On choice of technique in the Robinson–Solow–Srinivasan model

M. Ali Khan* and Tapan Mitra†

We report results on the optimal "choice of technique" in a model originally formulated by Robinson, Solow and Srinivasan. By viewing this model as a specific instance of the general theory of intertemporal resource allocation associated with Brock, Gale and McKenzie, we resolve long-standing conjectures in the form of theorems on the existence and price-support of optimal paths, and on their long-run behavior. We also examine policies, due to Stiglitz, as a cornerstone for a theory of transition dynamics in this model. We present examples to show that: (i) an optimal program can be periodic; (ii) a Stiglitz' program can be bad; and (iii) a Stiglitz production program can be non-optimal. We then provide sufficient conditions under which the policies proposed by Stiglitz coincide with optimal behavior.

Key words choice of technique, long run, transition dynamics, Stiglitz program, cycling

JEL classification D90, C62, O21

Accepted 29 July 2004

1 Introduction

In the late 1960s and early 1970s, under the general heading of "technical choice under full employment in a socialist economy," Robinson (1960, 1969), Okishio (1966) and Stiglitz (1968, 1970, 1973) studied the problem of optimal economic growth in a model of an economy originally formulated by Robinson (1960), Solow (1962b) and Srinivasan (1962a) (henceforth, the RSS model).[1] The work generated controversy. Stiglitz argued,

* Department of Economics, The Johns Hopkins University, Baltimore, Maryland, USA. Email: akhan@jhu.edu.

† Department of Economics, Cornell University, Ithaca, New York, USA.

The work reported here is part of a project with a long gestation period: it was initiated during Mitra's visit to the Department of Economics at the University of Illinois in 1986, received invaluable impetus from Professor Robert Solow's presentation at the Srinivasan Conference held at Yale in March 1998, was continued when Khan visited the Department of Economics at Cornell in November 1998, October 2000, July 2003 and July 2004, and the EPGE, Fundação Getilio Vargas in January 2001 and December 2002. The authors are grateful to all of these institutions for their hospitality, as well as to the Center for Analytic Economics at Cornell and the Center for a Livable Future at Johns Hopkins for research support. Two anonymous referees of *IJET* and Mr Chris Metcalf gave the manuscript a careful reading. Khan also thanks Professors Abhijit Banerjee, Jimmy Chan, Avinash Dixit and Debraj Ray for stimulating conversation.

[1] In Khan (2000), the model is referred to as the Solow–Srinivasan model; also see Solow (1962b) and Khan (2000) for the way it is seen in earlier work.

with justification, that the Robinson–Okishio assumption of a fixed labor allocation between the consumption and investment sectors had no place in an exercise that sought to determine the optimal growth path and, thereby, an optimum time-path of the allocation of labor.[2] He identified development policies, henceforth Stiglitz' policies,[3] under which there is investment only in the type of machine σ that minimizes effective labor costs and simultaneously maximizes the steady-state consumption, and a utilization of only those types of machines whose output per man ratios are higher than the effective labor cost of producing σ. Stiglitz observed that the "number of workers working in the consumption-goods sector increases monotonically (capital 'widening' occurs in a smooth way), output of consumption goods need not be monotonically increasing,"[4] and prescribed for the economy at any point in time an optimal choice of techniques, both to use and to produce, and, thereby, the (instantaneous) optimal levels of technological obsolescence: prescriptions that are all independent of the felicity function. Robinson commented on Stiglitz' solution by criticizing his assumption of a fixed positive discount rate, continuous time and the linearity assumption in the specification of the planner's felicity function.[5]

Robinson's objections were explicitly acknowledged by Stiglitz and,[6] as a first approximate step,[7] he extended his earlier analysis to the case of a minimum consumption constraint in a setting with continuous time and a positive rate of discount. However, he emphasized that the important modifications concerned transition, rather than long-run, dynamics.[8]

> Even if there is a minimum consumption constraint and a finite gestation period, the path of development will, after an initial "adjustment" period, look exactly as I have described it. [Unlike] long-run neoclassical models with malleable capital [where] the optimal policy is always of the so-called bang-bang variety – if the initial capital labor ratio is less than its long-run equilibrium value there is always a period of zero consumption, after which consumption jumps to its long-run equilibrium value, whereas in our *ex-post* fixed coefficients model consumption increases steadily to its long-run value.

Given the primary interest in the undiscounted case, Stiglitz interpreted the undiscounted case as a situation when the discount rate is "negligible"; he developed the intuitive ideas in discrete time and then chose to translate them to the continuous-time framework.[9]

---

[2] See (1968, 1970). Stiglitz (1970, p. 421) writes. "There may be some special situations . . . where the employment allocation is the same for all steady-state paths, but even then, in going from one steady-state path to another, one cannot infer that the employment allocation is unchanged – and it is this dynamic problem that we are discussing."

[3] This is formalized in Definitions 8 and 9 below. As we shall see in the sequel, Stiglitz' policies can be usefully compared to Faustmann's solution to the forestry problem, as formalized in Mitra and Wan (1986).

[4] See Stiglitz (1973, pp. 143–4). In the discussion of his policy, Stiglitz also drew attention to preliminary investigations of Bruno (1967).

[5] We restate Robinson's criticisms in our own terminology; she phrases them in terms of a "discount rate chosen once and for all, . . . negligible gestation periods, . . . [and] ceasing to consume and living on air during the first phase of the plan."

[6] See Stiglitz (1973) and also Cass and Stiglitz (1969).

[7] Therefore, Cass and Stiglitz (1969) saw the instantaneous utility function $U(C) = -\infty$ for $C < \bar{C}$ and $U(C) = C$ for $\geq \bar{C}$, $\bar{C}$ a minimum consumption level, as "one approximation to the general instantaneous utility function satisfying $U'(C) > 0$ with $\lim_{C=0} U'(C) = \infty$ and $U''(C) \leq 0$."

[8] Stiglitz's (1970) response is important for the record; this essay can also be seen as a further investigation into the substance of this response.

[9] See footnotes 1 and 3 on page 608 and the discussion of the "correct" pricing system on page 606 in Stiglitz (1968).

In his recent revisit of Srinivasan (1962a), Solow (2000, p. 7) asks for a solution to the "Ramsey problem for this model." Because Stiglitz had already provided a solution with a "linear utility and positive time preference," the open questions concern a rigorous treatment of the undiscounted case and of the discounted case with a "strictly concave social utility function for current per capita consumption". Like Robinson, Solow also mentions that an "adoption of this [linear utility] criterion can indeed lead to unjustifiable neglect of early consumption," and if one was to share "Ramsey's belief that the only ethically defensible social rate of time preference is zero, a sufficiently sharply-concave utility function would enforce a closer approach to intergenerational equality." In short, Solow's question remains unanswered, and the generalization of Stiglitz's work in the directions it prompts remains yet to be accomplished.[10] In the present paper, we address this general question and, most importantly from a methodological point of view, do so in the setting of the modern theory of optimal intertemporal allocation initiated originally by Ramsey (1928) and von Neumann (1935–1936) and brought to completion at the hands of Brock, Gale and McKenzie.[11] Because this theory was being finalized at the same time as the "capital controversy" between the two Cambridges,[12] it has not been brought to bear on the fundamental issues.

