Set-Valued Anal (2011) 19:135–156
DOI 10.1007/s11228-010-0168-2

The Concavity Assumption on Felicities and Asymptotic Dynamics in the RSS Model

M. Ali Khan · Adriana Piazza

Received: 30 November 2009 / Accepted: 18 October 2010 /
Published online: 30 October 2010

Abstract An analysis of the RSS model in mathematical economics involves the study of an infinite-horizon variational problem in discrete time. Under the assumption that the felicity function is upper semicontinuous and "supported" at the value of the maximally-sustainable level of a production good, we report a generalization of results on the equivalence, existence and asymptotic convergence of optimal trajectories in this model. We consider two parametric specifications, and under the second, identify a "symmetry" condition on the zeroes of a "discrepancy function" underlying the objective function that proves to be necessary and sufficient for the asymptotic convergence of good programs. With a concave objective function, as is standard in the antecedent literature, we show that the symmetry condition reduces to an equivalent "non-interiority" condition.

Keywords Good program · Maximal program · Optimal program · Value-loss · Non-differentiability · Discrepancy function · Non-interiority · Existence of optimal programs · Asymptotic convergence

JEL Classification C62 · D90

Mathematics Subject Classifications (2010) 52A41 · 91B55 · 49J45 · 37B25 · 39A06

M. A. Khan (✉)
Department of Economics, The Johns Hopkins University, Baltimore, MD 21218, USA
e-mail: akhan@jhu.edu

A. Piazza
Departamento de Matemática, Universidad Técnica Federico Santa María, Avda. España 1680, Casilla 110-V, Valparaíso, Chile
e-mail: adriana.piazza@usm.cl

1 Introduction

The mathematics of the general theory of (undiscounted) intertemporal resource allocation is developed in terms the triple (Ω, u, x_0) where $\Omega \subseteq I\!R^n_+ \times I\!R^n_+$, $u : \Omega \longrightarrow I\!R_+$ and $x_o \in I\!R_+$, where $I\!R^n$ is $n-$dimensional Euclidean space and $I\!R^n_+$ its non-negative orthant. The set Ω is referred to as the *transition possibility set* and its individual element (x, x') is interpreted as a pair of states such that the state of the system tomorrow x' is technologically feasible given the state of the system today x. The real-valued function u defined on the feasible pair of states (x, x') is interpreted as the benefit $u(x, x')$ received today. The vector x_o is interpreted as the initial state of the dynamical system in which time is formalized in discrete units $t \in I\!N$, where $I\!N$ is the set of natural numbers, including zero. The theory involves the following basic concepts.

Definition 1.1 A *program* from the initial state x_0 is a sequence $\{x(t)\}$ such that $x(0) = x_0$ and for all $t \in I\!N$, $\big(x(t), x(t+1)\big) \in \Omega$.

Definition 1.2 A program $\{x^*(t)\}$ is *optimal* if for any program $\{x(t)\}$ such that $x(0) = x^*(0)$ we have

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

A program is *maximal* if the lim sup operator above is substituted for the lim inf operator.[1]

Definition 1.3 A program $\{x^*(t)\}$ is *stationary* if $x^*(t)$ is constant for all $t \in I\!N$. A program is a *stationary optimal (maximal)* program if it is stationary and optimal (maximal).

The basic results of the theory are now well-known and available in [23, Section II.6] in the form of two theorems; also see the texts [1, 2, 32]. The first result (Theorem 6) asserts the existence of a stationary optimal program, as well as that of an optimal program from any initial state, under convexity, closedness and boundedness assumptions on Ω, and on strict concavity and continuity assumptions on u. The second result (Theorem 7) asserts the existence of a maximal program by weakening the strict concavity of u to concavity and by requiring the additional assumption of uniqueness of the stationary maximal program. We refer the reader to [23] and his references for a precise statement and bibliographic details.

In [8], a model due to Robinson, Solow and Srinivasan (referred to as the RSS model and to be precisely defined below) is reformulated so that its objects conform to the basic theory delineated above. It is shown that the general theorem on the existence of optimal programs does not directly apply by virtue of the fact that the reformulated benefit function of the RSS model is never strictly concave. A more direct mathematical argumentation needs to be furnished, one that appeals to the methods of proof rather than the results themselves of the theory. Such an argument revolves

[1] In [4, Chapter 2], the authors refer to these notions as *overtaking optimal* and *weakly overtaking optimal* programs. There is a dissonance in terminological usage as between discrete and continuous time, but also see [27] for the case of continuous time.

around the model's original felicity function $w(\cdot)$ defined on consumption levels rather than on the reformulated benefit function $u(\cdot,\cdot)$ defined on feasible ordered pair of stocks, and referred to above. The authors of [8] also limit themselves to the maximality notion, and show that maximal and stationary maximal programs exist, and, under strict concavity of $w(\cdot)$, the former converge asymptotically to the latter provided it is unique. Two aspects of this reformulation have attracted attention. First, through three decisive counterexamples, the results are shown to conflict rather dramatically with analogous results for an identical model in continuous time; see [29] and [8, Section 6.2–6.4]. Second, a sufficient summary statistic ξ_σ (to be precisely defined below) has been identified, and shown to govern the dynamic properties of the maximal program under the linearity assumption on $w(\cdot)$. Subsequent to [8], these results have been refined and elaborated also for the optimality notion and towards the development of a rather complete (so-called) turnpike theory for the model; see [9–11, 13, 16, 31, 33]. Indeed, these results for the RSS model have been used towards the development of a general theory in [34]. Nonetheless, this dichotomy of one set of results for a linear case of $w(\cdot)$, and another for a strictly concave case, has remained a rather unattractive aspect of the theory. A natural question has remained unanswered as to whether the results can be unified through a condition that is sufficient, and perhaps even necessary, for a unification of the results.

We present such a mathematical condition in this paper. But perhaps more than the condition itself, what is of interest, and to the authors of some surprise, is the considerably generalized context in which such a condition can be set. The essential mathematical contribution of this paper is that rather than the linear and strictly concave cases of the felicity function functions handled separately, one need work only with upper semicontinuous felicities that are supported at the unique consumption value of the stationary maximal plan. In such a context, we present (i) an equivalence theorem (Theorem 6.1 below) that shows maximal and optimal programs to be identical and (ii) an asymptotic theorem (Theorem 5.1 below) establishing their convergence to the (unique) maximal or optimal stationary program if *either* the summary statistic ξ_σ identified above is not unity *or* the sufficient condition holds. And if the felicity functions are concave, the sufficient condition reduces to one that is necessary and sufficient for the asymptotic convergence of programs that include the optimal and maximal ones; see Condition 4 below and the discussion following it. It is important to be clear that, leave alone for the general theory, such results have not even been conjectured, much less formulated and proved, even for the RSS model. The extent to which the results of this paper can be lifted to the general theory must remain a tantalizing open question for future investigation.

With this introduction, the plan of the paper is straightforward. In Section 2, we present a brief but mathematically self-contained statement of the RSS model as a special case of the general model specified above. Section 3 focusses on the golden-rule configurations and the underlying value-loss function. Sections 4 and 5 concern good programs, the former proves the existence of good programs while the later characterizes their asymptotic convergence. Section 6 presents the principal equivalence between maximal, optimal and minimal value loss programs. Finally, we conclude the paper in Section 7 by relating our results in more detail to the continuous-time literature in mathematical economics and optimal control theory.

2 The RSS Model

We begin with a brief description of the RSS model and its basic benchmarks as the special case of the model considered in the introduction; the reader is referred to [8, 10, 34] for more details.

Let the total labor force of the economy be stationary and positive, we normalize it to be unity. We work with a model of an economy capable of producing n types of machines using only labor: $a_i > 0$ units of perfectly divisible labor are needed to produce 1 unit of a perfectly divisible machine of type i, for $i = 1, \dots, n$. Let a denote the n-vector $(a_1, \cdots, a_n)$. The *state* of the economy will be represented by a point $x \in \mathbb{R}^n_+$, where x_i represents the stock of machines of type $i = 1, \dots, n$.[2] If x represents the stock of machines today, x' the stock tomorrow and $d \in (0, 1)$ the common rate of depreciation of machines of any type, then $z = x' - (1 - d)x$ stands for the number of machines produced during the period. The constraints in the $z \geq 0$ and $az \leq 1$ represent respectively the irreversibility of investment and the maximum labor available. With these definitions, the *transition possibility set* is given by

$$\Omega = \left\{(x, x') \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ : x' - (1 - d)x \geq 0 \text{ and } a\left(x' - (1 - d)x\right) \leq 1\right\}. \qquad (1)$$

One unit of labor together with one unit of machine of type $i = 1, \dots, n$, can produce $b_i > 0$ units of a single consumption good, thus, defining the output-coefficients vector $b = (b_1, \dots, b_n) \in \mathbb{R}^n_+$, the total good production is given by by, with $y \in \Lambda(x, x')$.

