srinivasan (RSS) model. Mimeo, The Johns Hopkins University.

[20] Li, L. and Lin, S., 2008. Existence and uniqueness of steady-state equilibrium in a two-sector overlapping generations model. Journal of Economic Theory, 141, 255-275.

[21] Majumdar M., Mitra T. and Nishimura K., (eds.), 2000. Optimization and Chaos. Studies in Economic Theory, 11. Springer-Verlag, Berlin.

[22] McKenzie, L. W., 1968. Accumulation programs of maximum utility and the von Neumann facet. In Wolfe, J. N. (ed.), Value, Capital and Growth, 353-383. Edinburgh University Press, Edinburgh.

[23] McKenzie L.W., 1986. Optimal economic growth, turnpike theorems and comparative dynamics. In Arrow, K. J., Intrilligator, M. (eds.), Handbook of Mathematical Economics, Vol. III, 1281-1355. North-Holland, Amsterdam.

[24] McKenzie L.W., 2002. Classical general equilibrium theory. The MIT Press, Cambridge, MA.

[25] Miller, R. A. and Voltaire, K., 1980. A sequential stochastic tree problem. Economics Letters 5, 135-140.

[26] Mitra, T., 1983. Sensitivity of optimal programs with respect to changes in target stocks. Journal of Economic Theory 29, 172-184.

[27] Mitra, T., 2000. Introduction to dynamic optimization theory. In Majumdar, M., Mitra, T., Nishimura, K. (eds.), Optimization and Chaos, 31-108. Springer, New York.

[28] Mitra, T., 2005. Intergenerational Equity and the forest management problem. In Kant, S. and Berry, R. A. (eds.) Economics, Sustainability, and Natural Resources: Economics of Sustainable Forest Management, Springer, Berlin, 137-173.

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

[30] Mitra, T. and Nishimura, K., 2006. On stationary optimal stocks in optimal growth theory: existence and uniqueness results. In Le Van et al. 2006. Handbook of Optimal Growth.

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

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

[33] Mitra, T., Ray, D. and Roy, R., 1991. The economics of orchards: an exercise in point-input flow-output capital theory. Journal of Economic Theory 53, 12-50.

[34] Nyarko, Y. and Olson, L. J., 1991. Stochastic dynamic models with stock dependent rewards. Journal of Economic Theory 55, 161-167.

[35] Nyarko, Y. and Olson, L. J., 1991. Stochastic growth when utility depends on both consumption and the stock level. Economic Theory 4, 791-798.

[36] Piazza, A., 2009, The optimal harvesting problem with a land market: a characterization of the asymptotic convergence. Economic Theory (40)1, 133-138.

[37] Piazza, A., 2009, About optimal harvesting policies for a multiple species forest without discounting, mimeo. Paper to be presented at a President-Elect Featured Session of the *Western Economic Association* Meetings to be held in Vancouver, June 29-July 3.

[38] Rapaport A., Sraidi S. and Terreaux J.P., Optimality of greedy and sustainable policies in the management of renewable resources. Optimal Control Application and Methods 24, 23-44 (2003).

[39] 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.

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

[41] Salo S. and Tahvonen O., 2004. Renewable resources with endogenous age classes and allocation of the land. American Journal of Agricultural Economics 86, 513-530.

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

[43] Wendner, R., 2003. Existence, uniqueness and stability of equilibrium in an OLG economy. Economic Theory 23, 165-174.

[44] Zaslavski A. J., 2005. Optimal programs in the RSS model. International Journal of Economic Theory 1, 151-165.

[45] Zaslavski A. J., 2006. Turnpike properties in the calculus of variations and optimal control. Nonconvex Optimization and its Applications 80. Springer, New York.

[46] Zaslavski, A. J., 2007. Turnpike results for discrete-time optimal control systems arising in economic dynamics. Nonlinear Analysis 67, 2024-2049.

[47] Zaslavski, A. J., 2009. Good locally maximal programs for the Robinson-Solow-Srinivasan model. Advances in Economic Theory, forthcoming.