In terms of specifics, we truly treat the Ramsey problem; that is, consider a formulation in which there is no discounting of future utilities and, therefore, no appeal to the assumption of structural stability of the model at the zero discount rate, an assumption at best roundabout and at worst dubious. We are by this time very familiar with the overtaking criterion of Atsumi (1965) and von Weiszäcker (1965) and, under this criterion, an optimal path in the undiscounted case can be shown to exist and its properties can be rigorously studied. Our treatment of time is discrete: the general theory of intertemporal allocation is developed in the simplest and most elegant way in such a setting (see McKenzie (1986) for a masterly presentation), and we can work with a reduced form of the Robinson–Solow–Srinivasan (RSS) model [13] in which the technological possibilities are given by a transition possibility set, and the objective function by a (reduced-form) utility function defined on this set (that is, defined on beginning and end of period capital stock vectors). We establish the existence of a golden-rule stock, with support prices, and show that the golden-rule stock is unique. We appeal to the methods of Brock (1970) and McKenzie (1968) to show the existence of an optimal program and, furthermore, to establish that, starting from an arbitrary initial stock, it converges asymptotically to a subset of the transition set, the so-called von Neumann facet, consisting of all plans which have "zero value-loss" at the golden-rule support prices. In the case of a strictly concave felicity function, the von Neumann facet shrinks to a point, and so we have asymptotic convergence to the golden-rule

[10] For some partial attempts at solution, and for a numerical example, see Stiglitz (1973); also see Stiglitz (1968) and Cass and Stiglitz (1969). However, in Khan (2000, p. 15) the situation is expressed as: "The loose end remains loose."

[11] The relevant papers are Gale (1967), McKenzie (1968) and Brock (1970). In the sequel, when we refer to the "general theory of optimal intertemporal allocation," we shall have these papers in mind.

[12] It is, of course, not our intention to revisit this debate here: the interested reader might want to see Birner (2002) and his references.

[13] We choose to work with the version presented in Stiglitz (1968) rather than that in Solow (1962b) or Srinivasan (1962a). All of these variants can be viewed as special cases of the models considered in Bruno (1967) or the more general treatment in Koopmans (1971) and Koopmans and Hansen (1972). We leave the analysis of these papers as a task for future research; also see the third paragraph of the concluding Section 8 below.

stock. These results furnish a complete resolution of the problem of the long-run choice of technique and, thereby, illustrate the power and elegance of the general theory.

Because we have a complete resolution of the problem of the long-run choice of technique, the natural question arises as to the choice of technique in *transition* to the steady state: a determination of the type and amounts of machines that are produced and used in the short run. Unfortunately, it is on this hard problem of transition dynamics that the general theory has little to offer, with the published literature lacking concrete results of any generality.[14] However, Stiglitz' prescriptions as to the choice of techniques can be identified as a basis for the development of a full-scale theory of transition dynamics: an analytical marker at which one can aim. It is here that the results yield surprises: even questions once seen as resolved are now starkly revealed not to be so through simple and compelling counterexamples.

Moving on to positive results, for an economy with a linear felicity function, we offer a (novel) set of sufficient conditions, pertaining only to the parameters of the type of machine σ used in the long run, under which the Stiglitz program is optimal, and uniquely so. For economies with a general felicity function and, *a fortiori,* for an economy with a linear utility function and a minimum consumption constraint, we also present sufficient conditions for the optimality of a Stiglitz production program. These conditions, in pointing to an interesting distinction between choice of technique that is appropriate in the short run from that which is appropriate in the long run, also connects to the published literature of the 1960s on planning in India (and elsewhere)[15] that comes as close to stating the problem as precisely as can be expected in the pre-Pontryagin period.[16] However, this published literature and earlier work notwithstanding, it will generally be recognized today that whether we are interested in this issue from a planning perspective or from the modern perspective of a representative agent, the problem of an appropriate choice of technique should really be viewed as part of the general theory of economic growth. A subsidiary motivation of the present paper is to facilitate this reorientation.[17]

It is important to appreciate the methodological significance of this reformulation of the RSS model. In the standard treatment based on Pontryagin's principle,[18] as in the work of Stiglitz (1968), Sen (1968) and others, one appeals to the transversality conditions in the study of the differential equations pertaining to the state and auxiliary variables obtained by substituting the values of the controls that maximize the instantaneous Hamiltonian.[19] Therefore, the relevance of the rest points is established only towards the end of the analysis. Here, we begin with the rest points, the golden-rule stock and the golden-rule prices, and use the value-loss function and the so-called average turnpike property of good programs

---

[14] See the third paragraph of the concluding remarks in Section 8 below.

[15] In addition to Raj and Sen (1961) and Sen (1960) in particular, also see Dobb (1956, 1960, 1961, 1967), Halevi (1987), Mirrlees (1962), Naqvi (1963), Solow (1962a); and for an open-economy perspective, Bardhan (1971).

[16] See Raj and Sen (1961). We remind the reader that an earlier version of the Raj and Sen paper (with a different title) was published in *Arthaniti* in 1959, and that Naqvi (1963) is a follow up to the Raj and Sen paper as his leading footnote clearly indicates.

[17] For an early emphasis on this, see Mirrlees (1962) and Srinivasan (1962b); Okishio (1987) represents the alternative perspective. This point was independently underscored to Khan by Debraj Ray as a comment on Khan (2000).

[18] Thus Dixit (1990, p. 5) writes, "Nowadays Hamiltonians and phase diagrams are everyday stuff for the typical second-year graduate student," and quotes Frank Hahn's reference to the "unseemly haste to get down to the Hamiltonian."

[19] In the context of Stiglitz (1968), see his footnote 2 both on page 605 and on page 607.

to yield the optimal program.[20] Stiglitz investigates the convergence (turnpike property) of a path that follows his (derived) policy prescriptions as to the choice of techniques,[21] while we need to investigate whether the optimal path and its turnpike property is sustained by these prescriptions. As mentioned above, this cannot be established, in general, for either a linear or a strictly concave felicity function, but only in special (identified) cases of either formulation. Therefore, through the introduction of a new conceptual vocabulary, a difficult step in one perspective is rendered straightforward in another.[22]

In sum, our results exhibit in a dramatic way both the strength and the weakness of the general theory of intertemporal allocation alluded to earlier and, thereby, reveal exactly why the choice of an appropriate technique is such a difficult and multifaceted problem. The application has the advantage of illustrating the power and flexibility of the modern theory: it is ideally suited to deal with this problem, and the general results of this theory can be readily applied through the use of extremely elementary methods. As such it is perhaps overdue. However, a secondary benefit of this application concerns the theory itself; it offers insights into its scope and suggests directions along which it may find fruitful extension.[23] It also points clearly to issues to which the general theory has (and by its very nature, will have) very little to offer, thereby indicating that even after much theoretical progress has been made, some of the questions that were asked 50 years ago about the appropriate choice of technique remain hard unanswered problems that need to be approached case by case.

