IJET International Journal of Economic Theory

doi: 10.1111/j.1742-7363.2009.00122.x

On uniform convergence of undiscounted optimal programs in the Mitra–Wan forestry model: The strictly concave case

M. Ali Khan* and Adriana Piazza†

In this paper, we show that the 1986 Mitra–Wan result establishing asymptotic convergence of maximal programs to the unique golden-rule forest in the case of undiscounted, strictly concave felicity functions can be strengthened, in the same setting, to the *uniform* asymptotic convergence of *optimal* programs to the unique golden-rule forest. We work with a notationally reformulated version of the model that may have independent interest.

Key words forest management, maximal programs, optimal programs, uniform asymptotic convergence, value-loss function, von Neumann facet

JEL classification C62, D90, Q23

Accepted 19 September 2009

1 Introduction

In two remarkable papers in the mid-eighties, Mitra and Wan (1985, 1986) reformulated the economics of forest management so as to exploit its connection to the general theory of intertemporal resource allocation in discrete time as developed by Gale (1967) and McKenzie (1968, 1986).[1] Such a reformulation conceives of a forest on a given area at any point in time as a probability measure on a given measurable space, and the forest management problem as a choice of an infinite sequence of probability measures such

* Department of Economics, Johns Hopkins University, Baltimore, USA and Department of Economics, National University of Singapore, Singapore. Email: akhan@jhu.edu

† Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaiso, Chile. Email: adriana.piazza@usm.cl

This work was initiated during Piazza's visit to Johns Hopkins in May 2008 and completed during Khan's visit to the National University of Singapore in July–December 2008. The authors acknowledge invaluable correspondence with Tapan Mitra as well as his generosity in making his unpublished work available to us, the encouragement of Alejandro Jofré, and the comments of an anonymous referee. Adriana Piazza gratefully acknowledges the financial support of Programa Basal PFB 03, Centro de Modelamiento Mathemático, Universidad de Chile and that of FONDECYT under projects 3080059 and 11090254.

[1] There are of course important precursors of the theory; for references to the relevant papers of Ramsey, von Neumann, Malinvaud, Samuelson, Koopmans and Radner, the reader is referred to McKenzie's (1986, 2002) texts.

that it maximizes a sequence (or its infinite sum when it is defined) of values obtained by integrating a given continuous function with respect to each measure. Whereas the given continuous function parameterizing lumber values goes towards the specification of the objective function, a natural ordering on the probability measures, stemming from the fact that there cannot be more land devoted to trees aged i years than that devoted to trees aged $(i-1)$, $i \geq 1$, last year, goes towards the specification of the technology of the problem. The essence of the Mitra–Wan reformulation, their seminal take on the problem, so to speak, is to focus on measures with a finite support, the cardinality of this support being exogenously given and constant over time. With this simplification, a forest reduces to a point in a given finite-dimensional simplex, and the continuous function of lumber values to a nonnegative tuple of the same dimension.[2]

To be sure, this is one, and not the only, way to view the Mitra–Wan forestry model, but this particular perspective allows an appreciation of the fecundity of the Mitra–Wan simplification, while at the same time, affords a general background view of the larger and more formidable problem that it has simplified.[3] At any rate, the basic outlines of the solution are by now well-known and well-understood. Although the discounted setting is beset with decisive and destructive counterexamples, the available undiscounted theory can be presented rather simply.[4] Under the assumption of values guaranteeing a unique composition of a maximally sustainable stationary forest, an invariant maximal measure so to speak, the solution falls into two distinct categorizations: with felicities linear in lumber values, optimal programs are periodic, the periodicity depending on the initial composition of the forest, leading to the so-called Faustman solution; and with felicities strictly concave in lumber values, optimal programs require any initial composition of the forest to converge to the stationary composition. As Mitra–Wan (1986, p. 229) understatedly write in their abstract,

> These results may be viewed as a possible resolution to the debate in forestry economics about what constitutes an optimal policy in forest management.

The issue then turns from long-run to transition dynamics: given the current composition of a forest this year (today), which of its vintages should be cut down and which allowed to grow next year (tomorrow)? Rather than the issue of asymptotic convergence, this involves the more difficult analytical problem of pinning down the optimal policy function, and, indeed, showing, first of all, that it is a function rather than a correspondence. As stated above, this has been resolved only for the case of linear felicities, and despite some substantial progress for the case of a dual-aged forest,[5] the problem for the full multi-aged setting remains very much an open one. It is entirely conceivable that a simple prescription

---

[2] This viewpoint is also reflected in the notation that we work with in the sequel, and in the corresponding simplification of the arguments. However, it is worthy of emphasis that this can hardly be seen as a contribution in a return, 22 years later, to an attempt that was nothing if not pioneering.

[3] The reader might want to look at Samuelson's (1976) formulation.

[4] What follows regarding this presentation is inevitably informal and imprecise; the reader is referred to Mitra and Wan (1985, 1986) for details.

[5] See Mitra (2006) for sufficient conditions on the curvature of the strictly concave felicity function under which the policy function can be characterized or its range delimited.

for an optimal action tomorrow given the state of affairs today may prove elusive for a multi-sectoral case, and that rather than a simply-stated and operationally-implementable rule, the solution may only be obtained computationally for specific functional forms. In this sense, the linear setting may be the exception rather than a guiding precedent to strive for.

What we offer in this paper is much more modest than the delineation of the optimal policy function for the Mitra–Wan forestry model in the undiscounted strictly concave case. It is rather a focus on an intermediate question: given that the composition of the forest ought to be "close" in the long-run to that of the maximally sustainable composition, is there a length of the run independent of the initial composition of the forest that the forest manager is endowed with? This is to strengthen the Mitra–Wan conclusion of asymptotic convergence to that of uniform asymptotic convergence; and in the absence of an explicit policy function, such a result might be of some reassurance to a planner who has only to go by the fact that the maximally-sustainable composition is more or less optimal in the long run. The question is new to the Mitra–Wan (1986) investigation. However, if the initial composition was restricted to *pure* forests, which is to say forests consisting only of a single vintage, the answer would follow trivially from their results; and the analytical interest in what we present here surely lies in the fact that mixtures are allowed, and that therefore there may be uncountably many of them. However, the result reported below remains a "qualitative existential assertion," and, surely, the next step must be to provide explicit computable bounds on the upper bound of the time it takes to get within an ϵ-distance of the maximally sustainable forest irrespective of how *disorganized* the forest initially is.[6]

A final introductory remark. The general theory of intertemporal resource allocation has greatly benefitted from an interplay of abstract mathematical structures and specific applied examples:[7] this interplay has identified the particularities of the example and the sharper results that they can thereby furnish, and nudged the theory towards a form that incorporates these particularities. The principal results, Theorems 4 and 6, reported here should also be viewed from this perspective. Mitra–Wan (1986) deal only with what we call here maximal programs, and Theorem 4 emphasizes that at least in the strictly concave setting, Brock's (1970) scepticism of the optimality criterion, and the effect of his counterexample, ought not to be overstated.[8] In particular, it does not pertain to the Mitra–Wan forestry model. As regards Theorem 6, in an earlier classification of McKenzie's, it offers a uniform turnpike theorem of the third kind; and relative to an analogous theorem for the Robinson–Solow–Srinivasan (RSS) model, shows how the technological structure of the Mitra–Wan forestry model can be exploited to furnish more direct and constructive

[6] Put differently, the next step is to figure out a proof of Theorem 3 below in which T_o can be explicitly computed. We emphasize this in Remark 4 below, and in Remarks 5 and 6, indicate where the present proof fails in this regard. This step-by-step approach to mathematical economics, one battle at a time, is again something that the authors have learnt from Tapan Mitra.