Given the pair $(x, x') \in \Omega$ and denoting by e the sum vector of $\mathbb{R}^n$, the stock of machines that may be devoted to the consumption goods sector is given by the correspondence

$$\Lambda(x, x') = \left\{y \in \mathbb{R}^n : 0 \leq y \leq x \text{ and } ey \leq 1 - a\left(x' - (1 - d)x\right)\right\}. \qquad (2)$$

Welfare is derived only from the consumption good and is represented by an upper semicontinuous felicity function $w : \mathbb{R}_+ \to \mathbb{R}_+$ that is assumed to be "supported" at one particular consumption level, the so-called *golden-rule* consumption level, to be fully characterized in the following section. In terms of the triple (Ω, u, x_o) pertaining to the general theory, and mentioned in the first sentence of the introduction, we can now define

$$u(x, x') = \arg\max\left\{w(by) : y \in \Lambda(x, x')\right\} \text{ for all } (x, x') \in \Omega.$$

In the antecedent literature on the RSS model, it is usual to define a program as $\{x(t), y(t)\}_{t \in \mathbb{N}}$ where $y(t)$ indicates the part of the stock devoted to the production of the consumption good, $y(t) \in \Lambda\left(x(t), x(t+1)\right)$. We take a different approach, defining

[2] Throughout this work we sometimes refer to the state x as stock of machines and sometimes simply as *stock*.

a program as a sequence of states $\{x(t)\}$ without specifying explicitly the sequence $\{y(t)\}$. Towards this end, we focus on the subset of $\Lambda(x, x')$ given by

$$\lambda(x, x') = \arg\max \left\{ b\,y : y \in \Lambda(x, x') \right\}.$$

In this formulation, the production levels at time t always belong to $\lambda\big(x(t), x(t+1)\big)$, i.e., they always yield the maximal possible total production.[3] It may well be that $\lambda\big(x(t), x(t+1)\big)$ is a correspondence rather than a function but given that every $y \in \lambda\big(x(t), x(t+1)\big)$ provides the same total production ($b\,y$ is constant for all $y \in \lambda(x, x')$) we are indifferent to the actual allocation of machines and labor to produce consumption goods. By abuse of notation we will sometimes refer to $\max\{b\,y : y \in \Lambda(x, x')\}$ as $b\,\lambda(x, x')$ even though $\lambda(x, x')$ is not always a single element vector set.

Since we shall be working directly with $w(\cdot)$ rather than $u(\cdot, \cdot)$, we modify Definition 1.2 above to

Definition 2.1 A program $\{x^*(t)\}$ is *optimal* if for any program $\{x(t)\}$ such that $x(0) = x^*(0)$ we have

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

where $y(t) \in \lambda\big(x(t), x(t+1)\big)$ and $y^*(t) \in \lambda\big(x^*(t), x^*(t+1)\big)$. A program is *maximal* if the lim sup operator above is substituted for lim inf.

Definitions 1.1 and 1.3 remain the same.

3 Golden-Rule Stocks and Minimum Value-Loss Programs

In this section, we develop three fundamental concepts for the RSS model that are based on the "supportability" assumption: the golden-rule stocks, the value-loss function and the von Neumann facet. Note that we depart from the literature in that we first propose a value-loss function and then use this function to characterize the golden rule stocks and von-Neumann facet. This is done to accommodate our non-concave, non-monotonic and non-differentiable felicities. We leave as exercises for the reader proofs which routinely extend the standard results.

As referenced in [8], the quotient $(1 + a_i d)/b_i$ represents the effective labor cost of producing a unit of output, and its reciprocal $b_i/(1 + a_i d)$, the maximal sustainable amount of output, both estimates with reference to a machine of type *i*.

Definition 3.1 Let the *golden-rule* consumption level $\hat{c}$ be given by $\max\{b_i/(1 + a_i d) : i = 1, \ldots, n\}$, and let $I \subseteq [1, \ldots, n]$ be the set of indices where the maximum is attained, i.e., $I = \{i : \hat{c} = b_i/(1 + a_i d)\} = \arg\max\{b_i/(1 + a_i d) : i = 1, \ldots, n\}$.

[3] This change of approach implies no loss in generality. Indeed, consider any program defined in the traditional way $\{x(y), y(t)\}_{t\in I\!N}$ with $y(t) \in \Lambda(x(t), x(t+1))$, and suppose that there is $\tilde{t}$ such that $y(\tilde{t}) \notin \lambda(x(\tilde{t}), x(\tilde{t}+1))$. There exists another program $\{x(t), y'(t)\}_{t\in I\!N}$ with $y'(t) \in \lambda(x(\tilde{t}), x(\tilde{t}+1))$ (observe that the sequence of stock of machines $\{x(t)\}$ remains unchanged). The new program provides a strictly larger benefit at $t = \tilde{t}$ and the same benefit if $t \neq \tilde{t}$. Hence, it is obvious that $\{x(y), y(t)\}_{t\in I\!N}$ cannot be maximal or optimal. Even more, there is no need to consider it as an alternative program in the definition of maximal or optimal program, it suffices to consider only $\{x(y), y'(t)\}_{t\in I\!N}$.

We can now state the already-announced supportability condition on the felicity function w, that will be used throughout this work:

Supportability condition on the felicity function *There exists* $\pi > 0$ *such that* $w(c) \leq w(\hat{c}) + \pi(c - \hat{c})$ *for all* $c \in I\!R_+$. *Let* $S_w \subseteq I\!R_+$ *be the set where the equality is attained,*

$$S_w = \{c \in I\!R_+ : w(c) = w(\hat{c}) + \pi(c - \hat{c})\} \tag{3}$$

The set S_w is closed by the upper semi-continuity of the function w. Furthermore, in the particular cases where w is linear, concave or strictly concave, the set S_w is respectively $I\!R_+$, a closed interval or $\{\hat{c}\}$. In fact, the identity $S_w = \{\hat{c}\}$ corresponds to the family of *strictly-supported* felicity functions,

$$S_w = \{\hat{c}\} \iff w(c) < w(\hat{c}) + \pi(c - \hat{c}) \qquad \text{for all } c \neq \hat{c}.$$

Remark 3.1 Note that the supportability assumption is based on the primitive parameters of the RSS model; namely $(a, b, w(\cdot))$.

Remark 3.2 Note that our notion of strictly-supported felicity function does not necessarily imply a unique support, as in Fig. 1c.

Definition 3.2 Let $\hat{p} = \hat{c}a \in I\!R^n_+$. The corresponding *value loss* associated with any $(x, x') \in \Omega$ is given by

$$\delta(x, x') = w(\hat{c}) - w(b\lambda(x, x')) - \pi\hat{p}\,(x' - x).$$

The following lemma proves that $\delta(x, x') \geq 0$ for any $(x, x') \in \Omega$, and also determines an inequality on the stocks that will be used afterwards in the characterization of the von Neumann facet. Remark 3.4 below notes the difference between the value-loss function used in the antecedent literature stemming from [8], when w is assumed to be either linear or strictly concave and differentiable, and our own.

Proposition 3.1 *For any* $(x, x') \in \Omega$ *we have* $\delta(x, x') \geq 0$.

Proof By the supportability condition on the felicity function, we obtain

$$w(\hat{c}) - w(by) \geq \pi\left(\hat{c} - by\right) \text{ for all } y \in \Lambda(x, x'), \text{ for all } (x, x') \in \Omega. \tag{4}$$

We can now appeal to the following identity

$$\begin{aligned}\hat{c} - by &= \sum_{i=1}^{n}\left(\hat{c} - \frac{b_i}{1 + da_i}\right)y_i + \hat{c}(1 - ey) + \sum_{i=1}^{n}\left(\frac{b_i}{1 + da_i} - b_i\right)y_i \\ &= \sum_{i=1}^{n}\left(\hat{c} - \frac{b_i}{1 + da_i}\right)y_i + \hat{c}(1 - ey) - \sum_{i=1}^{n}\frac{b_i a_i d}{1 + da_i}y_i\end{aligned} \tag{5}$$