The remainder of the paper is organized as follows. Sections 2, 3 and 4 present the basic theory of the RSS model when it is converted to its Gale–McKenzie reduced form. In particular, under a standing hypothesis on the finite set of parameters that define the RSS model, we show the existence and uniqueness of the golden-rule stock, and the existence of a program that is optimal starting from any given initial stock of machines. With this standard theory (Theorems 1 to 2) as an (indispensable) background, we can turn to the central results of the paper. In Section 5, we consider the question of the correct choice of technique for the long run, and through the identification of the von Neumann facet, present results for both linear and strictly concave felicity functions. In Section 6, we turn to transition dynamics through the identification and formalization of the policy prescriptions due to Stiglitz (1968), and present examples that decisively refute plausible intuitions concerning these prescriptions. In Section 7, we present a sufficient parameterization under which a Stiglitz program and/or a Stiglitz production program is optimal. Section 8 lists the salient results and identifies problems that remain open. The technical and computational details of some of the proofs are collected in an appendix.[24]

[20] As the reader will see below, we appeal to McKenzie's (1986) price-support property only to establish our final result pertaining to transition dynamics in the case of a strictly concave felicity function (Theorem 7 below).

[21] As alluded to in Footnote 3, we see the Stiglitz policy as the analogue of the Faustmann solution in the economics of forestry, and it would be interesting to pursue the analytics of this analogy; see Section 8.

[22] Stiglitz (1968) recognizes that the results for the linear utility function might not carry over to the general concave case, and in a subsequent analysis of the problem, one with a minimum consumption constraint, he refers to the difficulty of showing that there is only one type of machine that maximizes $p(x, t)$; see Stiglitz (1973). He also refers in this connection to Bliss (1968) and Cass and Stiglitz (1969). For this scepticism concerning the linear case, also see Solow (2000) in addition to Robinson (1969).

[23] Therefore, a distinction has to be drawn between applying a theorem from applying its methods of proof. As the reader will see in the sequel, the hypotheses of Brock's existence theorem and of McKenzie's price-support property are not literally fulfilled by the RSS model, but their methods of proof do apply.

[24] The full details of the proofs and computations are collected in *CAE Working Paper* No. 04-13 and, as indicated in the sequel, available from either author on request.

2 The model and its reduced form

We begin with some preliminary notation. Let $\mathbb{N}(\mathbb{N}_+)$ be the set of non-negative (positive) integers, $\mathbb{R}(\mathbb{R}_+)$ the set of real (non-negative) numbers. We shall work in finite-dimensional Euclidean space $\mathbb{R}^n$ with non-negative orthant $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_i \geq 0, i = 1, \ldots, n\}$. For any x, y in $\mathbb{R}^n$, let the inner product $xy = \sum_{i=1}^n x_i y_i$, and $x \gg y$, $x > y$, $x \geq y$ have their usual meaning. Let $e(i)$, $i = 1, \ldots, n$, be the i^{th} unit vector in $\mathbb{R}^n$, and e be an element of $\mathbb{R}^n_+$ all of whose coordinates are unity. For any $x \in \mathbb{R}^n$, let $\|x\|$ denote the Euclidean norm of x. The empty set is denoted by $\emptyset$ and set-theoretic subtraction between A and B by A/B.

Our choice of $\mathbb{R}^n$ is dictated by the consideration of an economy capable of producing a finite number n of alternative types of machines. For every $i = 1, \ldots, n$, one unit of machine of type i requires $a_i > 0$ units of labor to construct it, and together with one unit of labor, each unit of it can produce $b_i > 0$ units of a single consumption good. Therefore, the production possibilities of the economy can be represented by an (labor) input-coefficients vector, $a = (a_1, \ldots, a_n) \gg 0$ and an output-coefficients vector, $b = (b_1, \ldots, b_n) \gg 0$. Without loss of generality we shall assume that the types of machines are numbered such that $b_1 \geq b_2 \cdots \geq b_n$.[25]

We shall assume that all machines depreciate at a rate $d \in (0, 1)$. Therefore, the effective labor cost of producing a unit of output on a machine of type i is given by $(1 + da_i)/b_i$: the direct labor cost of producing unit output, and the indirect cost of replacing the depreciation of the machine in this production.[26] We shall work with the reciprocal of the effective labor cost, the effective output that takes the depreciation into account, and denote it by c_i for the machine of type i.[27] Throughout the present paper, we shall assume that there is a unique machine type σ at which this effective labor cost $(1 + da_i)/b_i$ is minimized, or at which the effective output per man $b_i/(1 + da_i)$ is maximized. Therefore, we shall assume:

$$\text{There exists } \sigma \in \{1, \ldots, n\} \text{ such that for all } \; i = 1, \ldots, n, i \neq \sigma, c_\sigma > c_i. \qquad (1)$$

For each date $t \in \mathbb{N}$, let $x(t) = (x_1(t), \ldots, x_n(t)) \geq 0$ denote the amounts of the n types of machines that are available in time-period t, and let $z(t+1) = (z_1(t+1), \ldots, z_n(t+1)) \geq 0$ be the gross investments in the n types of machines during period $(t+1)$. Hence, $z(t+1) = (x(t+1) - x(t)) + dx(t)$, the sum of net investment and depreciation. Let $y(t) = (y_1(t), \ldots, y_n(t))$ be the amounts of the n types of machines used for production of the consumption good, $by(t)$, during period $(t+1)$.[28] Let the total labor force of the economy be stationary and positive. We shall normalize it to be unity. Clearly, gross investment, $z(t+1)$ representing the production of new machines of the various types, will require $az(t+1)$ units of labor in period t. Also, $y(t)$ representing the use of available machines for manufacture of the consumption good, will require $ey(t)$ units of labor in period t. Therefore, the availability of labor constrains employment in the consumption and investment sectors by $az(t+1) + ey(t) \leq 1$. Note that both the flow of

[25] Note that Stiglitz (1968) assumes that $b_i > b_j$ implies that $a_i > a_j$; whereas this is a natural hypothesis, we make no such assumption.

[26] See Stiglitz (1968, pp. 608–9) on a "labor theory of value" interpretation.

[27] As we shall see below, c_i is the value of the steady-state consumption per man if only machines of type i are used and produced, a consideration that governs our choice of notation.

[28] The reader may choose to think of the consumption in period t as the scalar $c(t+1)$, with c_i reserved for $b_i/(1 + da_i)$; we avoid this notation in the text to prevent any ambiguity.

consumption and of investment (new machines) are in gestation during the period and available at the end of it. We now give a formal summary of this technological structure.