[7] In addition to one-sector and two-sector models, one can count models of forestry and resource economics, vintage-capital models and the Weitzman model; see Le Van, Dana, Mitra, and Nishimura (2006) and the references therein to earlier work.

[8] Mitra (2005, 2006) too is limited to maximal programs, although the former also presents an alternative criterion to which maximality is equivalent. The reader is warned that there is a proliferation of confusing terminology, and that we conform to Definitions 6 and 7 below. See McKenzie (2002) in this connection.

proofs.[9] It is perhaps of some methodological interest that the Mitra–Wan analysis formed the bulwark for the argumentation and results developed for the RSS model in Khan–Mitra (2005); now we rely on the analysis of Zaslavski (2005) and Khan–Zaslavksi (2007) to sharpen the Mitra–Wan results.[10] In particular, our emphasis on the coincidence of optimal and maximal programs in this paper is taken from the former. Hopefully, in the longer run, all of these exercises will coalesce towards a general model of vintage capital theory.

2 Model and preliminary results

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

Let the total surface be equal to 1 and let n be the age after which a tree dies or losses its economic value. We consider that the timber content per unit of area is related only to the age of the trees, through the biomass coefficient vector $b = (b_1, \ldots, b_n)$.

We make the assumption that:

$$\text{there exists } \sigma = \{1, \ldots, n\} \text{ such that } \frac{b_\sigma}{\sigma} > \frac{b_i}{i} \quad \text{for all } i \in \{1, \ldots, n\} \setminus \{\sigma\}.$$

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

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

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

This model is in fact an equivalent formulation of the one proposed by Mitra and Wan (1985, 1986).

Definition 1 *A sequence $\{x(t)\}$ is called a program if for each $t \geq 0$*

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

Definition 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 for each t satisfying $T_1 \leq t < T_2$ relations (1) hold.*

---

9 See in particular the proofs of Proposition 9 and Lemma 4 below.

10 The Stiglitz policy for the RSS model has stood as the guiding analogue for the Faustman policy. This dialectical dependence between the two models can also be seen in the relationship between Mitra (2006) and Khan–Mitra (2008).

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,

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

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

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

Remark 1 $(x, x') \in \Omega \Leftrightarrow x, x' \in \Delta$ and $\lambda(x, x') \geq 0$.

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

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

Definition 3 *A* golden-rule stock $\hat{x} \in I\!R^n_+$ *is such that* $(\hat{x}, \hat{x})$ *is a solution to the problem:*

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

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

Proof: The existence is proved in Mitra and Wan (1986, corollary 4.4) and the uniqueness comes from Mitra and Wan (1986, theorem 6.1). □

Definition 4 *Set* $\hat{p} \in I\!R^n_+$, $\hat{p} = w'(\frac{b_\sigma}{\sigma})\frac{b_\sigma}{\sigma}(1, 2, \ldots, n)$ *and* $\hat{c} = \lambda(\hat{x}, \hat{x})$*; namely,* $\hat{c}_\sigma = \frac{1}{\sigma}$ *and* $\hat{c}_i = 0$ *for all* $i \neq \sigma$. *We define the* value loss *associated with any* $(x, x') \in \Omega$ *to be*[11]

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

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

Lemma 1 *For any* $(x, x') \in \Omega$ *we have*

$$\delta(x, x') \geq w'\left(\frac{b_\sigma}{\sigma}\right)\left[\sum_{i=1}^{n-1}\left(\frac{b_\sigma}{\sigma} - \frac{b_i}{i}\right) i\,(x_i - x'_{i+1}) + \left(\frac{b_\sigma}{\sigma} - \frac{b_n}{n}\right) n x_n\right] \geq 0. \quad (2)$$

[11] 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.

PROOF: The nonnegativity of the value loss function is already established in Mitra and Wan (1985). We present an alternative proof that determines a disaggregated lower bound of the value loss function that will be used in the sequel to characterize the von Neumann facet.

As w is concave, we know that

$$w(y) \le w\left(\frac{b_\sigma}{\sigma}\right) + w'\left(\frac{b_\sigma}{\sigma}\right)\left(y - \frac{b_\sigma}{\sigma}\right) \quad \text{for all } y \in I\!R.$$

In particular, by taking $y = b\lambda(x, x')$, we obtain:

$$\begin{aligned}\delta(x, x') \ \ge\ & w\left(\frac{b_\sigma}{\sigma}\right) - w\left(\frac{b_\sigma}{\sigma}\right) - w'\left(\frac{b_\sigma}{\sigma}\right)\left[\sum_{i=1}^{n-1} b_i(x_i - x'_{i+1}) + b_n x_n - \frac{b_\sigma}{\sigma}\right] \\ & - \sum_{i=1}^{n} w'\left(\frac{b_\sigma}{\sigma}\right)\frac{b_\sigma}{\sigma} i(x'_i - x_i) \\ = {} & w'\left(\frac{b_\sigma}{\sigma}\right)\Bigg[-\sum_{i=1}^{n-1} b_i(x_i - x'_{i+1}) - b_n x_n + \frac{b_\sigma}{\sigma} - \frac{b_\sigma}{\sigma}\underbrace{\sum_{i=1}^{n} i(x'_i - x_i)}_{(A)}\Bigg].\end{aligned}$$

Rearranging the last summation we obtain:

$$\begin{aligned}(A) &= \sum_{i=1}^{n} i x'_i - \sum_{i=1}^{n} i x_i = \sum_{i=0}^{n-1}(i+1)x'_{i+1} - \sum_{i=1}^{n} i x_i \\ &= \sum_{i=0}^{n-1} x'_{i+1} + \sum_{i=1}^{n-1} i(x'_{i+1} - x_i) - n x_n = 1 - n x_n + \sum_{i=1}^{n-1} i(x'_{i+1} - x_i). \qquad (3)\end{aligned}$$

We substitute this expression in the previous inequality,

$$\begin{aligned}\delta(x, x') \ \ge\ & w'\left(\frac{b_\sigma}{\sigma}\right)\Bigg\{-\sum_{i=1}^{n-1} b_i(x_i - x'_{i+1}) - b_n x_n + \frac{b_\sigma}{\sigma} \\ & - \frac{b_\sigma}{\sigma}\left[1 - n x_n + \sum_{i=1}^{n-1} i(x'_{i+1} - x_i)\right]\Bigg\} \\ = {} & w'\left(\frac{b_\sigma}{\sigma}\right)\left[\sum_{i=1}^{n-1}\left(\frac{b_\sigma}{\sigma} - \frac{b_i}{i}\right) i\,(x_i - x'_{i+1}) + \left(\frac{b_\sigma}{\sigma} - \frac{b_n}{n}\right) n x_n\right] \ge 0,\end{aligned}$$

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

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

Definition 5 *A program* $\{x(t)\}$ *is called* good *if there exists* $M \in I\!R$ *such that for all* $T \ge 0$, $\sum_{t=0}^{T}[w(bc(t)) - w(\frac{b_\sigma}{\sigma})] \ge M$, *where* $c(t) = \lambda(x(t), x(t+1))$. *A program is* bad *if* $\lim_{T\to\infty} \sum_{t=0}^{T}[w(bc(t)) - w(\frac{b_\sigma}{\sigma})] = -\infty$.