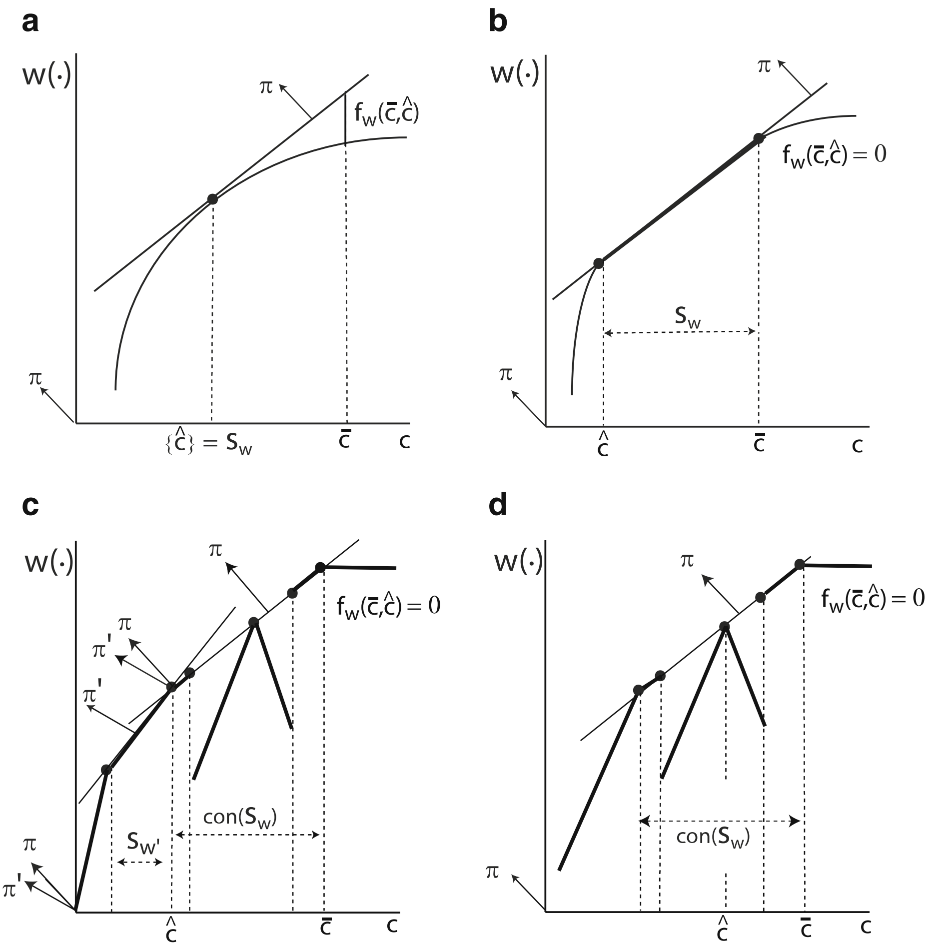

Fig. 1 Interiority and non-interiority of $\hat{c}$ in con(S_w)

Since $y \in \Lambda(x, x')$, we obtain from (2), $1 - ey \geq a(x' - x) + dax \geq a(x' - x) + day$. On substituting this in (5), we obtain

$$\hat{c} - b\,y \geq \sum_{i=1}^{n} \left(\hat{c} - \frac{b_i}{1 + da_i} \right) y_i + \hat{c}\,a(x' - x) + \hat{c}\,day - \sum_{i=1}^{n} \frac{b_i a_i d}{1 + da_i} y_i$$

$$\geq \hat{c}\,a(x' - x) + \sum_{i=1}^{n} \left(\hat{c} - \frac{b_i}{1 + da_i} \right) (1 + a_i d) y_i$$

$$\geq \hat{c}\,a(x' - x) = \hat{p}\,(x' - x) \tag{6}$$

But then this implies from (4)

$$w(\hat{c}) - w(b\,y) \geq \pi\,\hat{p}(x' - x),$$

and hence that $\delta(x, x') \geq 0$. The proof is finished. □

Remark 3.3 Note that in contrast to the antecedent literature there is no presumption in the proof that the golden-rule consumption level $\hat{c}$ is attained at a unique stock $\hat{x}$.

Indeed, we can now define

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

$$(P) \quad \max\{b\lambda(x, x) : (x, x) \in \Omega\}.$$

Theorem 3.1 *The set of golden-rule stocks is*

$$\hat{X} = \left\{x \in \mathbb{R}^n_+ : (x, x) \in \Omega \text{ and } bx = \hat{c}\right\}.$$

Furthermore, for every $x \in \hat{X}$ *we have* $x_j = 0$ *for all* $j \notin I$. *When* I *is a singleton equal to* $\{\sigma\}$, *the golden-rule stock is unique and given by* $\hat{x} = (1/(1 + a_\sigma d))e_\sigma$, *where* e_σ *is the* σ*-th unit vector.*

Before getting into the proof of the theorem we present an easy technical lemma,

Lemma 3.1 *Let* $x \in X$ *such that* $(x, x) \in \Omega$. *Then*

(i) *If* $b\lambda(x, x) = \hat{c} \Longrightarrow \lambda(x, x) = \{x\}$
(ii) *If* $bx = \hat{c}$ *and* $x_j = 0$ *for all* $j \notin I \Longrightarrow \lambda(x, x) = \{x\}$

Proof

(i) Consider $y \in \lambda(x, x)$. One of the two equalities: (i) $x = y$ or (ii) $1 - adx - ey = 0$ must hold. Of course, if (i) holds, our claim is proved. Assuming that (i) does not hold, we have

$$\begin{aligned} ey &= 1 - adx < 1 - ady \leq 1 + \textstyle\sum_{i=1}^n \left(1 - \left(b_i/\hat{c}\right)\right) y_i \\ &= 1 + ey - \left(1/\hat{c}\right) \textstyle\sum_{i=1}^n b_i y_i = ey \end{aligned}$$

which is obviously absurd.

(ii) We claim that $x \in \Lambda(x, x)$. Indeed,

$$1 - a(x - (1-d)x) - ex = 1 - \textstyle\sum_{i \in I} (1 + da_i x_i) = 1 - \sum_{i \in I} \left(b_i/\hat{c}\right) x_i = 0. \quad (7)$$

Hence, $bx \leq \max\{by : y \in \Lambda(x, x)\} \leq \max\{by : y \leq x\} = bx$, implying $x \in \lambda(x, x)$ and $b\lambda(x, x) = \hat{c}$. Finally, by (i) we conclude. □

Proof of Theorem 3.1 To prove that $\hat{X}$ is the solution set to the optimization problem (P) in Definition 3.3, it suffices to check that $\alpha(P) = \hat{c}$, where $\alpha(P)$ denotes the optimal value of (P). Using (6) we get that $\alpha(P) \leq \hat{c}$. On the other hand, it is easy to check that given $\sigma \in I$ the state $x = 1/(1 + a_\sigma d)e(\sigma)$ satisfies the constraint of (P) and that $bx = \hat{c}$. Using Lemma 3.1(ii) we get that $b\lambda(x, x) = \hat{c} \leq \alpha(P)$. Thus, $\alpha(P) = \hat{c}$.

Given any $x \in \hat{X}$, we know that $\lambda(x, x) = \{x\}$ thanks to Lemma 3.1, which means that $0 \leq 1 - adx - ex$ and hence

$$0 \leq 1 - (ad + e)x \leq 1 - \textstyle\sum_{i=1}^n (b_i/\hat{c})x_i = 0$$

implying that all the inequalities above must hold with equality. Consequently, $x_j = 0$ for all $j \notin I$.

If $I = \{\sigma\}$, then $x_j = 0$ for $j \neq \sigma$ and $b_\sigma x_\sigma = \hat{c} \Longrightarrow x_\sigma = 1/(1 + a_\sigma d)$ and $\hat{X} = \{\hat{x}\}$. □

In the next proposition, we denote the convex hull of a set A by $co(A)$.

Proposition 3.2 $\hat{X} = co\big\{(1/(1 + a_i d))e_i : i \in I\big\}$.

Proof It is evident that $\{(1/(1 + a_i d))e(i) : i \in I\} \subseteq \hat{X}$, and as $\hat{X}$ is convex we readily conclude that $co\{(1/(1 + a_i d))e(i) : i \in I\} \subseteq \hat{X}$.

Let $x \in \hat{X}$ we want to prove that x can be written as $x = \sum_{i\in I} \lambda_i (1/(1 + a_i d))e(i)$, where $\lambda_i \geq 0$ and $\sum_{i\in I} \lambda_i = 1$. Thanks to Theorem 3.1 we know that $x_j = 0$ for all $j \notin I$. Then we have $(ad + e)x = \sum_{i\in I}(a_i d + 1)x_i = \sum_{i\in I}(b_i/\hat{c})x_i = 1$ and defining $\lambda_i = (a_i d + 1)x_i$, we conclude as desired. □

Proposition 3.3 *The von Neumann facet, i.e., the set of pairs* $(x, x') \in \Omega$ *such that* $\delta(x, x') = 0$ *is given by*

$$\Omega_{vN} = \Big\{(x, x') \in \Omega :\; 1 - ax' = [e - (1 - d)a]x,\;\; x_j = 0 \textit{ for all } j \notin I \;\textit{ and }\; bx \in S_w\Big\}.$$

Furthermore, $\{x\} = \lambda(x, x')$ *for all* $(x, x') \in \Omega_{vN}$.