Definition 1 *A program from x_o in $\mathbb{R}^n_+$ is a sequence*[29] $\{x(t), y(t)\}$ *with* $(x(t), y(t)) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+$ *such that* $x(0) = x_o$, *and for all* $t \in \mathbb{N}$, $(i) x(t+1) \geq (1-d)x(t)$, $(ii)\, 0 \leq y(t) \leq x(t)$, $(iii)\, a(x(t+1) - (1-d)x(t)) + ey(t) \leq 1$. *A program* $\{x(t), y(t)\}$ *is simply a program from* $x(0)$.

Definition 2 *Associated with any program* $\{x(t), y(t)\}$ *is a gross investment sequence* $\{z(t+1)\}$ *with* $z(t+1) \in \mathbb{R}^n_+$, *and a consumption sequence* $\{by(t)\}$ *such that for all* $t \in \mathbb{N}$, $z(t+1) = x(t+1) - (1-d)x(t)$.

Definition 3 *A program* $\{x(t), y(t)\}$ *is stationary if for all* $t \in \mathbb{N}$, $(x(t), y(t)) = (x(t+1), y(t+1))$.

We conclude this subsection with a result on the boundedness property of programs.

Proposition 1 *For any program* $\{x(t), y(t)\}$, *there exists* $m(x(0)) \in \mathbb{R}_+$ *such that* $x(t) \leq m(x(0))e$ *for any* $t \in \mathbb{N}$.

PROOF: The case $t = 0$ is a triviality. For $t \in \mathbb{N}_+$, $ax(t) \leq 1 + (1-d)ax(t-1) \leq \sum_{\tau=0}^{t-1}(1-d)^\tau + (1-d)^t ax(0)$. Because $0 < d < 1$, we obtain $ax(t) \leq (1/d) + ax(0)$. Let $a_j = \min_{1 \leq i \leq n} a_i$. Because $a_i > 0$ for all $i = 1, 2, \ldots n$, we obtain $x_i(t) \leq (1/a_j)((1/d) + ax(0)) \equiv m(x(0))$ and complete the proof. □

The preferences of the planner are represented by a felicity function, $w : \mathbb{R}_+ \longrightarrow \mathbb{R}$, which is assumed to be continuous, strictly increasing and concave, and differentiable.[30] We suppose, as in the published literature taking its lead from Ramsey (1928), that future welfare levels are treated like current ones in the planner's objective function. The notion of optimality that we use is due to Brock (1970), and the notion of overtaking is due to Atsumi (1965) and von Weiszäcker (1965).[31]

Definition 4 *A program* $\{x^*(t), y^*(t)\}$ *from* x_o *is optimal if for every program* $\{x(t), y(t)\}$ *from* x_o,

$$\liminf_{T\to\infty} \sum_{t=1}^{T} [w(by(t)) - w(by^*(t))] \leq 0.$$

A program is a stationary optimal program if it is stationary and optimal.

Note that the optimality notion can be restated to say that there does not exist any other program $\{x(t), y(t)\}$, $x(0) = x_o$, a number $\varepsilon > 0$ and a time period t_ε such that $\sum_{t=1}^{T}[w(by(t)) - w(by^*(t))] > \varepsilon$ for all $T \geq t_\varepsilon$. Therefore, an optimal program is one in comparison to which no other program from the same initial stock is eventually significantly better, for any given level of significance.

[29] Note $\{x(t), y(t)\}$ is an abbreviation of $\{x(t), y(t)\}_{t\in\mathbb{N}}$; we use it for notational simplicity.

[30] We leave it to the reader to check that differentiability of w is not needed, and derivatives of w can be replaced uniformly by (for example) the right-hand derivative of w. These exist because w is concave and the point of evaluation of the (right-hand) derivative is always positive.

[31] Brock (1970) uses the terminology of "weakly maximal" programs for what we call optimal programs. The notion of optimality in Atsumi and von Weiszäcker is stronger, and creates problems in proving existence of optimal programs in many reasonable models.

Following McKenzie (1968), we convert the above model into its reduced form, and as emphasized in the introduction, thereby exploit as far as possible the results of the general theory of intertemporal allocation for our particular case. Define the *transition possibility set* Ω as a collection of pairs (x, x'), such that it is possible to obtain the amounts of the n types of machines x' in the next period (tomorrow) from the amounts of the n types of machines x available in the current period (today). Formally,

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

For any $(x, x') \in \Omega$, one can consider the amounts y of the n types of machines available for the production of the consumption good. Formally, we have a correspondence $\Lambda : \Omega \longrightarrow \mathbb{R}^n_+$ given by

$$\Lambda(x, x') = \{y \in \mathbb{R}^n_+ : 0 \leq y \leq x \text{ and } ey \leq 1 - a(x' - (1-d)x)\}.$$

For any $(x, x') \in \Omega$, we shall denote the number of machines that are produced in the period $(x' - (1-d)x)$ by z. Note that $z \geq 0$. Finally, the reduced form utility function, $u : \Omega \longrightarrow \mathbb{R}_+$, is defined on Ω such that

$$u(x, x') = \max\{w(by) : y \in \Lambda(x, x')\}.$$

We leave it to the readers to check for themselves that our assumptions on w imply that the reduced form utility function, u, is upper semicontinuous[32] and concave on Ω, and that it is increasing in its first argument and decreasing in its second argument.

Given the description of the transition possibility set Ω, and of the reduced form utility function, u, it is clear that for any program $\{x(t), y(t)\}$ from x_o, $(x(t), x(t+1)) \in \Omega$ and $y(t) \in \Lambda(x(t), x(t+1))$ for all $t \in \mathbb{N}$. Also, for any optimal program $\{x^*(t), y^*(t)\}$ from x_o, $w(by^*(t)) = u(x^*(t), x^*(t+1))$ for all $t \in \mathbb{N}$, and for every program $\{x(t), y(t)\}$ from x_o,

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

In summary, the basic data of the model denoted by the triple $(w, (a_i, b_i)_{i=1}^n, d)$ summarizing the felicity function w, the technology $(a_i, b_i)_{i=1}^n$, and the depreciation rate d, is converted to the pair (u, Ω) summarizing the reduced-form utility function u and the transition possibility set Ω.

3 Existence and uniqueness of a golden-rule stock

A stationary optimal program is of special significance, and in this section we take the first step in establishing the existence of such a program. We show the existence and uniqueness of a golden-rule stock and, simultaneously, provide a price support property of such a stock.[33] We exploit the concrete structure of the RSS model to provide a purely constructive

[32] It is now well understood that continuity of w does not necessarily imply the continuity of u; see Dutta and Mitra (1989) for details.

[33] We show in Section 4 (see Theorem 2) that this golden-rule stock defines a stationary optimal program.

proof of our claims. This has the additional advantage that we can identify the shadow prices in terms of the basic data of the model.

We begin with a definition.

Definition 5 *A golden-rule stock is* $\hat{x} \in \mathbb{R}^n_+$ *such that* $(\hat{x}, \hat{x})$ *is a solution to the problem: maximize* $u(x, x')$ *subject to* (i) $x' \geq x$, (ii) $(x, x') \in \Omega$.