The following general result of Gale applies to the Mitra–Wan forestry model.

Proposition 1 *Any program that is not good is bad.*

PROOF: Due to Lemma 1 we know that $w(bc(t)) - w(\frac{b_\sigma}{\sigma}) \leq \hat{p}(x(t) - x(t+1))$ for any $(x(t), x(t+1)) \in \Omega$ and $c(t) = \lambda(x(t), x(t+1))$. Adding up these inequalities along a given program $\{x(t)\}$ we get

$$\sum_{t=T_1}^{T-1} w(bc(t)) - w\left(\frac{b_\sigma}{\sigma}\right) \leq \hat{p}(x(T_1) - x(T)).$$

Using the fact that Δ is a compact set of $I\!R^n$ we know that there is M such that

$$\sum_{t=T_1}^{T-1} w(bc(t)) - w\left(\frac{b_\sigma}{\sigma}\right) \leq M \quad \text{for all } T_1,\ T \in I\!N. \tag{4}$$

For any program that is not good and for any $N \in I\!R$ there is T_N such that $\sum_{t=0}^{T_N}[w(bc(t)) - w(\frac{b_\sigma}{\sigma})] \leq N - M$. By choosing $T_1 = T_N$ in (4) we obtain

$$\sum_{t=0}^{T}\left[w(bc(t)) - w\left(\frac{b_\sigma}{\sigma}\right)\right] \leq N \quad \text{for all } T \geq T_N.$$ □

It follows immediately that programs are partitioned into good and bad ones. The proposition below shows an equivalent characterization of good and bad programs.

Proposition 2

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

PROOF: The proof of (ii) follows from the equality

$$\sum_{t=0}^{T-1} \delta(x(t), x(t+1)) = -\sum_{t=0}^{T-1}\left[w(bc(t)) - w\left(\frac{b_\sigma}{\sigma}\right)\right] - \hat{p}(x(T) - x(0))$$

and the fact that the inner product $\hat{p}(x(T) - x(0))$ is bounded for all $x(T)$. Then (i) follows from (ii) because programs are partitioned into good and bad programs. □

We establish next the existence of at least one good program from any initial state.

Remark 2 Given x_0, consider the program:

$$
\begin{array}{ll}
x(0)=x_0 & c(0)=x_0 \\
x(1)=(1,0,0,0\ldots,0) & c(1)=\left(\frac{\sigma-1}{\sigma},0,0,\ldots,0\right) \\
x(2)=\left(\frac{\sigma-1}{\sigma},\frac{1}{\sigma},0,0,\ldots,0\right) & c(2)=\left(\frac{\sigma-2}{\sigma},0,\ldots,0\right) \\
x(3)=\left(\frac{\sigma-1}{\sigma},\frac{1}{\sigma},\frac{1}{\sigma},0,\ldots,0\right) & c(3)=\left(\frac{\sigma-3}{\sigma},0,\ldots,0\right) \\
\vdots & \vdots \\
x(t)=\hat{x} & c(t)=\hat{c} \qquad \text{for all } t\geq\sigma.
\end{array} \tag{5}
$$

It is evident that this program is good, because $\delta(\hat{x},\hat{x})=0$.

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

The remark above implies that $\mu(x_0)<\infty$. The following result can now be established.

Proposition 3 *From any $x_0\in\Delta$ there exist a program $\{x(t)\}$ such that*

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

PROOF: Let us define the function $\gamma(\{x(t)\})=\sum_{t\in I\!N}\delta(x(t),x(t+1))$ in the space $\Pi=\prod_{t=0}^{\infty}\Delta$. Observe that $\gamma:\Pi\to[0,\infty]$. Let $\gamma_T(\{x(t)\}_{t=0}^{T})=\sum_{t=0}^{T-1}\delta(x(t),x(t+1))$. For every T, γ_T is convex and continuous in the product topology and $\gamma_T\leq\gamma_{T+1}$. Hence, γ is the increasing limit of continuous convex functions; it is, therefore, convex, lower semi-continuous. We know there is at least one good program, so $\min\gamma(\cdot)<\infty$. Then there is a minimizer $\{x^*(t)\}$ on Π which is compact, non-empty in the product topology such that $\gamma(\{x^*(t)\})<\infty$.[12] □

Theorem 2 *For each good program $\{x(t)\}$, $\lim_{t\to\infty}x(t)=\hat{x}$.*

PROOF: See Mitra and Wan (1986, lemma 6.4). □

We suppose, as in the literature taking its lead from Ramsey, and, subsequently, Atsumi and von Weizsäcker, that future welfare levels are treated like current ones in the planner's objective function.[13]

Definition 6 *A program $\{x^*(t)\}$ is* optimal *if for any program $\{x(t)\}$ such that $x(0)=x^*(0)$ we have*

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

[12] The proof we present is a straightforward adaptation of that of Le Van, Dana, Mitra, and Nishimura (2006, proposition 1.4.2). Being a rather brief one, we include it for the sake of completeness.

[13] For references to these papers, see McKenzie (2002). For a discussion of these definitions, see Brock (1970) and McKenzie (2002).

Definition 7 *A program* $\{x^*(t)\}$ *is* maximal *if for any program* $\{x(t)\}$ *such that* $x(0) = x^*(0)$ *we have*

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

However, these two notions coincide for the Mitra and Wan model, as we show immediately. We start with an easy technical lemma showing that

Lemma 2 *Every maximal program is good.*

PROOF: Let $\{x^*(t)\}$ be a maximal program from x_0 and $\{x(t)\}$ a good program from x_0. We know that

$$\begin{aligned}\sum_{t=0}^{T-1} w(bc(t)) - w(bc^*(t)) &= \sum_{t=0}^{T-1}\left[w(bc(t)) - w\left(\frac{b_\sigma}{\sigma}\right)\right] - \sum_{t=0}^{T-1}\left[w(bc^*(t)) - w\left(\frac{b_\sigma}{\sigma}\right)\right]\\ &= -\sum_{t=0}^{T-1}\delta(x(t), x(t+1)) - \hat{p}(x(T) - x_0)\\ &\quad + \sum_{t=0}^{T-1}\delta(x^*(t), x^*(t+1)) + \hat{p}(x^*(T) - x_0)\\ &= -\sum_{t=0}^{T-1}\delta(x(t), x(t+1)) + \sum_{t=0}^{T-1}\delta\big(x^*(t), x^*(t+1)\big)\\ &\quad + \hat{p}(x^*(T) - x(T))\\ &\ge -\sum_{t=0}^{T-1}\delta(x(t), x(t+1)) + \sum_{t=0}^{T-1}\delta(x^*(t), x^*(t+1)) + M. \quad (6)\end{aligned}$$