Proof Using (6) we get for every $y \in \lambda(x, x')$,

$$\delta(x, x') = 0 \Longrightarrow w(\hat{c}) - w(by) = \pi \hat{p}\,(x' - x) \leq \pi(\hat{c} - by).$$

On the other hand, the Supportability Assumption yields: $w(\hat{c}) - w(by) \geq \pi(\hat{c} - by)$ and putting the two things together we get

$$\delta(x, x') = 0 \Longrightarrow w(\hat{c}) - w(by) = \pi \hat{p}\,(x' - x) = \pi(\hat{c} - by).$$

Considering the equation above and (2) we get for every $y \in \lambda(x, x')$,

$$\begin{aligned} 0 \leq 1 - a(x - x') - adx - ey \;&=\; 1 - \frac{1}{\hat{c}}\big(\hat{c} - by\big) - adx - ey \\ \leq \left[\frac{b}{\hat{c}} - (ad + e)\right] y \;&=\; \sum_{i=1}^{n} \frac{1 + a_i d}{\hat{c}} \left(\frac{b_i}{1 + a_i d} - \hat{c}\right) y_i \;\leq\; 0. \end{aligned}$$

Thus, in every step the equality must hold, implying that $x = \lambda(x, x')$ and $x_j = 0$ for all $j \notin I$ and $1 - ax' = [e - (1 - d)a]x$. And, from the definition of the set S_w we get $bx = by \in S_w$. The fact that $\{x\} = \lambda(x, x')$ can be deduced from $1 - ax' = [e-(1 - d)a]x$.

On the other hand, it is very easy to see that the points $(x, x') \in \Omega_{vN}$ satisfy that $\delta(x, x') = 0$, and the proposition follows. □

Corollary 3.1 $(x, x) \in \Omega_{vN}$ *for all* $x \in \hat{X}$.

Remark 3.4 In [8] and subsequent work, a different value-loss function is proposed in the differentiable framework. This function, adapted to our framework and notation would be,

$$\delta_{KM}(x, x') = w(\hat{c}) - w\left(b\,\lambda(x, x')\right) - \sum_{i=1}^{n} \pi\left((a_i c_i)\,/\,(1 + a_i d)\right)\left(x'_i - x_i\right).$$

It is easy to see that $\delta_{KM}(x, x') \geq \delta(x, x') \geq 0$ for all $(x, x') \in \Omega$. The corresponding von Neumann facet is

$$\left\{(x, x') \in \Omega_{vN} : x'_j = 0 \text{ for all } j \notin I\right\}.$$

The fact that the function $\delta_{KM}(x, x')$ produces a smaller von Neumann facet, suggests that stronger results can be obtained with it instead of $\delta(x, x')$. But this is not the case, because our proofs will not concern the von Neumann facet itself but the programs that produce zero accumulated value loss, and these coincide for both value loss functions.

Remark 3.5 Note that other than a subsidiary claim in Theorem 3.1, the concepts and results presented in this section do not rely on I being a singleton.

4 Good Programs: Existence and Limit Points

In this section, we present some preliminary results on *good* programs, a concept due to Gale. We supplement the methods of Brock and Gale by utilizing the concept of the set of limit points of such programs, a set that has been used to advantage in [31] and [12].

Definition 4.1 A program $\{x(t)\}$ is called *good* if there exists $M \in I\!R$ such that for all $T \geq 0$,

$$\sum_{t=0}^{T}\left[w\left(b\,\lambda\big(x(t), x(t{+}1)\big)\right) - w(\hat{c})\right] \geq M.$$

A program is *bad* if $\lim_{T\to\infty}\sum_{t=0}^{T}\left[w\left(b\,\lambda\big(x(t), x(t{+}1)\big)\right) - w(\hat{c})\right] = -\infty$.

As shown in [8, Proposition 4], the following general result of Gale applies.

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

We present here a proof of the existence of good programs from any initial stock, suitable to our non-concave, non-differentiable framework. The majority of the proofs presented in the literature are based on building a particular program that converges asymptotically, and sufficiently fast, to the golden rule stock. Our approach again differs from this generality and is similar to the one presented in [22, Lemma 7.4.4] as we build a program that in a finite number of steps, reaches the golden rule consumption level. To this end, we consider $\sigma \in I$ and by investing in an appropriate way, we assure that $x_\sigma(t) = 1/(1 + a_\sigma d)$ for all $t \geq T$ for some finite $T > 0$ that depends on the initial stock x. After T, the investment and production of consumption goods is made in a way such that not only the value of $x_\sigma(t)$ remains constant but the benefit at each step is the golden-rule consumption level $\hat{c}$.

Proposition 4.2 *There exists a good program from any arbitrary initial stock* $x(0) \in \mathbb{R}^n_+$.

Proof Let $\sigma \in I$. We build a program $\{x(t)\}$ that does not invest in machines of type $j \neq \sigma$. Hence, $x_j(t) = (1-d)^t x_j(0) \to 0$ for all $j \neq \sigma, t \in \mathbb{N}$.

1. If $x_\sigma(0) < \hat{x}_\sigma$ we initially apply an investment policy where the totality of labor is devoted to building machines of type σ, $z(t) = 1/a_\sigma e(\sigma)$. This yields the following

$$x_\sigma(t) = (1-d)^t x_\sigma(0) + \frac{1}{a_\sigma} \sum_{i=0}^{t-1} (1-d)^t = (1-d)^t x_\sigma(0) + \tfrac{1}{da_\sigma}\left[1-(1-d)^t\right]. \tag{8}$$

This policy is kept until the first stage, t_0, such that the increasing sequence $\{x_\sigma(t)\}$ satisfies

$$\hat{x}_\sigma - (1-d)x_\sigma(t_0) < 1/a_\sigma \tag{9}$$

Indeed, the condition above is fulfilled by a finite value of t_0, because $x_\sigma(t) \to_{t\to\infty} 1/(da_\sigma)$ and $\hat{x}_\sigma - (1-d)/(da_\sigma) < 1/a_\sigma$. At t_0, the investment is $z(t_0) = [\hat{x}_\sigma - (1-d)x_\sigma(t_0)]e(\sigma)$ assuring that $x_\sigma(t_0+1) = \hat{x}_\sigma$. From there on, the investment is kept at the level $z(t) = d\hat{x}_\sigma e(\sigma)$ and we get $x_\sigma(t) = \hat{x}_\sigma$ for all $t > t_0$.
It can be seen that the evolution of the sequence $x_\sigma(t)$ can be compactly expressed as $x_\sigma(t+1) = \min\{(1-d)x_\sigma(t) + 1/a_\sigma,\ \hat{x}_\sigma\}$ and the corresponding investment policy as $z_\sigma(t) = \min\{1/a_\sigma,\ \hat{x}_\sigma - (1-d)x_\sigma(t)\}$.
2. If $x_\sigma(0) > \hat{x}_\sigma$ we set the investment policy $z(t) = 0$ until the first time stage t_1 when the stock of machines of type σ, $x_\sigma(t) = (1-d)^t x_\sigma(0)$, falls below the golden rule level $\hat{x}_\sigma$. We then invest $z_\sigma(t_1) = \hat{x}_\sigma - (1-d)x_\sigma(t_1)$ to get $x_\sigma(t_1+1) = \hat{x}_\sigma$. A brief computation shows that $x_\sigma(t_1) \le \hat{x}_\sigma \le x_\sigma(t_1-1)$ implies that $\hat{x}_\sigma - (1-d)x_\sigma(t_1) \le 1/a_\sigma$, hence the proposed investment is feasible. For all $t > t_1$, we set $z(t) = d\hat{x}_\sigma e(\sigma)$ assuring that $x_\sigma(t) = \hat{x}_\sigma$ for all $t > t_1$.
3. If $x_\sigma(0) = \hat{x}_\sigma$, we then set $z_\sigma(t) = d\hat{x}_\sigma$ for all t getting $x_\sigma(t) = \hat{x}_\sigma$ for all t.

Let $\hat{t}$ be the time when $x_\sigma(t)$ reaches $\hat{x}_\sigma$ (according to the initial condition: $\hat{t} = t_0+1$ or $\hat{t} = t_1+1$ or $\hat{t} = 0$). As we saw above $x_\sigma(t) = \hat{x}_\sigma$ for all $t \ge \hat{t}$. It is easy to see that $y(t) = \hat{x}_\sigma e(\sigma) \in \lambda\big(x(t), x(t+1)\big)$ for all $t \ge \hat{t}$. For every $T \ge \hat{t}$ we have,

$$\begin{aligned}\sum_{t=0}^{T} w\left(b\lambda\big(x(t), x(t+1)\big)\right) - w(\hat{c}) &\ge \sum_{t=0}^{\hat{t}-1} w\left(b\lambda\big(x(t), x(t+1)\big)\right) - w(\hat{c}) \\ &\quad + \sum_{t=\hat{t}}^{T} \underbrace{w\left(b\hat{x}_\sigma e(\sigma)\right) - w(\hat{c})}_{=0} \\ &= \sum_{t=0}^{\hat{t}-1} w\left(b\lambda\big(x(t), x(t+1)\big)\right) - w(\hat{c}) \in \mathbb{R}.\end{aligned}$$

From this, we can conclude that the proposed program is good. □

It is interesting to notice that an explicit superior bound of $\hat{t}$ can be easily found, and we supplement Proposition 4.2 by such a result.