If we limit ourselves to a stationary program in which only a machine of type i is used and produced, the constraint of labor allows us to maintain the stock $(1/(1 + da_i)$ and obtain a stationary consumption stream in the amount $b_i/(1 + da_i) = c_i$.[34] Because we have assumed (in (1) above) that a machine of type σ is the one that uniquely minimizes effective labor costs, we see that it is also the type that uniquely maximizes the consumption per unit of labor.[35] Denote $\hat{y} = (1/(1 + da_\sigma))e(\sigma)$, and note that if we are in such a stationary state, $b\hat{y} = (b_\sigma/(1 + da_\sigma))$ and $w'(b\hat{y})$ is the marginal utility of output produced. Furthermore, because the labor cost of a machine of type i is a_i, and a unit of labor is worth $((1 + da_i)/b_i)^{-1}$ units of output, a machine is worth $a_i \times (b_i/(1 + da_i)$ in terms of output, and $w'(b\hat{y})(a_i \times (b_i/(1 + da_i)))$ in terms of utils. We can then identify a stationary price system ($\hat{q}$ in terms of the consumption good and $\hat{p}$ in terms of utils)[36] for the various types of machines as $\hat{q}_i = (a_i b_i/(1 + da_i))$ and $\hat{p}_i = w'(b\hat{y})\hat{q}_i$ for each $i = 1, \ldots, n$.

We can now present a simple but important result.

Lemma 1 $w(b\hat{y}) \geq w(by) + \hat{p}x' - \hat{p}x$ *for any* $(x, x') \in \Omega$, *and for any* $y \in \Lambda(x, x')$.

PROOF: For any $(x, x') \in \Omega$ and $y \in \Lambda(x, x')$, we have:[37]

$$\begin{aligned}
b\hat{y} - by - \hat{q}(x' - x) &= c_\sigma - by - \hat{q}(x' - x) \\
&= c_\sigma - by - \hat{q}(x' - (1 - d)x) + d\hat{q}x \\
&= c_\sigma(1 - ey - az) + c_\sigma ey + c_\sigma az - by - \hat{q}z + d\hat{q}x \\
&= c_\sigma(1 - ey - az) + \sum_{i=1}^{n}(c_\sigma - b_i)y_i \\
&\quad + \sum_{i=1}^{n}(c_\sigma - c_i)a_i z_i + d\hat{q}x \qquad (2) \\
&= c_\sigma(1 - ey - az) + \sum_{i=1}^{n}(c_\sigma - c_i)y_i \\
&\quad + \sum_{i=1}^{n}(c_\sigma - c_i)a_i z_i + d\hat{q}(x - y). \qquad (3)
\end{aligned}$$

[34] The labor requirements of the consumption sector in the amount $(1/(1 + da_i)$ plus those of the investment sector arising from replacement for depreciation in the amount $da_i/(1 + da_i)$ add up to the total labor available.

[35] As alluded to in Footnote 27 above.

[36] When the felicity function is linear, the magnitudes of $\hat{p}$ and $\hat{q}$ are identical, although their units remain different. Note also the identities $\hat{q}_i = a_i c_i$ and $c_i + d\hat{q}_i = b_i$ for all i.

[37] Note that in the derivation of (2) and (3) below, we appeal to the identities referred to in Footnote 36.

Because $(x, x') \in \Omega$, $z \geq 0$. Because $y \in \Lambda(x, x')$, $x \geq y$ and $1 - ey - az \geq 0$. We can now appeal to our standing hypothesis as described in (1) to assert that

$$by - b\hat{y} \leq \hat{q}x - \hat{q}x'. \tag{4}$$

Given our hypotheses on the felicity function w, we obtain as a consequence of (4),

$$w(by) - w(b\hat{y}) \leq w'(b\hat{y})(by - b\hat{y}) \leq w'(b\hat{y})(\hat{q}x - \hat{q}x') = (\hat{p}x - \hat{p}x').$$

A simple transposition of terms completes the proof. □

We can now state the principal result of this section.[38]

Theorem 1 *There exists a unique golden-rule stock* $\hat{x} = (1/(1 + da_\sigma))e(\sigma)$.

PROOF: Let $\hat{y} = \hat{x} = (1/(1 + da_\sigma))e(\sigma)$, and check that $(\hat{x}, \hat{x}) \in \Omega$, and $\hat{y} \in \Lambda(\hat{x}, \hat{x})$. Next, appeal to Lemma 1 to assert that $(\hat{x}, \hat{x})$ is a solution to the problem specified in Definition 5 and, hence, that $\hat{x}$ is a golden-rule stock.

We can also show that it is a unique solution to this problem. Suppose, in contrast, that $(\tilde{x}, \tilde{x}')$ is another solution with a corresponding $\tilde{y} \in \Lambda(\tilde{x}, \tilde{x}')$ and $\tilde{z}' = \tilde{x}' - (1 - d)\tilde{x}$. Because $w(\cdot)$ is strictly increasing, $b\tilde{y} = b\hat{y} = c_\sigma$. On substituting $\tilde{x}$, $\tilde{y}$ and $\tilde{z}$ for x, y and z in (3) above, we obtain the fact that the right-hand side of (3) equals zero, which implies that each of its four terms is zero. This implies that $\tilde{y}_i = 0 = \tilde{z}_i$ for $i \neq \sigma$, that $\tilde{x}_i = \tilde{y}_i$ for all i, and that $\tilde{y}_\sigma + a_\sigma \tilde{z}_\sigma = 1$. Coupling the first assertion with the equality $b\tilde{y} = c_\sigma$, we obtain that $\tilde{y}_\sigma = 1/(1 + da_\sigma)$ and, hence, from the third assertion that $\tilde{x} = 1/(1 + da_\sigma)e(\sigma)$. From the last assertion we can then obtain that $\tilde{z}_\sigma = d/(1 + da_\sigma)$ and, hence, that $\tilde{x}' = \tilde{z} - (1 - d)\tilde{x} = (d/(1 + da_\sigma) + (1 - d)/(1 + da_\sigma))e(\sigma) = (1/(1 + da_\sigma))e(\sigma)$. The demonstration is complete. □

4 Existence of an optimal program

In this section, we prove the existence of an optimal program from an arbitrarily given initial stock. We follow the methods of Brock (1970), which in turn build on those of Gale (1967): this methodology relies on the concept of a *good* program and then exploits the assumption of a unique golden-rule stock to deduce the average turnpike property of such a program. We follow the same conceptual benchmarks in the context of the RSS model and present a unified treatment both to highlight certain steps that are crucial for subsequent argumentation and to avoid possibly confusing cross-referencing.[39]

Definition 6 *A program* $\{x(t), y(t)\}$ *is good if there exists* $G \in \mathbb{R}$ *such that* $\sum_{t=0}^{T}(w(by(t)) - w(b\hat{y})) \geq G$ for all $T \in \mathbb{N}$. *A program is bad if* $\lim_{T\to\infty} \sum_{t=0}^{T}(w(by(t)) - w(b\hat{y})) = -\infty$.