Taking $\lim\inf_{T\to\infty}$ at both sides we get $0 \ge -\sum_{t=0}^{\infty}\delta(x(t), x(t+1)) + \sum_{t=0}^{\infty}\delta(x^*(t), x^*(t+1))$, implying that $\{x^*(t)\}$ is good because $\sum_{t=0}^{\infty}\delta(x^*(t), x^*(t+1)) \le (\sum_{t=0}^{\infty}\delta(x(t), x(t+1)) - M) \in I\!R$. □

Proposition 4 *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 (6) 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 $\lim_T \sum_{t=0}^{T}\delta(x(t), x(t+1))$ is well defined, as well as $\lim_T \sum_{t=0}^{T}\delta(x^*(t), x^*(t+1))$ and also that $\lim_T x^*(T) = \lim_T x(T) = \hat{x}$. Then, considering (6) again and on letting $T\to\infty$, we see that the limit of the right-hand side is well-defined, and, hence, it is the limit of the left-hand side:

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

Theorem 3 *For each initial state $x_0 \in \Delta$ there exists an optimal program $\{x(t)\}$ that satisfies $x(0) = x_0$.*

PROOF: (Mitra and Wan 1986, theorem 4.1) establishes the existence of a maximal program from every x_0. This result can also be deduced from Proposition 3 and the following theorem. $\square$

3 Main results

We can now present our principal results.

Theorem 4 *Let $\{x(t)\}$ be a program from x_0. The following conditions are equivalent:*

(i) $\sum_{t=0}^{\infty} \delta(x(t), x(t+1)) = \mu(x(0))$.
(ii) $\{x(t)\}$ *is optimal.*
(iii) $\{x(t)\}$ *is maximal.*

Remark 3 It is worth reminding the reader that the essence of Brock's (1970) contribution is to show that minimum value-loss programs are maximal, but that the converse is not considered. As has now been shown for the two-sector RSS model with linear felicities, it may be false; see Khan and Mitra (2008) for references. The reader is also referred to Khan and Zaslavski (2007, theorem 3.1) for an analogue of this theorem for the RSS model.

Theorem 5 *Let $\epsilon > 0$. There exists $\delta > 0$ such that for each optimal program $\{x(t)\}$ satisfying $\|x(0) - \hat{x}\|_\infty < \delta$ the following inequality holds:*

$$\|x(t) - \hat{x}\|_\infty < \epsilon \quad \textit{for all } t \geq 0.$$

Theorem 6 *Let $\epsilon > 0$. There exists a natural number T_0 such that for each optimal program $\{x(t)\}$ the following inequality holds:*

$$\|x(t) - \hat{x}\|_\infty < \epsilon \quad \textit{for all } t \geq T_0.$$

Remark 4 See Remarks 5 and 6 for an emphasis on the non-constructive nature of the proofs of these theorems. Given the motivation of this paper and behind this exercise, explicit computable proofs remain an important open question.[14]

4 Proof of Theorem 4

The proof of Theorem 4 is simple enough that it can be offered without any additional machinery.

[14] Also see footnote 6 in this connection.

PROOF OF THEOREM 4: Proposition 4 states that (ii) and (iii) are equivalent. We show now the equivalence between (i) and (iii).

We know from Proposition 3 that γ has a good minimizer $\{x^*(t)\}$. Let us show that it is optimal. Let $\{x(t)\}$ be any other program from x_0. Either it is good or bad. In the first case, let $T \to \infty$ in (6) to obtain

$$\lim_T \sum_{t=0}^{T} w(bc(t)) - w(bc^*(t)) = -\gamma(\{x(t)\}) + \gamma(\{x^*(t)\}) \leq 0$$

since $\{x^*(t)\}$ is a minimizer of γ. In the second case when the program is bad,

$$\lim_T \sum_{t=0}^{T} w(bc(t)) - w(bc^*(t)) = \lim_T \sum_{t=0}^{T} \left[w(bc(t)) - w\left(\frac{b_\sigma}{\sigma}\right)\right] - \lim_T \sum_{t=0}^{T} \left[w(bc^*(t)) - w\left(\frac{b_\sigma}{\sigma}\right)\right] = -\infty.$$

Hence, $\{x^*(t)\}$ is optimal.

Conversely, if $\{x^*(t)\}$ is optimal, then from Proposition 3 it is good. If $\{x(t)\}$ is any good program, we have

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

because $\{x^*(t)\}$ is optimal. If $\{x(t)\}$ is a bad program then $\gamma(\{x(t)\}) = +\infty$. Hence, $\{x^*(t)\}$ is a minimizer of γ. □

5 Auxiliary results for Theorems 5 and 6

Proposition 5 *There exists $m > 0$ such that for any program $\{x(t)\}$, $x(t) \leq me$ for all $t \geq 0$, where e is the unit vector.*[15]

Proposition 6 *The von Neumann facet $\{(x, x') \in \Omega : \delta(x, x') = 0\}$ is*

$$\left\{ (x, x') \in \Omega \Big/ \begin{array}{l} x'_{i+1} = x_i \quad \text{for all } i = 1, \ldots, n-1,\ i \neq \sigma \\ x'_{\sigma+1} = x_\sigma - \dfrac{1}{\sigma}, \qquad x_n = 0 \end{array} \right\}.$$

PROOF: We recall (2) proved in Lemma 1:

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

15 For the Mitra–Wan model being considered here, this claim is obvious and hardly merits being introduced as a proposition; we do so for the readers interested in comparing the analysis of the Mitra–Wan model presented here with that in Khan and Zaslavski (2007) for the RSS model.

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

$$\delta(x, x') = 0 \text{ implies } x_i = x'_{i+1} \quad \text{for all } i \neq \sigma,\ i < n \text{ and } x_n = 0. \tag{7}$$

It remains to prove that $x'_{\sigma+1} = x_\sigma - \frac{1}{\sigma}$. Using (7) and (3) we can find a much simpler expression for $\sum_{i=1}^{n} i(x'_i - x_i)$, as given by

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

and so,

$$\delta(x, x') = 0 \Rightarrow \delta(x, x') = w\left(\frac{b_\sigma}{\sigma}\right) - w(b_\sigma c_\sigma) - w'\left(\frac{b_\sigma}{\sigma}\right) b_\sigma \left[\frac{1}{\sigma} - (x_\sigma - x'_{\sigma+1})\right] = 0.$$

The function $f : I\!R \to I\!R$, $f(c) = w(\frac{b_\sigma}{\sigma}) - w(b_\sigma c) - w'(\frac{b_\sigma}{\sigma}) b_\sigma (\frac{1}{\sigma} - c)$ is a strictly convex, nonnegative function such that $f(c) > f(\frac{1}{\sigma}) = 0$ for all $c \neq \frac{1}{\sigma}$. Hence, $\delta(x, x') = 0$ implies that $c_\sigma = \frac{1}{\sigma}$. □

Proposition 7 *Let a program $\{x(t)\}$ satisfy $\delta(x(t), x(t+1)) = 0$ for all $t = T, \ldots, T + \sigma - 1$. Then $x(t) = \hat{x}$ for all $t = T, \ldots, T + \sigma$.*

PROOF: Without loss of generality take $T = 0$. Thanks to Proposition 6, we know that $\delta(x(t), x(t+1)) = 0$ implies that $c(t) = \hat{c}$. It is easy to see that $c(t) = \hat{c}$ for all $t = 0, \ldots, \sigma - 1$ implies $x_i(\sigma) = (1/\sigma)$ for all $i \leq \sigma$. The area balance now shows that $x_i(\sigma) = 0,\ i > \sigma$ and so $x(\sigma) = \hat{x}$.