For example, in case 1 where $x_\sigma(0) < \hat{x}_\sigma$, condition (9) is fulfilled if t satisfies

$$t + 1 \geq \ln\left[(1 + da_\sigma)\,(1 - da_\sigma x_\sigma(0))\right] / \ln\left[(1-d)^{-1}\right]$$

Hence the value of t_0 is

$$\begin{aligned} t_0 &= \left\lfloor \ln\left[(1 + da_\sigma)\,(1 - da_\sigma x_\sigma(0))\right] / \ln\left[(1-d)^{-1}\right]\right\rfloor \\ &\leq \ln(1 + da_\sigma)/\ln\left[(1-d)^{-1}\right] \qquad \text{for all } x_\sigma(0) \leq \hat{x}_\sigma. \end{aligned} \tag{10}$$

In case 2, where $x_\sigma(0) > \hat{x}_\sigma$, we can only find a bound depending on the value of $x_\sigma(0)$. In fact, it is easy to see that the value of t_1 must be

$$t_1 \leq \ln\left[x_\sigma(0)(1 + da_\sigma)\right] / \ln\left[(1-d)^{-1}\right] + 1. \tag{11}$$

Expressions (10) and (11) lead us to the following corollary:

Corollary 4.1 *For every initial state x_0, let*

$$T = \max\left\{\left\lfloor \ln(1 + da_\sigma) / \ln\left[(1-d)^{-1}\right]\right\rfloor\ ,\ \left\lceil \ln\left[x_\sigma(0)\,(1 + da_\sigma)\right] / \ln\left[(1-d)^{-1}\right]\right\rceil\right\}.$$

There is a good program $\{x(t)\}$ from x_0 satisfying $x_\sigma(t) = \hat{x}_\sigma$ for all $t \geq T$.

The proposition below states the classical equivalence between good programs and finite aggregate value loss programs, a proof adapted to the RSS model can be found in [8, Proposition 7].

Proposition 4.3 *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.*

We can now establish the existence of a program that attains minimum aggregate value-loss, for every initial stock. This is a benchmark result in the literature, but the proof we present adapts an argument in [6, Proposition 1.4.2], circumventing Cantor's diagonalization argument in [3], and following him, in [26] and [8].

Proposition 4.4 *From any $x_0 \in \mathbb{R}^n_+$ there exists a program $\{x(t)\}$ starting from x_0 that gives a minimal aggregate value-loss among all programs that start from x_0.*

$$\sum_{t=0}^{\infty} \delta\big(x(t), x(t+1)\big) = \Delta(x_0). \tag{12}$$

Proof As we stated above, the proof can be adapted from [6, Proposition 1.4.2]. Two remarks are in order. First, the proof presented in [6] considers continuous functions, whereas we deal with $\delta\big(x(t), x(t+1)\big)$ lower semi-continuous. It is an easy Analysis exercise to check that the proof is still valid with this weaker assumption. Second, the proof we are referring to is presented in a compact state space. We appeal to [8, Proposition 1] where it is proved that it is possible to define for each initial stock x_0 a compact set $X(x_0) \subseteq \mathbb{R}^n_+$ such that any program that starts from x_0 stays in $X(x_0)$. □

The aggregate value loss is always non-negative, and we are particularly interested in characterizing every program and every sequence $\{x(t)\}_{t=-\infty}^{\infty}$ with zero aggregate value loss, to make use of the following lemma taken from Zaslavski [31, Lemma 3.1]. We state a version adapted to our notation,[4]

Lemma 4.1 *Assume that* $\{x(t)\}$ *is a good program and that* (x, x') *is an accumulation point of* $\big(x(t), x(t+1)\big)$. *Then* $(x, x') \in \Omega$ *and* $\delta(x, x') = 0$, *and there exists a sequence* $\{u(t)\}_{t=-\infty}^{\infty}$ *such that* $\big(u(0), u(1)\big) = (x, x')$ *and*

$$\big(u(s), u(s+1)\big) \text{ are accumulation points of } \big(x(t), x(t+1)\big) \text{ for all } s.$$

In particular, the sequence defined above fulfills $\big(u(s), u(s+1)\big) \in \Omega$ and $\delta\big(u(s), u(s+1)\big) = 0$ for all integer s. We will use Lemma 4.1 to characterize the set of accumulation points of good programs:

Definition 4.2 $V = \{x : x$is an accumulation point of a good program.$\}$

Evidently, $\hat{X} \subseteq V$. Indeed, by Corollary 3.1 we know that for every state $x \in \hat{X}$, the constant program $x(t) = x$ for all t has zero aggregate value loss, and therefore it is good. The question arises as to the converse, and we present below a sufficient condition that does not rely on the fact that the golden-rule stock is a singleton.

Condition 1 $S_w = \{\hat{c}\}$.

Proposition 4.5 *If Condition* 1 *holds, then* $\hat{X} = V$.

Proof We already know that $\hat{X} \subseteq V$. To see the other inclusion, consider $x \in V$. Take x' such that (x, x') is an accumulation point of $\big(x(t), x(t+1)\big)$.[5] Lemma 4.1 yields that $\delta(x, x') = 0$. Using Proposition 3.3 we conclude that $x_j = 0$, $j \notin I$ and $bx \in S_w = \{\hat{c}\}$. To prove that $x \in \hat{X}$, we only need to show that $(x, x) \in \Omega$. The first inequality in the definition of Ω (see (1)) follows trivially, and to see the second and last one, observe that

$$1 - a\left[x - (1-d)x\right] = 1 - adx = 1 - \textstyle\sum_{i\in I} a_i d x_i = 1 + \sum_{i\in I}\left(1 - \frac{b_i}{\hat{c}}\right)x_i = \sum_{i\in I} x_i \geq 0$$

□

Our final proposition concerns the convergence of the Cesàro means of every good program. To the best of our knowledge the proofs presented in the literature rely heavily on the concavity of the felicity function, we provide a new proof suitable to our non-concave framework. Note also that for the moment, we continue to avoid the assumption that the golden-rule stock is a singleton.

[4]The proof of the lemma can easily be adapted to our framework from [31, Lemma 3.1] by observing that, neither the concavity of the function nor its differentiability play a role in the proof, and that every time that the continuity of w is used, it is actually sufficient to require only the lower semi-continuity of w to obtain the same result.

[5]We appeal again, as in the proof of Proposition 4.4, to the proof of [8, Proposition 4] where it is stated that the program $\{x(t)\}$ is included in a compact set $X(x_0)$.

Lemma 4.2 *The Cesàro means of every good program $\{x(t)\}$ converge to $\hat{X}$; namely,*

$$\bar{x}(t) = \frac{x(0) + \ldots + x(t-1)}{t} \quad \textit{is such that } dist(\bar{x}(t), \hat{X}) \to 0 \quad \textit{when } t \to \infty. \tag{13}$$

Proof We first observe that the convexity of Ω implies $(\bar{x}(t), \bar{x}'(t)) = \left(\frac{x(0)+\ldots+x(t-1)}{t}, \frac{x(1)+\ldots+x(t)}{t}\right) \in \Omega$. 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

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

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

$$(\bar{x}, \bar{x}) \text{ belongs to the closed set } \Omega. \tag{14}$$

Consider a good program $\{x(t)\}$ and let $y(t) = \lambda\big(x(t), x(t+1)\big)$, thanks to the Supportability Assumption we have

$$w(\hat{c}) - w(b\,y(t)) \geq \pi(\hat{c} - b\,y(t))$$

Adding the equation above from $t = 0, \ldots, T-1$ we obtain

$$-M \geq \sum_{t=0}^{T-1} w(\hat{c}) - w(b\,y(t)) \geq \pi\left[T\hat{c} - \sum_{t=0}^{T-1} b\,y(t)\right]$$

and dividing by $\pi T > 0$ we get

$$-M/\pi T + \frac{1}{T}\sum_{t=0}^{T-1} b\,y(t) \geq \hat{c}.$$

The concavity of $\lambda(x, x')$, yields $\lambda(\bar{x}(T), \bar{x}'(T)) \geq \frac{1}{T}\sum_{l=0}^{T-1} b\,y(t)$. These last two inequalities together yield

$$-M/\pi T + \lambda\left(\bar{x}(T), \bar{x}'(T)\right) \geq \hat{c}$$

and letting $T \to \infty$ we get

$$\lambda(\bar{x}, \bar{x}) = \lim_k \lambda\left(\bar{x}(t_k), \bar{x}'(t_k)\right) \geq \hat{c}$$

So we have that $(\bar{x}, \bar{x}) \in \Omega$ such that $\lambda(\bar{x}, \bar{x}) \geq \hat{c}$, implying that $\bar{x} \in \hat{X}$. □

5 Good Programs: Asymptotic Convergence

The literature on the RSS model uniformly assumes that there is a unique machine type at which the effective output per man is maximized. We name this additional assumption, the Brock-Koopmans-Mitra uniqueness condition.