Proposition 2 *There exists a good program from any arbitrary initial stock* $x_o \in \mathbb{R}^n_+$.

PROOF: For each $t \in \mathbb{N}$, let $z(t + 1) = d\hat{x}$. Define $y(0) = 0$, and $y(t + 1) = (1 - d)y(t) + d\hat{x}$ for $t \in \mathbb{N}$. Then, $y(t)$ is monotonically non-decreasing,

[38] We remind the reader of our standing hypothesis as expressed in (1).

[39] Note that we cannot directly apply the relevant theorems in Gale (1967), Brock (1970) or McKenzie (1968, 1987) because the assumptions of these theorems are not directly satisfied; instead, the concrete structure of the RSS model allows a simplification of the arguments.

and converges to $\hat{x}$ as $t \to \infty$. Given an arbitrary initial stock, x_o, define $x(0) = x_o$, and for each $t \in \mathbb{N}$, $x(t+1) = (1-d)x(t) + z(t+1)$. Then, it is easy to check that $\{x(t), y(t)\}$ is a program from x_o. Given the definition of the sequence $\{y(t)\}$, we have $(by(t) - b\hat{x}) = (1-d)^t(by(1) - b\hat{x})$ for $t \geq 2$, and $by(t) \geq db\hat{x}$ for $t \in \mathbb{N}_+$. Therefore, we have for $t \in \mathbb{N}_+$,

$$[w(b\hat{x}) - w(by(t))] \leq w'(by(t))(b\hat{x} - by(t)) \leq w'(db\hat{x})(b\hat{x} - by(1))(1-d)^{t-1}.$$

Therefore, the sequence $\{w(b\hat{x}) - w(by(t))\}$ is summable, and so $\{x(t), y(t))\}$ is a good program from x_o. □

Proposition 3 *For any program* $\{x(t), y(t)\}$, *there exists* $M(x(0)) \in \mathbb{R}_+$ *such that for any* $t_1 \in \mathbb{N}$, *and any integer* $t_2 \geq t_1$, $\sum_{t=t_1}^{t_2}(w(by(t)) - w(b\hat{y})) \leq M(x(0))$.

PROOF: From Lemma 1, for any $t_2 \geq t_1$, $\sum_{t=t_1}^{t_2} w(by(t)) - w(b\hat{y}) \leq \hat{p}(x(t_1) - x(t_2+1)) \leq \hat{p}x(t_1) \leq m(x(0)) \sum_{j=1}^{n} \hat{p}_j$. Let $M(x(0)) = m(x(0))w'(b_\sigma/1 + da_\sigma) \sum_{j=1}^{n} a_i b_i/(1 + da_i)$ to complete the proof. □

Proposition 4 *Any program that is not good is bad.*

PROOF: For any program $\{x(t), y(t)\}$ that is not good, and for any $N \in \mathbb{R}$, there exists T_N such that $\sum_{\tau=0}^{T_N}(w(by(\tau)) - w(b\hat{y})) \leq N - M(x(0))$, $M(x(0))$ the real number whose existence is asserted in Proposition 3. By choosing $t_1 = T_N + 1$ and $t_2 = t > T_N + 1$ in Proposition 3, we obtain that $\sum_{\tau = T_N + 1}^{t}(w(by(\tau)) - w(b\hat{y})) \leq M(x(0))$ for all $t > T_N + 1$. On adding these two expressions, we obtain that $\sum_{\tau=0}^{t}(w(by(\tau)) - w(b\hat{y})) \leq N$ for all $t > T_N$, and complete the proof. □

Definition 7 *A program* $\{x(t), y(t)\}$ *exhibits the average turnpike property if* $\lim_{T\to\infty}(\bar{x}(T), \bar{y}(T)) = (\hat{x}, \hat{y})$, *where* $\bar{x}(T) = (1/T)\sum_{t=0}^{T-1} x(t)$ and $\bar{y}(T) = (1/T)\sum_{t=0}^{T-1} y(t)$ for all $T \in \mathbb{N}_+$.

The proofs of Propositions 5, 8 and 9 are technical, and available on request.

Proposition 5 *Every good program exhibits the average turnpike property.*

For any $y \in \Lambda(x, x')$ and any $(x, x') \in \Omega$, let

$$\delta_{(\hat{x},\hat{p})}(x, x') = w(b\hat{y}) - w(by) - \hat{p}(x' - x) = \hat{p}(x - x') - (w(by) - w(b\hat{y})). \quad (5)$$

Whenever there is no possibility of confusion, we shall abbreviate $\delta_{(\hat{x},\hat{p})}(x(t), x(t+1))$ by $\delta(t)$ for any program $\{x(t), y(t)\}$. We shall refer to $\{\delta(t)\}$ as the value-loss sequence associated with the program $\{x(t), y(t)\}$.

Proposition 6 *The value-loss sequence* $\{\delta(t)\}_{t\in\mathbb{N}}$ *of any program* $\{x(t), y(t)\}$ *is non-negative, and*

$$\sum_{t=0}^{T}(w(by(t)) - w(b\hat{y})) = \hat{p}(x(0) - x(T+1)) - \sum_{t=0}^{T}\delta(t) \text{ for all } T \in \mathbb{N}.$$

PROOF: For each $t \in \mathbb{N}$, let $\delta(t) = \hat{p}(x(t) - x(t+1)) - (w(by(t)) - w(b\hat{y}))$. Because $\{x(t), y(t)\}$ is a program, we can appeal to Lemma 1 to assert that $\delta(t) \geq 0$ for all $t \in \mathbb{N}$. On summing over t, and rearranging, we complete the proof of the assertion. □

We now define the aggregate value-loss associated with any program as

$$\Delta(x_o) = \inf\left\{\sum_{t=0}^{\infty}\delta(t) : \{x(t), y(t)\} \text{ is a program from } x_o\right\}.$$

Our next two results assert that this infimum is a finite number and that it can be attained.