Furthermore, Proposition 6 implies that $\delta(x, \hat{x}) = 0 \Longrightarrow x = \hat{x}$ and then the proof follows by backwards induction. □

Proposition 8

$$\sup\{\mu(x) : x \in \Delta\} < \infty.$$

PROOF: Given any $x \in \Delta$ consider the program $\{x(t)\}$ given by (5) which reaches $\hat{x}$ in σ steps. Hence, the accumulated value loss will be a finite sum

$$\sum_{t \geq 0} \delta(x(t), x(t+1)) = \sum_{t=0}^{\sigma-1} \delta(x(t), x(t+1)).$$

However, given any pair $(x, x') \in \Omega$, there is M such that

$$\delta(x, x') = w\left(\frac{b_\sigma}{\sigma}\right) - w(bc) - \hat{p}(x' - x) \leq w\left(\frac{b_\sigma}{\sigma}\right) + \hat{p}\, x \leq M.$$

Hence, for every $x \in \Delta$, the proposed program has an accumulated value loss bounded above by σM,

$$\sup\{\mu(x) : x \in \Delta\} \leq \sup\left\{\sum_{t=0}^{\sigma-1} \delta(x(t), x(t+1)) : x \in \Delta\right\} \leq \sigma M < \infty. \quad \square$$

Proposition 9 *Let $\epsilon > 0$. There is $\delta > 0$ such that for each $x \in \Delta$ satisfying $\|x - \hat{x}\|_\infty < \delta$ the inequality $\mu(x) < \epsilon$ holds.*

PROOF: The proof of this proposition is rather technical and it is based on *explicitly* finding a program from x whose accumulated value loss is less that $\epsilon > 0$.

Thanks to the continuity of $\delta(\cdot, \cdot)$ at the point $(\hat{x}, \hat{x})$, we know that given $\epsilon > 0$ there is δ_1 such that given any pair $(x, x') \in \Omega$ satisfying

$$\|(x, x') - (\hat{x}, \hat{x})\|_\infty \le \delta_1 \quad \text{implies} \quad \delta(x, x') \le \epsilon/\sigma.$$

Take $\delta = \min\{\frac{\delta_1}{n}, \frac{1}{n\sigma}\}$, we will see that for any x satisfying $\|x - \hat{x}\|_\infty \le \delta$ there is a program $\{x(t)\}$ from x such that

$$\begin{cases} \|x(t) - \hat{x}\|_\infty < \delta_1, & t = 1, \dots, \sigma - 1 \\ x(t) = \hat{x}, & t \ge \sigma \end{cases} \tag{8}$$

implying that $\sum_{t=0}^{\infty} \delta(x(t), x(t+1)) = \sum_{t=0}^{\sigma-1} \delta(x(t), x(t+1)) < \epsilon$.

Given any x such that $\|x - \hat{x}\|_\infty < \delta$, x is of the form

$$x = \left(\frac{1}{\sigma} + \phi_1, \dots, \frac{1}{\sigma} + \phi_\sigma, \phi_{\sigma+1} \dots, \phi_n\right), \text{ where } \sum_i \phi_i = 0 \text{ and } |\phi_i| < \delta \ \forall i \ \text{ and } \phi_i > 0 \text{ for } i > \sigma.$$

We claim that the following sequence is a program from x fulfilling relations (8):

$$\begin{aligned} x(0) &= x_0 \\ x(t) &= \Big(\underbrace{\frac{1}{\sigma}, \dots, \frac{1}{\sigma}}_{t}, \underbrace{\frac{1}{\sigma} + \phi_1, \dots, \frac{1}{\sigma} + \phi_{\sigma-t-1}}_{\sigma-t-1}, x_\sigma(t), x_{\sigma+1}(t), \underbrace{0, \dots, 0}_{n-\sigma-1}\Big) \quad 0 < t < \sigma \\ x(t) &= \hat{x} \quad t \ge \sigma, \end{aligned}$$

where $x_\sigma(t) = \frac{1}{\sigma} + \phi_{\sigma-t} + \min(0, \sum_{i=\sigma-t+1}^{n} \phi_i)$ and $x_{\sigma+1}(t) = \max(0, \sum_{i=\sigma-t+1}^{n} \phi_i)$.

We check easily that $\|x(t) - \hat{x}\|_\infty < n\delta \le \delta_1$. It is left to see that the proposed sequence is in fact a program. Of course, this is so for $t \ge \sigma$. To check it for $t < \sigma$ we prove first that $x(t) \in \Delta$ for $t = 1, \dots, \sigma - 1$. From the definition of the sequence and $\delta \le \frac{1}{n\sigma}$ it is evident that $x(t) \in I\!R^n_+$ for $t = 1, \dots, \sigma - 1$. Let us see that $\sum_i x_i(t) = 1$:

$$\begin{aligned} \textstyle\sum_i x_i(t) &= t\frac{1}{\sigma} + \sum_{i=1}^{\sigma-t-1} \left(\frac{1}{\sigma} + \phi_i\right) + \frac{1}{\sigma} + \phi_{\sigma-t} \\ &\quad + \min\left(0, \sum_{i=\sigma-t+1}^{n} \phi_i\right) + \max\left(0, \sum_{i=\sigma-t+1}^{n} \phi_i\right) \end{aligned}$$

$$= \frac{1}{\sigma}\big[t + (\sigma - t - 1) + 1\big] + \sum_{i=1}^{\sigma-t-1} \phi_i + \phi_{\sigma-t} + \sum_{i=\sigma-t+1}^{n} \phi_i$$

$$= 1 + \sum_{i=1}^{n} \phi_i = 1.$$

The proof is completed by showing that $\lambda(x(t), x(t+1)) \geq 0$ for $t = 0, \ldots, \sigma - 1$ (see Remark 1).

Case: t = 0

$$\lambda(x_0, x(1)) = \left(\underbrace{0, \ldots, 0}_{\sigma-2}, \; -\min\left(0, \sum_{i=\sigma}^{n} \phi_i\right), \frac{1}{\sigma} + \phi_\sigma - \max\left(0, \sum_{i=\sigma}^{n} \phi_i\right), \underbrace{\phi_{\sigma+1} \ldots, \phi_n}_{n-\sigma} \right).$$

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

Case: t = 1, ..., σ − 2

$$\lambda\big(x(t), x(t+1)\big) = \big(\underbrace{0, \ldots, 0}_{\sigma-2}, c_{\sigma-1}(t), c_\sigma(t), c_{\sigma+1}(t), \underbrace{0, \ldots, 0}_{n-\sigma-1}\big),$$

where

$$c_{\sigma-1}(t) = -\min\left(0, \sum_{i=\sigma-t}^{n} \phi_i\right)$$

$$c_\sigma(t) = \frac{1}{\sigma} + \phi_{\sigma-t} + \min\left(0, \sum_{i=\sigma-t+1}^{n} \phi_i\right) - \max\left(0, \sum_{i=\sigma-t}^{n} \phi_i\right)$$

$$c_{\sigma+1}(t) = x_{\sigma+1}(t).$$

Again, the nonnegativity of the $c_i(t)$ for $i \neq \sigma$ is evident and after some computations we can see that

$$c_\sigma(t) \geq \frac{1}{\sigma} - |\phi_{\sigma-t}| - \left| \sum_{i=\sigma-t+1}^{n} \phi_i \right| \geq \frac{1}{\sigma} - (n - \sigma + t)\delta > 0$$

and, finally,

Case: t = σ − 1

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

□

Remark 5 Note that with explicit assumptions on the curvature of the felicity function $w(\cdot)$, one could furnish a bound on δ_1 as a function of ϵ.[16] This bound would then

[16] We thank an anonymous referee for sensitizing us to these computational considerations.

furnish a bound on δ defined in the next step of the argument. In any case, by definition, $\delta \leq (1/n\sigma)$.