Assumption 1 (Uniqueness) *There exists $\sigma \in \{1, \ldots, n\}$ such that $I = \{\sigma\}$.*

We saw in Theorem 3.1 that $\hat{X} = \{\hat{x}\} = \{1/(1 + a_\sigma d)e(\sigma)\}$ whenever the condition above is satisfied. And as has been understood since [3], this uniqueness result has a fundamental impact on the results reported here. We define the notion of *marginal rate of transformation* of type i machines today into type i machines tomorrow when there is no consumption or investment in machines of any other type. This is to say that we consider

$$\xi_i \equiv \big[(1/a_i) - (1-d)\big] > -1.$$

Proposition 5.1 *Let the sequence* $\{x(t)\}$ *be such that* $\sum_{t\in\mathbb{N}} \delta\big(x(t), x(t+1)\big) = 0$*. Then under Assumption* 1*, we obtain the following*

1. *If* $\xi_\sigma > 1,\ x(t) = \hat{x}$ *for all t.*
2. *If* $|\xi_\sigma| < 1,\ bx(t) \in S_w$ *for all t and* $\lim_{t\to\infty} x(t) = \hat{x}$*.*
3. *If* $\xi_\sigma = 1,\ b_\sigma x_\sigma(0) \in S_w \cap (\{(b_\sigma/a_\sigma)\} - S_w)$ *and*

$$x(t) = \begin{cases} x(0) & t = \dot{2} \\ 1/a - x(0) & t \neq \dot{2} \end{cases}$$

where the notation $\dot{n}$ *stands for a multiple of n.*

Proof Thanks to Proposition 3.3, we know that if $\delta\big(x(t), x(t+1)\big) = 0$ for all t, then

$$x_j(t) = 0 \text{ for all } j \neq \sigma, \text{ and } b_\sigma x_\sigma(t) \in S_w \tag{15}$$

and $1 - a_\sigma x_\sigma(t+1) = [1 - (1-d)a_\sigma]x_\sigma(t)$. From this last condition we easily get

$$\begin{aligned} x_\sigma(t) &= \frac{1}{a_\sigma} - \xi_\sigma x_\sigma(t-1) = \frac{1}{a_\sigma} \textstyle\sum_{i=0}^{t-1} (-\xi_\sigma)^i + (-\xi_\sigma)^t x_\sigma(0) \\ &= \frac{1}{a_\sigma}\frac{1-(-\xi_\sigma)^t}{1+\xi_\sigma} + (-\xi_\sigma)^t x_\sigma(0) \\ &= \frac{1}{1+a_\sigma d} + (-\xi_\sigma)^t \left[x_\sigma(0) - \frac{1}{1+a_\sigma d}\right] \end{aligned} \tag{16}$$

1. $\xi_\sigma > 1$. If the expression in brackets is different from zero, the value of $x_\sigma(t)$ oscillates without bounds, and it is negative for sufficiently large values of t, violating the definition of program. Hence, the only possibility is to have $x_\sigma(0) = \frac{1}{1+a_\sigma d} = x_\sigma(t)$ and thus $x(t) = \hat{x}$ for all t.
2. $|\xi_\sigma| < 1$. In this case, by (16) we have that $\lim_t x_\sigma(t) = \frac{1}{1+a_\sigma d}$.
3. $\xi_\sigma = 1$. By (16) we immediately get that $x_\sigma(t) = x_\sigma(0)$ for all $t = \dot{2}$ and $x_\sigma(t) = 1/a_\sigma - x_\sigma(0)$ for all $t \neq \dot{2}$ and then to assure that $b_\sigma x_\sigma(t) \in S_w$, it is enough to know that $b_\sigma x_\sigma(0) \in S_w$ and $b_\sigma(1/a_\sigma - x_\sigma(0)) \in S_w$. These two conditions can be equivalently expressed as $b_\sigma x_\sigma(0) \in S_w \cap (\frac{b_\sigma}{a_\sigma} - S_w)$. Observe that $\hat{c} \in S_w \cap (\frac{b_\sigma}{a_\sigma} - S_w)$, hence we retrieve what we already knew: $x(t) = \hat{x}$ for all t has zero aggregate value loss. □

We use this proposition and Lemma 4.1 to characterize V, recalling that $\hat{X} \subseteq V$ and that under Condition 1, $\hat{X} = \{\hat{x}\}$. We present the argument in three separate cases.

1. $|\xi_\sigma| > 1$. We claim that $\{\hat{x}\} = V$. Indeed, let us assume that there is $x \neq \hat{x}$, accumulation point of a good program $\{x(t)\}$. Take x' such that (x, x') is an accumulation point of $\big(x(t), x(t+1)\big)$.[6] From Lemma 4.1, there exists then a sequence $\{u(s)\}_{s=-\infty}^{\infty}$ such that $\big(u(s), u(s+1)\big)$, $\delta\big(u(s), u(s+1)\big) = 0$ for all integer s and $u(0) = x$. But due to (16) this sequence oscillates without bound when $s \to \infty$ and a contradiction arises.
2. $|\xi_\sigma| < 1$. As in the previous point, we claim that $\{\hat{x}\} = V$. The proof is similar to the one above, except that the contradiction arises for the negative values of s: there is no sequence $\{u(s)\}_{s=-\infty}^{\infty} \subset I\!R^n_+$ such that $u(0) = x \neq \hat{x}$ and (16) is fulfilled for all $s < 0$.
3. $|\xi_\sigma| = 1$. The third case offers a different result. We have already characterized the zero accumulated value loss programs as the 2-periodic programs that fulfill (15) and (16). Hence the set of accumulation points of zero aggregate value loss programs is

$$\left\{x \in I\!R^n_+ : \ x_j = 0 \text{ for all } j \notin I,\ bx \in S_w \cap (b_\sigma/a_\sigma - S_w)\right\}$$

and it is obviously included in V. We claim that this inclusion is an equality, i.e., every accumulation point of a good program is also an accumulation point of a zero aggregate value loss program. Indeed, if there was $x \notin V$ accumulation point of a good program, by Lemma 4.1, there should be a zero accumulated value loss sequence $\{u(s)\}_{s=-\infty}^{\infty}$ such that $u(0) = x$. Due to (15) and (16) we deduce: $u_j(s) = 0$ for all $j \neq \sigma$ and

$$u_\sigma(s) = \begin{cases} x_\sigma & s = \dot{2} \\ 1/a_\sigma - x_\sigma & s \neq \dot{2}. \end{cases}$$

Hence, x is an accumulation point of the zero value loss program $\{u(s)\}_{s=0}^{\infty}$, implying that $x \in V$.

In Cases 1 and 2, it is evident that along any good program $x(t) \to \hat{x}$. In the last case, we only know that $\mathrm{dist}\,(x(t), V) \to 0$. We summarize these results in the following theorem and we also show that in Case 3, $V = \{\hat{x}\}$ if and only if the following assumption is fulfilled

Condition 2 (Symmetry) $\{\hat{c}\} = S_w \cap (\frac{b_\sigma}{a_\sigma} - S_w)$

Theorem 5.1 *Let $\{x(t)\}$ be a good program. Then under Assumption 1, we have the following:*

1. *If $\xi_\sigma > 1$, $x(t) \to \hat{x}$.*
2. *If $|\xi_\sigma| < 1$, $x(t) \to \hat{x}$.*
3. *If $\xi_\sigma = 1$, $dist\,(x(t), V) \to 0$. In particular, if Condition 2 also holds, $x(t) \to \hat{x}$.*

[6] We appeal again, as in the proof of Proposition 4.4, to the proof of [8, Proposition 4] where it is stated that the program $\{x(t)\}$ is included in a compact set $X(x_0)$.

Proof Indeed, if the symmetry condition 2 holds, we have $b_\sigma x_\sigma(0) = b_\sigma x_\sigma(1) = \hat{c}$ and then $x_\sigma(t) = 1/(1 + a_\sigma d)$ for all t.

On the other hand, if there is $\mu \neq \hat{c}$ such that $\mu \in S_w \cap (b_\sigma/a_\sigma - S_w)$ then we can take $x(t) = \frac{c}{b_\sigma} e(\sigma)$ for all $t = \dot{2}$ and $x(t) = (\frac{1}{a_\sigma} - \frac{c}{b_\sigma})e(\sigma)$ for all $t \neq \dot{2}$. The program fulfills (16) and $b_\sigma x_\sigma \in S_w$ for all t, hence, it has zero aggregate value loss. □

Observe that if $\hat{c} \lesseqgtr b_\sigma x_\sigma(t)$, the opposite relation holds at the following step: $b_\sigma x_\sigma(t+1) = b_\sigma(\frac{1}{a_\sigma} - x_\sigma(t)) \lesseqgtr \hat{c}$. Hence, if $\hat{c}$ is not located between two elements of S_w there is no possibility of having two consecutive steps fulfilling $b_\sigma x_\sigma(t) \in S_w$, i.e., the following condition is sufficient to assure Condition 2.