Proposition 7 *The value-loss sequence $\{\delta(t)\}_{t\in\mathbb{N}}$ of any program $\{x(t), y(t)\}$ is summable if and only if it is good. Hence, $\lim_{t\to\infty}\delta(t) = 0$ for any good program.*

PROOF: For any good program, Proposition 6 allows us to assert the existence of $G \in \mathbb{R}$ such that for all $t \in \mathbb{N}$,

$$\begin{aligned}\sum_{t=0}^{T}\delta(t) &= \hat{p}(x(0) - x(T+1)) - \sum_{t=0}^{T}(w(by(t)) - w(b\hat{y})) \\ &\leq \hat{p}(x(0) - x(T+1)) - G \leq \hat{p}(x(0)) - G\end{aligned}.$$

Because $\sum_{t=0}^{\infty}\delta(t)$ is a finite number, certainly $\lim_{t\to\infty}\delta(t) = 0$. However, the first equality and Proposition 1 allows us to assert that a program with a summable value-loss sequence is good. □

Proposition 8 *There exists a program $\{x'(t), y'(t)\}$ from an arbitrary initial stock x_o such that its associated value-loss sequence $\{\delta'(t)\}$ satisfies $\sum_{t=0}^{\infty}\delta'(t) = \Delta(x_o)$ where $0 \leq \Delta(x_o) < \infty$.*

Proposition 9 *A program $\{x(t), y(t)\}$ whose associated value-loss sequence $\{\delta(t)\}$ satisfies $\sum_{t=0}^{\infty}\delta(t) = \Delta(x(0))$ is optimal.*

Theorem 2 *For any arbitrary initial stock, $x_o \in \mathbb{R}^n_+$, there exists an optimal program from x_o. If the initial stock x_o equals $\hat{x} = \hat{y} = (1/(1 + da_\sigma))e(\sigma)$, then the stationary program $\{\hat{x}, \hat{y}\}$ is an optimal program from x_o.*

PROOF: Proposition 9 guarantees that the program whose existence is asserted in Proposition 8 is optimal. For the second claim, simply note that the aggregate value-loss of the stationary program is (trivially) zero and that an appeal to Proposition 9 completes the argument. □

5 Choice of techniques in the long run

We are now in a position to describe what the economy looks like in the long run. Towards this end, we begin with a characterization of the von Neumann facet as described in McKenzie (1968, 1986). It is of interest that under our standing hypothesis as described in (1), this reduces to a line in the Euclidean space of dimension $2n$.

Lemma 2 *The von Neumann facet $\{(x, x') \in \Omega : \delta_{(\hat{p},\hat{x})}(x, x') = 0\}$ is a subset of $\{(x, x') \in \Omega : x'_i = x_i = 0, i \neq \sigma, x'_\sigma = (1/a_\sigma) + \xi_\sigma x_\sigma\}$, $\xi_\sigma = 1 - d - (1/a_\sigma)$, with equality if the felicity function w is linear. If the felicity function is strictly concave, the facet is the singleton $\{(\hat{x}, \hat{x})\}$.*

PROOF: Pick $(\tilde{x}, \tilde{x}') \in \Omega$ and $\tilde{y} \in \Lambda(\tilde{x}, \tilde{x}')$ such that $\delta_{(\hat{x},\hat{p})}(\tilde{x}, \tilde{x}') = 0$. From (5) we obtain $w(b\tilde{y}) - w(b\hat{y}) + \hat{p}(\tilde{x}' - \tilde{x}) = 0$. On appealing to the concavity of $w(\cdot)$, this reduces to:

$$w(b\tilde{y}) - w(b\hat{y}) \leq w'(b\hat{y})(b\tilde{y} - b\hat{y}) \Longrightarrow b\hat{y} - b\tilde{y} - q(\tilde{x}' - \tilde{x}) \leq 0. \tag{6}$$

This combined with (4) and (3) yields:

$$c_\sigma(1 - e\tilde{y} - a\tilde{z}) + \sum_{i=1}^{n}(c_\sigma - c_i)\tilde{y}_i + \sum_{i=1}^{n}(c_\sigma - c_i)a_i\tilde{z}_i + dq(\tilde{x} - \tilde{y}) = 0.$$

This implies that $\tilde{z}_i = 0 = \tilde{y}_i = \tilde{x}_i = \tilde{x}_i'$ for all $i \neq \sigma$. Furthermore, that $\tilde{y}_\sigma = \tilde{x}_\sigma$ and that

$$\tilde{y}_\sigma - a_\sigma\tilde{z}_\sigma = 1 \Longrightarrow \tilde{x}_\sigma + a_\sigma(\tilde{x}_\sigma' - (1-d)\tilde{x}_\sigma) = 1 \Longrightarrow \tilde{x}_\sigma' = (1/a_\sigma) + \xi_\sigma\tilde{x}_\sigma.$$

Now suppose that $w(\cdot)$ is strictly concave and that $b\tilde{y} \neq b\hat{y}$. We then obtain a strict inequality in (6) and, thereby, contradict (4). Therefore, $b\tilde{y} = b\hat{y} = c_\sigma$. On appealing to the computations above, we obtain that $\tilde{y}_\sigma = 1/(1 + da_\sigma) = \tilde{x}_\sigma$ and, hence, that $\tilde{x}_\sigma' = (1/a_\sigma) + \xi_\sigma\tilde{x}_\sigma = 1/(1 + da_\sigma)$.

For the reverse implication in the linear case, pick $(x, x') \in \Omega$ such that $x_\sigma' = (1/a_\sigma) + \xi_\sigma x_\sigma$, $x_i' = x_i = 0$, $i \neq \sigma$, and $y_\sigma = x_\sigma$. On substituting these values in the left-hand side of (3), we see that it is equal to zero. But that is precisely $\delta_{(\hat{x},\hat{p})}(x, x)$ in the linear case. □

Before we present the principal result of this section, we record the following observation.

Proposition 10 *Any optimal program is good.*

PROOF: Let $\{x(t), y(t)\}$ be an optimal program, and suppose it is not good. By Proposition 2, there exists a good program $\{x'(t), y'(t)\}$ starting from $x(0)$. Hence, there exists $G \in \mathbb{R}$ such that for all $T \in \mathbb{N}_+$, $\sum_{t=0}^{T}(w(by'(t)) - w(b\hat{y})) \geq G$. Pick any $\varepsilon > 0$, and appeal to Proposition 4 to guarantee the existence of t_ε such that $\sum_{t=0}^{T}(w(by(t)) - w(b\hat{y})) < G - \varepsilon$ for all $T \geq t_\varepsilon$. Putting these two expressions together, we obtain that $\sum_{t=0}^{T}(w(by'(t)) - w(by(t))) > \varepsilon$ for all $T \geq t_\varepsilon$ and, hence, a contradiction to the fact that $\{(x(t), y(t)\}$ is an optimal program. □

We can now present

Theorem 3 *Any optimal program* $\{x(t), y(t)\}$ *converges to the von Neumann facet and, therefore,* $\lim_{t\to\infty} x_i(t) = \lim_{t\to\infty} y_i(t) = \lim_{t\to\infty} z_i(t) = 0$ for all $i \neq \sigma$. If the felicity function $w(\cdot)$ is strictly concave, $\lim_{t\to\infty} x(t) = \lim_{t\to\infty} y(t) = (1/1 + da_\sigma)e(\sigma)$ *and* $\lim_{t\to\infty} z(t) = (da_\sigma/1 + da_\sigma)e(\sigma)$.