6 Proof of Theorems 5 and 6

We begin with an important preliminary lemma.

Lemma 3 *Let $\epsilon > 0$. There exists $\delta > 0$ such that for each sequence $\{x(t)\}_{t=-\infty}^{\infty} \subset \Delta$ satisfying*

$$(x(t),\, x(t+1)) \in \Omega \text{ for all integers } t,$$
$$\delta(x(t),\, x(t+1)) \leq \delta \text{ for all integers } t,$$

the following inequality holds:

$$\|x(t) - \hat{x}\|_{\infty} \leq \epsilon \text{ for all integers } t.$$

PROOF: The proof follows the same lines of that of Khan and Zaslavski (2007, lemma 6.1). Let us assume the converse. Then for each natural number k there exist a sequence $\{x^k(t)\}_{t=-\infty}^{\infty} \subset \Delta$ that satisfies

$$(x^k(t),\, x^k(t+1)) \in \Omega \text{ for all integers } t \tag{9}$$

$$\delta(x^k(t),\, x^k(t+1)) \leq \frac{1}{k} \text{ for all integers } t \tag{10}$$

and an integer τ_k such that

$$\|x^k(\tau_k) - \hat{x}\|_{\infty} \geq \epsilon. \tag{11}$$

We may assume without loss of generality that $\tau_k = 0$ for all $k \in \mathbb{N}$. Extracting subsequences, re-indexing and using a diagonalization process, we obtain that there exists a strictly increasing sequence of natural numbers $\{k_j\}_{j=1}^{\infty}$ such that for each integer s there is

$$\tilde{x}(s) = \lim_{j\to\infty} x^{k_j}(s). \tag{12}$$

Because Ω is a closed set, it follows from (9) and (12) that

$$(\tilde{x}(s),\, \tilde{x}(s+1)) \in \Omega \text{ for all integers } s. \tag{13}$$

Because the function $\delta(\cdot,\, \cdot)$ is nonnegative and continuous it follows from (10) and (12) that

$$\delta(\tilde{x}(s),\, \tilde{x}(s+1)) = 0 \text{ for all integers } s. \tag{14}$$

Relations (11), (12) and $\tau_k = 0$ for all k imply

$$\|\tilde{x}(0) - \hat{x}\|_{\infty} \geq \epsilon. \tag{15}$$

However, it follows from (13), (14) and Proposition 7 that

$$\tilde{x}(s) = \hat{x} \text{ for all integers } s,$$

which contradicts (15). This contradiction proves the lemma. □

Before turning to the proof of Theorem 5, we show another technical lemma.

Lemma 4 *Let $\epsilon > 0$. There is $\delta > 0$ such that given any x satisfying $\|x - \hat{x}\|_\infty < \delta$, there exists a sequence $\{x(t)\}_{t=-\infty}^{0}$ such that*

(i) $x(0) = x$
(ii) $\|x(t) - \hat{x}\|_\infty < n\delta$ *for all* $t < 0$
(iii) $(x(t), x(t+1)) \in \Omega$ *for all* $t < 0$ *and*
(vi) $\delta(x(t), x(t+1)) < \epsilon$ *for all* $t < 0$.

PROOF: Using the continuity of $\delta(\cdot, \cdot)$ at the point $(\hat{x}, \hat{x})$, we chose $\delta > 0$ such that $\delta(x(t), x(t+1)) < \epsilon$ whenever $\|(x(t), x(t+1)) - (\hat{x}, \hat{x})\|_\infty < n\delta$. Hence, (iv) follows from (ii). Let δ be such that $n\delta < 1/\sigma$.

Observe that if $\|x - \hat{x}\|_\infty < \delta$, then x is of the form $x = (\frac{1}{\sigma} + \phi_1, \ldots, \frac{1}{\sigma} + \phi_\sigma, \phi_{\sigma+1} \ldots, \phi_n)$, where $\sum_i \phi_i = 0$ and $|\phi_i| < \delta < 1/(n\sigma)$. Let $\phi = \max_i\{\phi_i\}$.

We again provide a constructive proof that is somewhat involved. To help the reader, we present first the case $\sigma = 3$ and $n = 8$ before providing the proof in the general case. Given x such that $\|x - \hat{x}\|_\infty < \delta$, one possible sequence satisfying all the announced properties would be:

$$
\begin{aligned}
x(t) &= \hat{x} \quad t < -8 & c(t) &= \hat{c} \quad t < -8\\
x(-8) &= \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0, 0, 0, 0\right) & c(-8) &= (0, 0, x_\sigma(-8) - \phi, 0, 0, 0, 0, 0)\\
x(-7) &= \left(\frac{1}{3} - \phi, \frac{1}{3}, \frac{1}{3}, \phi, 0, 0, 0, 0\right) & c(-7) &= (0, 0, x_\sigma(-7) - \phi, 0, 0, 0, 0, 0)\\
x(-6) &= \left(\frac{1}{3} - \phi, \frac{1}{3} - \phi, \frac{1}{3}, \phi, \phi, 0, 0, 0\right) & c(-6) &= (0, 0, x_\sigma(-6) - \phi, 0, 0, 0, 0, 0)\\
x(-5) &= \left(\frac{1}{3} - \phi, \frac{1}{3} - \phi, \frac{1}{3} - \phi, \phi, \phi, \phi, 0, 0\right) & c(-5) &= (0, 0, x_\sigma(-5) - \phi, 0, 0, 0, 0, 0)\\
x(-4) &= \left(\frac{1}{3} - 2\phi, \frac{1}{3} - \phi, \frac{1}{3} - \phi, \phi, \phi, \phi, \phi, 0\right) & c(-4) &= (0, 0, x_\sigma(-4) - \phi, 0, 0, 0, 0, 0)\\
x(-3) &= \left(\frac{1}{3} - 2\phi, \frac{1}{3} - 2\phi, \frac{1}{3} - \phi, \phi, \phi, \phi, \phi, \phi\right) & c(-3) &= (2\phi, 0, x_\sigma(-3) - \phi, 0, 0, 0, 0, \phi)\\
x(-2) &= \left(\frac{1}{3} + \phi, \frac{1}{3} - 4\phi, \frac{1}{3} - 2\phi, \phi, \phi, \phi, \phi, \phi\right) & c(-2) &= (0, 3\phi, x_\sigma(-2) - \phi, 0, 0, 0, 0, \phi)\\
x(-1) &= \left(\frac{1}{3} + \phi, \frac{1}{3} + \phi, \frac{1}{3} - 7\phi, \phi, \phi, \phi, \phi, \phi\right) & c(-1) &= (\phi - \phi_2, \phi - \phi_3, \tfrac{1}{3} - 7\phi - \phi_4,\\
& & &\quad \phi - \phi_5, \phi - \phi_6, \phi - \phi_7, \phi - \phi_8, \phi)\\
x(0) &= \left(\frac{1}{3} + \phi_1, \frac{1}{3} + \phi_2, \frac{1}{3} + \phi_3, \phi_4, \phi_5, \phi_6, \phi_7, \phi_8\right) = x,
\end{aligned}
$$

where in the last transition we are using $\sum_i \phi_i = 0$.