Condition 3 (Non-interiority) $\hat{c} \notin \operatorname{int} \operatorname{co}(S_w)$.

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

Condition 4 (Non-Interiority, Concave Case) $\hat{c} \notin \operatorname{int} S_w$.

To see the necessity of this condition, let us assume that it is not fulfilled. Then there is $\epsilon > 0$ such that $[\hat{c} - \epsilon, \ \hat{c} + \epsilon] \subseteq S_w$ and it is easy to see that taking any initial state $x(0) \in [(\hat{c} - \epsilon)/b_\sigma, \ (\hat{c} + \epsilon)/b_\sigma]$ there will be a zero accumulated value loss program from $x(0)$.

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

Remark 5.1 We leave it to the reader to check that Condition 2 can be equivalently stated as

$$\{S_w - \{\hat{c}\}\} \cap \{\{\hat{c}\} - S_w\} = \{0\}\,.$$

To sum up, in Sections 4 and 5 we give a complete characterization of V under Condition 1 or Assumption 1, leaving the study of the general case, i.e., $\{\sigma\} \subsetneq I$ and $\{\hat{c}\} \subsetneq S_w$, for future work.

6 Maximal and Optimal Programs

Next, we turn to the relations between optimal, maximal and minimal value-loss programs that are obtained when Assumption 1 holds. We prove that every optimal program is a minimal value-loss program, a fact that seems to have been missed out in the literature. We also prove that every minimal value loss program is maximal, a benchmark result in the literature concerning non-decreasing, differentiable concave felicities. We present a new proof valid in our more general framework.

Finally, when $\xi_\sigma \neq 1$, or the Symmetry Condition 2 holds, we present an equivalence theorem that shows the equivalence of optimal, maximal and minimal value-loss programs from which the existence of optimal programs for every initial state can be deduced.

We begin with the following result,

Lemma 6.1 *If Assumption* 1 *holds, the Cesàro means of every good program converges to the golden-rule stock,* $\bar{x}(t) \to \hat{x}$ *when* $t \to \infty$.

Proof This lemma is easily deduced from Theorem 3.1 and Lemma 4.2. □

As it was shown by Brock [3, Theorem 1], the convergence of the Cesàro means of any good program to the GRS is fundamental to the proof of the following proposition

Proposition 6.1 *Under Assumption* 1*, if* $\{x(t)\}$ *is a value loss minimizer from* x_0 *then it is maximal.*

We present now another classical result that will be needed in the proof of Proposition 6.2.

Lemma 6.2 *Any maximal program is good.*

Proof Let $\{x(t)\}$ be a maximal program from x_0 and $\{x'(t)\}$ be a good program from the same initial condition. To reach a contradiction suppose that $\{x(t)\}$ is bad. We have that

$$\begin{aligned}\sum_{t=0}^{T} w\left(b\,y'(t)\right) - w\left(b\,y(t)\right) &= \sum_{t=0}^{T} w\left(b\,y'(t)\right) - w(\hat{c}) + \sum_{t=0}^{T} w(\hat{c}) - w\left(b\,y(t)\right)\\ &\geq M - \sum_{t=0}^{T} w\left(b\,y(t)\right) - w(\hat{c})\end{aligned}$$

$$\Rightarrow \liminf_{T} \sum_{t=0}^{T} w\left(b\,y'(t)\right) - w\left(b\,y(t)\right) \geq M + \infty$$

contradicting the fact that $\{x(t)\}$ is maximal. □

The convergence of the Cesàro means proves itself also basic to prove that every optimal program is a minimal value-loss program, as we see in the following proposition.

Proposition 6.2 *Under Assumption* 1*, if* $\{x(t)\}$ *is an optimal program from* x_0*, then it is a minimal value-loss program.*

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 $\epsilon > 0$ such that[7]

$$\sum_{t=0}^{\infty} \delta(t) - \sum_{t=0}^{\infty} \delta'(t) > \epsilon.$$

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

$$\sum_{t=0}^{T-1} \delta(t) - \sum_{t=0}^{T-1} \delta'(t) > \epsilon_0 \ \text{ for all } T \geq T_0.$$

Using the very easy to check equality

$$\sum_{t=0}^{T-1} \left[w\left(b\,y'(t)\right) - w\left(b\,y(t)\right)\right] = \sum_{t=0}^{T-1} \delta(t) - \sum_{t=0}^{T-1} \delta'(t) + \hat{p}\ \left(x(T) - x'(T)\right), \quad (17)$$

[7] In the sequel, and when no confusion can arise, we will write simply $\delta(t)$ and $\delta'(t)$ instead of $\delta\left(x(t), x(t+1)\right)$ and $\delta\left(x'(t), x'(t+1)\right)$.

we deduce

$$\sum_{t=0}^{T-1}\left[w\left(b\,y'(t)\right)-w(b\,y(t))\right] > \epsilon_0+\hat{p}\ \left(x(T)-x'(T)\right) \qquad \text{for all } T\geq T_0.$$

We know as well that there is T_1 such that

$$\limsup_T \sum_{t=0}^{T-1}\left[w\left(b\,y'(t)\right)-w\left(b\,y(t)\right)\right]+\epsilon_0/2 > \sum_{t=0}^{T-1}\left[w\left(b\,y'(t)\right)-w(b\,y(t))\right]$$
$$\text{for all } T\geq T_1.$$

Using the last two inequalities and the optimality of $\{x(t)\}$ we get

$$\epsilon_0/2 > \epsilon_0+\hat{p}\ \left(x(T)-x'(T)\right) \qquad \text{for all } T\geq \max\{T_0, T_1\}$$

which leads to the following $-\epsilon_0/2\ \geq\ \lim_T\ \hat{p}\ (\bar{x}(T)-\bar{x}'(T))$.

Now observe that both programs are good. Indeed, it is trivial to see that any optimal program is maximal and, due to Lemma 6.2, must be good. In addition, $x'(t)$ is the value loss minimizer and is also good. Hence, the Cesàro means of both programs converge to $\hat{x}$, contradicting the previous inequality. □

As shown in [34, Theorem 2.4], it is the asymptotic convergence of the stocks of any good program to the GRS that turns out to be a key ingredient in the demonstration that any minimal value-loss program is optimal, and hence for the proof of the existence of such programs.

Theorem 6.1 *Under Assumption* 1*, if* $\xi_\sigma\neq 1$ *or Condition* 2 *holds, the following are equivalent:*

(i) $\{x(t)\}$ *is an optimal program from* x_0.
(ii) $\{x(t)\}$ *is a minimum value-loss program from* x_0.
(iii) $\{x(t)\}$ *is a maximal program from* x_0.

Proof Under Assumption 1, we know that (i)⇒(ii) and (ii)⇒(iii) due to Propositions 6.2 and 6.1, respectively. To prove the equivalence, it is only left to see that (iii)⇒(i).

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, we know that

$$\lim_T \sum_{t=0}^{T} w\left(b\,y'(t)\right)-w(\hat{c})=-\infty \quad\text{and}\quad \sum_{t=0}^{T} w(b\,y(t))-w(\hat{c})\geq M \ \text{ for all } T$$

then we can easily deduce that $\lim_T\sum_{t=0}^{T}w(b\,y'(t))-w(b\,y(t))=-\infty$.

If the alternative program is good then we know that $\lim_T\sum_{t=0}^{T}\delta'(t)$ is well defined as well as $\lim_T\sum_{t=0}^{T}\delta(t)$. We also know that $\lim_T x(T)=\lim_T x'(T)=\hat{x}$, thanks to Theorem 5.1.

Considering (17) and letting $T\to\infty$ we get that the limit of the right hand side is defined and so it is the limit of the left hand side, hence:

$$\limsup_T \sum_{t=0}^{T-1} w\left(b\,c'(t)\right)-w\left(b\,c(t)\right)=\liminf_T \sum_{t=0}^{T-1} w\left(b\,c'(t)\right)-w\left(b\,c(t)\right)\leq 0.$$

□

Corollary 6.1 *Thanks to Proposition* 4.4 *and Theorem* 6.1*, the existence of optimal programs follows.*

We conclude this section with an observation that Theorem 6.1 is tight in the sense that if $\xi_\sigma = 1$ and Condition 2 does not hold, we can show that there is no optimal program from *any* initial state. This proof follows the lines of [13, Theorem 2.4] where the non-existence of an optimal program was showed for a 2-sector, single machine-type case of the RSS model when the felicity function is linear and the marginal rate equals unity.