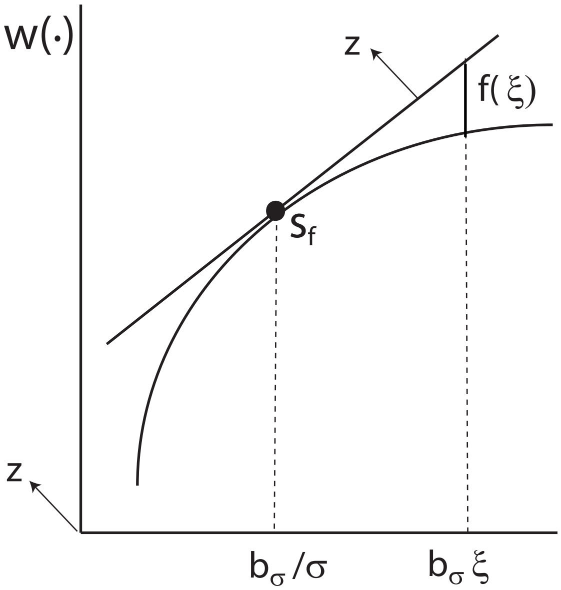

Figure 1a

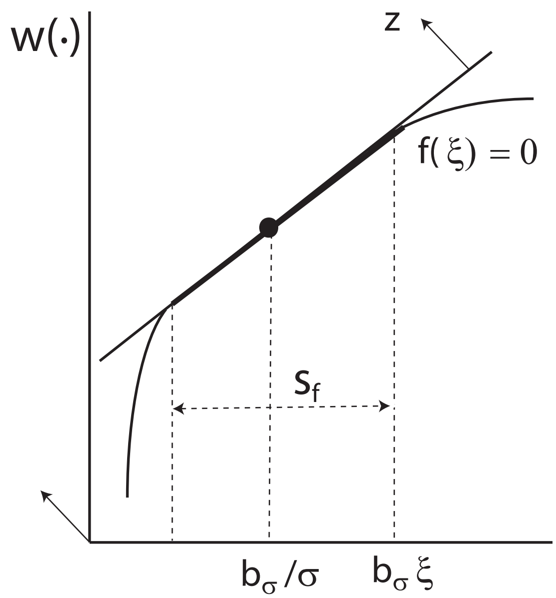

Figure 1b

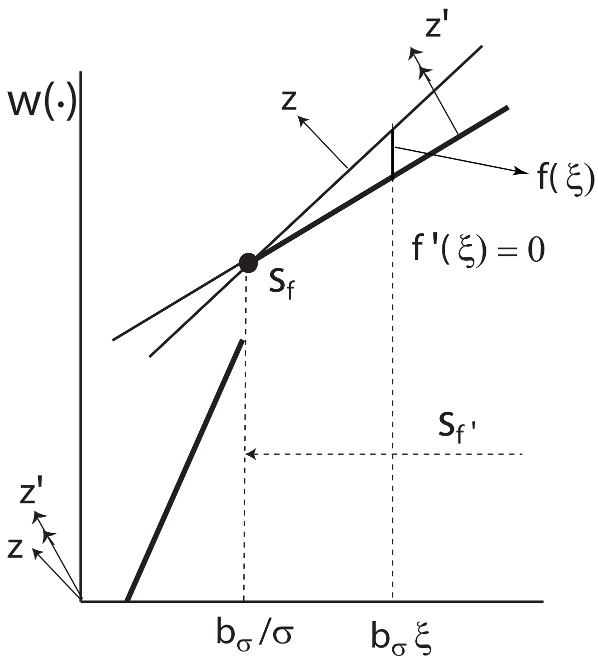

Figure 1c

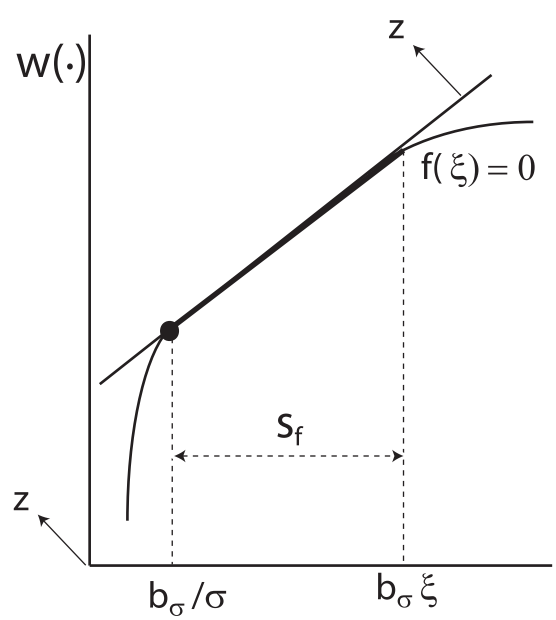

Figure 1d

Figure1: The Discrepancy Function $f(\cdot)$