Observe that in this example the harvesting policy remains constant for $t=[-n,\ldots,-\sigma+1]$; namely,

$$c(t)=c(x(t))=(0,\ldots,0,\underbrace{x_\sigma(t)-\phi}_{\sigma-th\ pos},0,\ldots,0) \qquad -n\le t<-\sigma.$$

At stages $[-\sigma,\ldots,-2]$, we modify the harvesting policy to assure that $x_1(t)=\frac{1}{\sigma}+\phi$; in fact, the new policy can be expressed as

$$c(t)=c(t,x(t))=(0,\ldots,0,\underbrace{\frac{1}{\sigma}+\phi-x_\sigma(t)}_{(t+\sigma+1)-th\ pos},0,\ldots,0,\underbrace{x_\sigma(t)-\phi}_{\sigma-th\ pos},0,\ldots,0,\phi).$$

We use the above remarks to provide the proof in the general case. Given σ and n, let $k\in I\!N$ and $j\in[0,\ldots,\sigma-1]$ be such that $j=n(\sigma)$ and $n=k\sigma+j$. The $(-\infty)$-tail of our sequence will be $x(t)=\hat{x}$ for all $t\le -n$ and $c(t)=\hat{c}$ for all $t<-n$. During $n-\sigma$ stages we apply the harvesting policy given by

$$c(x(t))=\big(0,\ldots,0,\underbrace{x_\sigma(t)-\phi}_{\sigma-th\ pos},0,\ldots,0,\underbrace{x_n(t)}_{=0}\big) \qquad -n\le t<-\sigma,$$

generating the sequence $\{x(t)\}_{t=-n+1}^{-\sigma}$,

$$\begin{aligned}
x(-n+1) &= \left(\frac{1}{\sigma}-\phi,\frac{1}{\sigma},\ldots,\frac{1}{\sigma},\phi,0,\ldots,0\right)\\
x(-n+2) &= \left(\frac{1}{\sigma}-\phi,\frac{1}{\sigma}-\phi,\frac{1}{\sigma},\ldots,\frac{1}{\sigma},\phi,\phi,0,\ldots,0\right)\\
&\vdots\\
x(-\sigma) &= \left(\underbrace{\frac{1}{\sigma}-k\phi,\ldots,\frac{1}{\sigma}-k\phi}_{j},\underbrace{\frac{1}{\sigma}-k\phi+\phi,\ldots,\frac{1}{\sigma}-k\phi+\phi}_{\sigma-j},\underbrace{\phi,\ldots,\phi}_{n-\sigma}\right).
\end{aligned}$$

Due to Remark 1 it is not difficult to check that the generated sequence is such that $(x(t),x(t+1))\in\Omega$ for $t<-\sigma$. Furthermore, $\|\hat{x}-x(t)\|_\infty<n\delta$ for $t\le-\sigma$.

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

$$\text{for } -\sigma\le t<-j,\quad c_i(t)=\begin{cases} k\phi & i=t+\sigma\\ x_\sigma(t)-\phi & i=\sigma\\ x_n(t) & i=n \qquad (x_n(t)=\phi)\\ 0 & \text{otherwise}\end{cases}$$

$$\text{and for } -j\le t<-1,\quad c_i(t)=\begin{cases} k\phi+\phi & i=t+\sigma\\ x_\sigma(t)-\phi & i=\sigma\\ x_n(t) & i=n \qquad (x_n(t)=\phi)\\ 0 & \text{otherwise,}\end{cases}$$

generating the sequence $\{x(t)\}_{t=-\sigma+1}^{-1}$,

$$x(-\sigma+1) = \left(\frac{1}{\sigma}+\phi, \frac{1}{\sigma}-2k\phi, \underbrace{\frac{1}{\sigma}-k\phi, \ldots, \frac{1}{\sigma}-k\phi}_{j-1}, \underbrace{\frac{1}{\sigma}-k\phi+\phi, \ldots, \frac{1}{\sigma}-k\phi+\phi}_{\sigma-j-1}, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\right)$$

$$x(-\sigma+2) = \left(\underbrace{\frac{1}{\sigma}+\phi, \frac{1}{\sigma}+\phi}_{2}, \tfrac{1}{\sigma}-3k\phi, \underbrace{\frac{1}{\sigma}-k\phi, \ldots, \frac{1}{\sigma}-k\phi}_{j-1}, \underbrace{\frac{1}{\sigma}-k\phi+\phi, \ldots, \frac{1}{\sigma}-k\phi+\phi}_{\sigma-j-2}, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\right)$$

$$\vdots \qquad \vdots$$

$$x(-j) = \left(\underbrace{\frac{1}{\sigma}+\phi, \ldots, \frac{1}{\sigma}+\phi}_{\sigma-j}, \tfrac{1}{\sigma}-(\sigma-j+1)k\phi, \underbrace{\frac{1}{\sigma}-k\phi, \ldots, \frac{1}{\sigma}-k\phi}_{j-1}, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\right)$$

$$x(-j+1) = \left(\underbrace{\frac{1}{\sigma}+\phi, \ldots, \frac{1}{\sigma}+\phi}_{\sigma-j+1}, \tfrac{1}{\sigma}-(\sigma-j+2)k\phi-\phi, \underbrace{\frac{1}{\sigma}-k\phi, \ldots, \frac{1}{\sigma}-k\phi}_{j-2}, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\right)$$

$$\vdots \qquad \vdots$$

$$x(-1) = \left(\underbrace{\frac{1}{\sigma}+\phi, \ldots, \frac{1}{\sigma}+\phi}_{\sigma-1}, \tfrac{1}{\sigma}+(1-n)\phi, \underbrace{\phi, \ldots, \phi}_{n-\sigma}\right).$$

And, finally, harvesting

$$c(-1) = \left(\phi-\phi_2, \ldots, \phi-\phi_\sigma, \underbrace{\frac{1}{\sigma}+(1-n)\phi-\phi_{\sigma+1}}_{\geq 0(\sigma-th\ pos.)}, \phi-\phi_{\sigma+2}, \ldots, \phi\right)$$

we obtain $x(0)=x$. Again, the fact that $c(t)\geq 0$ for $t=-\sigma, \ldots, -1$ implies that the proposed sequence satisfies $(x(t), x(t+1))\in\Omega$.[17] □

Proof of Theorem 5: In view of Lemma 3 there is $\epsilon_1>0$ such that for each sequence $\{x(t)\}_{t=-\infty}^{\infty}$ satisfying $(x(t), x(t+1))\in\Omega$ and $\delta(x(t), x(t+1))\leq\epsilon_1$ the following inequality holds

$\|x(t)-\hat{x}\|_\infty \leq \epsilon$ for all integers t.