7 Conclusion

We begin this conclusion by reminding the reader that the pioneering papers of Solow, Robinson and Stiglitz were in continuous time, and the discrete-time revisiting of the model by Khan-Mitra attracted attention because of the dramatic dissonance between the results in the two settings. Indeed, as the three decisive examples in [8] illustrate, the characterization of optimal policies that was obtained through the use of Pontryagin's maximum principle, no longer holds in discrete-time setting. In particular, Examples 2 and 3 in [8] established that for the discrete-time model optimal policies in [29] turn out to be *bad* policies, and the fundamental conclusion for development planning that only *one* type of machine is produced, no longer holds. More to the point, Stiglitz worked with linear benefit functions, and the model proved recalcitrant to analytical treatment for strictly concave benefit functions.[8] There have of course been advances over the years in the theory of continuous-time optimal control that are motivated by concerns identical to ours, as in [20, 21], though not yet given textbook and survey treatment (see, for example [4, 27, 28]). The genesis of this work dates to the mid-eighties where in his emphasis on a *minimal long-run average-cost growth rate,* Leizarowitz has focussed on cost functions that are lower semicontinuous and convex, as in [18, 19], or simply lower semicontinuous and bounded as in [17]. Indeed, one can trace Zaslavski's studies of the RSS model as the importation of techniques from this literature. However, in emphasizing supportability of a non-smooth, non-monotonic felicity function at a distinguished point, and then focussing primarily on the support in developing the value-loss or penalty function, we depart from the optimal control literature as much as we do from that in mathematical economics.[9] The extent to which our techniques

[8]In [5, 30], remaining in continuous time, Cass-Stiglitz and Stiglitz worked with linear benefit functions with a minimum consumption constraint, but could not provide an explicit characterization of optimal policies.

[9]This strong claim needs justification for the non-initiated reader, and we offer it in some detail. Our remarks all refer to the text [4]. Although, the integrands proposed in Section 4.9 are not strictly concave, but concave, our assumption on the benefit function is much weaker: the supportability condition that we offer can be interpreted as concavity at a point, weaker than concavity itself. Furthermore, the existence result in the book by Carlson et al. requires Assumptions 4.7, 4.8 and Property S. Assumption 4.7 is concavity and upper semi-continuity, Assumption 4.8 is the analogue of the Brock-Koopmans-Mitra uniqueness condition and Property S is the analogue of the requirement of convergence in the Von Neuman Facet, which does not always hold in our model. In Chapter 8, the analogue to Property S, now called Assumption 8.4, is proposed and the authors themselves say that it is very difficult to verify in terms of the primitives of the model, (something that can be done easily for the Symmetry Condition that we present in this paper) substituting it for the stronger (A4′) that also implies that the set $\mathcal{F}$ is the singleton golden rule stock in the non-delay case, see [4, pp. 208–209]. In Chapter 9, a convergence result is offered as Theorem 9.6 using what the authors

and concepts can be subsumed in an overall complementary and synthetic approach must remain a subject for future investigation.[10]

Acknowledgements The authors thank Tapan Mitra for insightful correspondence and Rabee Tourky for stimulating conversation. This version of the paper was completed when Khan was visiting the *China Economics and Management Academy (CEMA)* at the Central University of Finance and Economics in Beijing, China during August 2010. He would like to thank CEMA for its facilities and hospitality. Adriana Piazza gratefully acknowledges the financial support of Programa Basal PFB 03, CMM, U. de Chile and that of FONDECYT under Project 11090254.

References

1. Arkin, V., Evstigneev, I.: Stochastic Models of Control and Economic Dynamics. Academic, New York (1987)
2. Bewley, T. F.: General Equilibrium, Overlapping Generations Models and Optimal Growth Theory. Harvard University Press, Cambridge (2007)
3. Brock, W.A.: On existence of weakly maximal programmes in a multi-sector economy. Rev. Econ. Stud. **37**, 275–280 (1970)
4. Carlson, D.A., Haurie, A. B., Leizarowitz, A.: Infinite Horizon Optimal Control: Deterministic and Stochastic Systems. Springer, Berlin (1991)
5. Cass, D., Stiglitz, J.E.: The implications of alternative savings and expectation hypotheses for choices of technique and patterns of growth. J. Polit. Econ. **77**, 586–627 (1970)
6. Dana, R.A., Le Van, C., Mitra, T., Nishimura, K. (eds.): Handbook of Optimal Growth, vol. 1 (2006) Springer, Berlin.
7. Gale, D.: On optimal development in a multi-sector economy. Rev. Econ. Stud. **37**, 1–18 (1967)
8. Khan, M.A., Mitra, T.: On choice of technique in the Robinson–Solow–Srinivasan model. Int. J. Econ. Theory **1**, 83–109 (2005)
9. Khan, M.A., Mitra, T.: Undiscounted optimal growth under irreversible investment: a synthesis of the value-loss approach and dynamic programming. Econ. Theory **29**, 341–362 (2006)
10. Khan, M.A., Mitra, T.: Optimal growth in a two-sector RSS model without discounting: a geometric investigation. Japanese Economic Review **58**, 191–225 (2007)
11. Khan, M.A., Mitra, T.: Growth in the Robinson–Solow–Srinivasan model: undiscounted optimal policy with a strictly concave welfare function. J. Math. Econ. **44**, 707–732 (2008)
12. Khan, M.A., Piazza, A.: On the Mitra–Wan forestry model: a unified analysis. In: Publicación Técnica CMM, no. 226, pp. 1–21 (2009)
13. Khan, M.A., Piazza, A.: On the non-existence of optimal programs in the Robinson–Solow–Srinivasan (RSS) model. Econ. Lett. **109**, 94–98 (2010)
14. Khan, M.A., Piazza, A.: An overview of turnpike theory: towards the discounted deterministic case. Adv. Math. Econ. **14**, 39–68 (2011)
15. Khan, M.A., Zaslavski, A.J.: On a uniform turnpike of the third kind in the Robinson–Solow–Srinivasan model. J. Econ. **92**, 137–166 (2007)
16. Khan, M.A., Zaslavski, A.J.: On two classical turnpike results for the Robinson–Solow–Srinivasan (RSS) model. Adv. Math. Econ. **13**, 47–97 (2010)
17. Leizarowitz, A.: Infinite horizon autonomous systems with unbounded cost. Appl. Math. Optim. **13**, 19–43 (1985)
18. Leizarowitz, A.: Infinite horizon autonomous systems with unbounded cost. Math. Oper. Res. **10**, 450–461 (1985)
19. Leizarowitz, A.: Optimal trajectories of infinite-horizon deterministic control systems. Appl. Math. Optim. **19**, 11–32 (1989)
20. Leizarowitz, A., Zaslavski, A.J.: Infinite-horizon variational problems with non-convex integrands. SIAM J. Control Optim. **34**, 1099–1134 (1996)

refer to as Property C, the analogue of Property S (see [4, p. 242]), but after that, in Remark 9.2 (p. 247) the authors say that "at present there are no known conditions under which property C holds".

[10]For similar judgements, see Rockafellar's recent survey, [27].

21. Leizarowitz, A., Zaslavski, A.J.: On a class of infinite-horizon optimal control problems with periodic cost functions. J. Nonlinear Convex Anal. **6**, 71–91 (2005)
22. Majumdar, M.K.: Intertemporal allocation with a non-convex technology. In: Dana, R.A., Le Van, C., Mitra, T., Nishimura, K. (eds.) Handbook of Optimal Growth, vol. 1, chapter 7, pp. 171–202. Springer, Berlin (2006)
23. McKenzie, L.W.: Optimal economic growth, turnpike theorems and comparative dynamics. In: Arrow, K.J., Intrilligator, M. (eds.) Handbook of Mathematical Economics, vol. III, pp. 1281–1355. North-Holland, Amsterdam (1986)
24. McKenzie, L.W.: Classical General Equilibrium Theory. MIT Press, Cambridge (2002)
25. Mitra, T.: Characterization of the turnpike property of optimal paths in the aggregative model of intertemporal allocation. Int. J. Econ. Theory **1**, 247–275 (2005)
26. Mitra, T., Wan, H.W. Jr.: On the Faustmann solution to the forest management problem. J. Econ. Theory **40**, 229–249 (1986)
27. Rockafellar, R.T.: Hamiltonian trajectories and saddle points in mathematical economics. Control Cybern. **4**, 1565–1578 (2009)
28. Seierstad, A., Sydsaeter, K.: Optimal Control Theory With Economic Applications. North-Holland, Amsterdam (1987)
29. Stiglitz, J.E.: A note on technical choice under full employment in a socialist economy. Econ. J. **78**, 603–609 (1968)
30. Stiglitz, J.E.: Recurrence of techniques in a dynamic economy. In: Mirrlees, J., Stern, N.H. (eds.) Models of Economic Growth. Wiley, New York (1973)
31. Zaslavski, A.J.: Optimal programs in the RSS model. Int. J. Econ. Theory **1**, 151–165 (2005)
32. Zaslavski, A.J.: Turnpike Properties in the Calculus of Variations and Optimal Control. Springer, New York (2005)
33. Zaslavski, A.J.: Good programs in the RSS model with a nonconcave utility function. J. Ind. Manag. Optim. **2**, 399–423 (2006)
34. Zaslavski, A.J.: Turnpike results for discrete-time optimal control systems arising in economic dynamics. Nonlinear Anal. **67**, 2024–2049 (2007)