By Proposition 9 there is $\delta_1>0$ such that for each $x\in I\!R_+^n$ satisfying $\|x-\hat{x}\|_\infty\leq\delta_1$,

$\mu(x)\leq\epsilon_1$.

[17] We leave it to the reader to verify that the proposed sequence fulfills (*ii*) and (*iii*).

By Lemma 4 we know that there is δ_2 such that if $\|x-\hat{x}\|_\infty < \delta_2$, then the program $\{x(t)\}_{t=-\infty}^{0}$ defined in the lemma satisfies $\delta(x(t), x(t+1)) < \epsilon_1$ for all t.

Let $\delta = \min\{\delta_1, \delta_2\}$ and let x be any state $\|x-\hat{x}\|_\infty < \delta$ with $\{x(t)\}$ the corresponding optimal program from x. We build a sequence $\{x(t)\}_{t=-\infty}^{\infty}$ by concatenating the sequence given by Lemma 4 with the optimal program. We have the following properties

(i) $\mu(x) \leq \epsilon_1$ and, hence, $\delta(x(t), x(t+1)) < \epsilon_1$ for all $t \geq 0$.
(ii) $\delta(x(t), x(t+1)) < \epsilon_1$ for $t \in [-n, 0]$.
(iii) $\delta(x(t), x(t+1)) = 0$ for all $t < -n$.

Using (i) to (iii) and Lemma 3, we can write $\|x(t)-\hat{x}\|_\infty < \epsilon$ for all t. □

PROOF OF THEOREM 6: The preceding theorem guarantees that there exists $\delta > 0$ such that for each optimal program $\{x(t)\}$ satisfying $\|x(0)-\hat{x}\|_\infty \leq \delta$ we have:

$$\|x(t)-\hat{x}\|_\infty \leq \epsilon \text{ for all integers } t \geq 0. \tag{16}$$

We show now that there is a natural number τ_0 such that:

(P) For each optimal program $\{x(t)\}$ there exists an integer t' such that $0 \leq t' \leq \tau_0$ and $\|x(t')-\hat{x}\|_\infty < \delta$.

Let us assume the converse. Then for each natural number k there exists an optimal program $\{x(t)\}$ such that

$$\|x^k(t)-\hat{x}\|_\infty \geq \delta \text{ for all } t = 0, \dots, k. \tag{17}$$

Proposition 8 implies that there is D_0 such that $\mu(z) \leq D_0$ for all $z \in \Delta$. This property together with (17) and Theorem 4 imply that for each natural number k

$$\sum_{t=0}^{\infty} \delta(x(t), x(t+1)) = \mu(x^k(0)) \leq D_0. \tag{18}$$

Extracting subsequences, reindexing and using a diagonalization process we obtain that there exists a strictly increasing sequence of natural numbers $\{k_j\}_{j \in I\!N}$ such that for each integer $s \geq 0$ there exists

$$\tilde{x}(s) = \lim_{j \to \infty} x^{k_j}(s). \tag{19}$$

It is not difficult to see that $\{\tilde{x}(t)\}_{t=0}^{\infty}$ is a program. In view of (18), (19) and the continuity of the function $\delta(\cdot, \cdot)$:

$$\sum_{t=0}^{\infty} \delta(\tilde{x}(t), \tilde{x}(t+1)) \leq D_0;$$

namely, $\{\tilde{x}(t)\}_{t=0}^{\infty}$ is a good program. Then, Theorem 2 implies that $\lim_t \tilde{x}(t) = \hat{x}$.

However, it follows from (17) and (19) that $\|\tilde{x}(t)-\hat{x}\|_\infty \geq \delta$ for all $t \geq 0$. The contradiction we have reached proves that there is a natural number τ_0 such that property (P) holds.

Now assume that $\{x(t)\}$ is an optimal program. By property (P), there is an integer $t_0 \in [0, \tau_0]$ such that $\|x(t_0) - \hat{x}\|_\infty < \delta$. Clearly, the program $\{x(t + t_0)\}_{t \geq 0}$ is also optimal. The last inequality and the choice of δ (see (16)) imply:

$$\|x(t + t_0) - \hat{x}\|_\infty \leq \epsilon \text{ for all } t \geq 0. \qquad \square$$

We conclude this section with an elaboration of Remark 3 as

Remark 6 As indicated in Remark 5, with explicit assumptions on the curvature of the felicity function $w(\cdot)$, one could furnish a bound on δ as a function of ϵ in Proposition 7. In any case, it is less than $(1/n\sigma)$. However, the proof of Lemma 3, and those of Theorems 5 and 6, to the extent that they rely on Lemma 3, is totally non-constructive. Hence, from a computational point of view, a constructive proof of Lemma 3 remains an open question.[18]

References

Brock, W. A. (1970), "On existence of weakly maximal programmes in a multi-sector economy," *Review of Economic Studies* **37**, 275–80.

Gale, D. (1967), "On optimal development in a multi-sector economy," *Review of Economic Studies* **34**, 1–18.

Khan, M. Ali, and T. Mitra (2005), "Choice of technique in the Robinson–Solow–Srinivasan model," *International Journal of Economic Theory* **1**, 83–109.

Khan, M. Ali, and T. Mitra (2008), "Growth in the Robinson–Solow–Srinivasan model: Undiscounted optimal policy with a strictly concave welfare function," *Journal of Mathematical Economics* **44**, 707–32.

Khan, M. Ali, and A. J. Zaslavski (2007), "On a uniform turnpike of the third kind in the Robinson–Solow–Srinivasan model," *Journal of Economics* **92**, 137–66.

Le Van, C., R. A. Dana, T. Mitra, and K. Nishimura (2006), *Handbook of Optimal Growth: The Discrete Time Horizon*, Vol. I, Berlin: Springer-Verlag.

McKenzie, L. W. (1968), "Accumulation programs of maximum utility and the von Neumann facet," J. N. Wolfe, ed., *Value, Capital and Growth*, 353–83, Edinburgh: Edinburgh University Press.

McKenzie, L. W. (1986), "Optimal economic growth, turnpike theorems and comparative dynamics," K. J. Arrow, and M. Intrilligator, eds, *Handbook of Mathematical Economics*, Vol. III, 1281–355, Amsterdam: North-Holland.

McKenzie, L. W. (2002), *Classical General Equilibrium Theory*, Cambridge: MIT Press.

Mitra, T. (2005), "Intergenerational equity and the forest management problem," S. Kant, and R. A. Berry, eds, *Economics, Sustainability and Natural Resources: Economics of Sustainable Forest Mangement*, 137–73, Dordrecht: Springer-Verlag.

Mitra, T. (2006) "Notes on a model of forestry: The undiscounted case," Cornell University, 10 June, mimeo.

Mitra, T., and H. Wan Jr. (1985), "Some theoretical results on the economics of forestry," *Review of Economics Studies* **52**, 263–82.

Mitra, T., and H. Wan Jr. (1986), "On the Faustmann solution to the forest management problem," *Journal of Economic Theory* **40**, 229–49.

Samuelson, P. A. (1976), "Economics of forestry in an evolving society," *Economic Inquiry* **14**, 466–92.

Zaslavski, A. J. (2005), "Optimal programs in the RSS model," *International Journal of Economic Theory* **1**, 151–65.

[18] As in footnote 19, we thank an anonymous referee for alerting us to these computational considerations.