PROOF: Suppose that there exists $\varepsilon > 0$ such that for all $k \in \mathbb{N}_+$, there exists $t(k) \geq k$ such that $\sum_{i\neq\sigma} \|x(t(k))\| > \varepsilon$. We can then assert that for the value-loss sequence $\{\delta(t(k))\}_{k\in\mathbb{N}_+}$, there exists $\delta_o > 0$ and $k_o \in \mathbb{N}_+$ such that for all $k \geq k_o$, $\delta(t(k)) \geq \delta_o$. If the assertion is valid, we obtain a contradiction to Proposition 7 and complete the proof of the first claim. Therefore, suppose that the assertion is false. Then we can manufacture a sequence of integers $\{k_i\}_{i\in\mathbb{N}_+}$ such that $\lim_{i\to\infty} \delta(t(k_i)) = 0$. Now consider the sequence $\{(x(t(k_i)), x(t(k_i) + 1))\}_{i\in\mathbb{N}_+}$ and appeal to Proposition 1 to guarantee the existence of a subsequence that converges to a point $(\tilde{x}, \tilde{x}')$. Because Ω is closed, and $w(\cdot)$ is continuous, $\delta_{(\hat{p},\hat{x})}(\tilde{x}, \tilde{x}') = 0$. We now appeal to Lemma 2 to obtain a contradiction to our initial hypothesis.

For the case of a strictly concave felicity function, repeat the argument above, but with $\|x(t(k)) - \hat{x}\| + \|x(t(k+1)) - \hat{x}\| > \varepsilon$. In this case, $(\tilde{x}, \tilde{x}') = (\hat{x}, \hat{x})$, and we again appeal to Lemma 2 to obtain a contradiction to our initial hypothesis. □

6 Choice of techniques in transition

In this section, we turn to the question of which machines are optimally used and produced – the choice of techniques – not only in the long run, but also in the medium to short run. Our discussion revolves around the identification and formalization of a policy prescribed in Stiglitz (1968). Therefore, for the case of a linear felicity function, we present simple examples of economies, consisting of only a single type of machine and, thereby, posing no issue as to the choice of technique in production,[40] in which consumption and capital stock exhibit a two-period cycle along an optimal program or a four-period cycle along a Stiglitz program, which is, thereby, shown to be bad. The first shows the optimality of periodic over-building and over-consuming relative to the golden-rule levels, and the second, the non-optimality of a no excess-capacity policy: phenomena that Stiglitz (1968) apparently does not encounter in his continuous-time formulation of the model.[41] Leaving aside questions relating to the utilization of machines, and focusing only on their production, we also present an example of an economy consisting of two types of machines in which a machine other than the golden-rule machine is produced in the very first period along an optimal program. For a non-linear felicity function, this establishes the non-optimality of a Stiglitz production program (Definition 9 below), and gives an affirmative answer to the question as to whether there is a compelling reason to ever produce a type of machine that we know would eventually be depreciated to zero.[42]

6.1 Stiglitz' policy prescriptions

We present Stiglitz' policy prescriptions in the form of particular programs. Towards this end, let $D = \{i \in \{1, \ldots, n) : b_i \geq c_\sigma = b_\sigma/(1 + da_\sigma)\}$ be the set of machine-types whose output per unit labor ratios are not less than the effective output per unit labor ratio of machines of type σ.[43] We shall refer to such types as desirable and to those not in D as undesirable. Under a Stiglitz policy, labor is allocated in each time period to a set of available desirable machines with a higher type of machine having a priority over a lower one,[44] and

[40] The question of the choice of technique in terms of use of course remains: should all of the stocks of the machine be utilized until production is undertaken? This question is investigated in Section 6.3 below.

[41] It is worth reemphasizing in this connection that Brock's 1970 results were not available to Stiglitz in 1968.

[42] Stiglitz (1973) has an example of a four machine economy where such a phenomenon can occur, but it is with discounting and a minimum consumption constraint. In this work, Stiglitz is primarily interested in what he calls the phenomena of "recurrence": a situation when a machine once in service is put out of service to be brought back into service again.

[43] See Footnote 27 and the associated text. Note that $\sigma \in D$.

[44] Recall that without any loss of generality, the machine types have been numbered so that $b_i \geq b_{i+1}$ for all $i = 1, \ldots, n$.

any remaining labor allocated towards producing only one type of machine, that delineated by σ. More formally,

Definition 8 *A program* $\{x(t), y(t)\}$ *with an associated gross investment sequence* $\{z(t+1)\}$ *is said to be a Stiglitz program if for any* $t \in \mathbb{N}$ *the following policy prescriptions are followed.*

(i) *If* $0 \leq \sum_{i\in D} x_i(t) \leq 1$, *let* $y_i(t) = x_i(t)$ *for all* $i \in D$, $y_i(t) = 0$ *for all* $\notin D$, *and* $z(t+1) = ((1 - \sum_{i\in D} x_i(t))/a_\sigma)e(\sigma)$.
(ii) *If* $\sum_{i\in D} x_i(t) > 1$ *and* $x_1(t) \geq 1$, *let* $y(t) = e(1)$ *and* $z(t+1) = 0$.
(iii) *If* $\sum_{i\in D} x_i(t) > 1$ *and* $x_1(t) < 1$, *let* $y_i(t) = x_i(t)$ *for all* $i = 1, \ldots, i_o - 1$, $y_{i_o}(t) = 1 - \sum_{i=1}^{i_o-1} x_i(t)$, *where* $i_o \in D$ *such that* $\sum_{i=1}^{i_o-1} x_i(t) < 1$ *and* $\sum_{i=1}^{i_o} x_i(t) \geq 1$, *and* $z(t+1) = 0$.

It is perhaps uncontroversial that the crucial aspect of the issue of choice of technique relates to production rather than the utilization of the correct type of machine. In keeping with this, we are also interested in the following kind of programs that contain, as a strict subset, the set of Stiglitz programs.

Definition 9 *A program* $\{x(t), y(t)\}$ *with an associated gross investment sequence* $\{z(t+1)\}$ *is said to be a Stiglitz production program if for any* $t \in \mathbb{N}$, $z_i(t+1) = 0$ *for all* $i \neq \sigma$.

In a continuous-time framework of our model, with a linear felicity function, Stiglitz (1968) has argued that an optimal program must follow the policy prescriptions described above. This result turns out to be invalid in our framework, and we provide three examples in the next three subsections to illustrate this observation. In Section 7, we provide alternate sets of sufficient conditions under which the Stiglitz' assertion is valid in our framework.

6.2 Non-monotonicity of an optimal program

Stiglitz (1968, p. 607) notes that an implication of his policy prescriptions is that employment and output in the consumption good sector increases monotonically, if the economy is initially capital poor. Our first example shows that such a monotonicity property is invalid in general in our framework.

We present an example of an economy with a linear felicity function in which the optimal path cycles around the golden rule stock. The economy has available to it only one type of machine whose (labor) input and output coefficients (a_1, b_1) are given by (2/3,1), the depreciation rate d by 1/2, and the felicity function by $w(by) = y$. The reduced form of the economy is given by:

$$\begin{aligned}
\Omega &= \left\{(x, x') \in \mathbb{R}^2_+ : (1/2)x + (3/2) \geq x' \geq (1/2)x\right\},\\
\Lambda(x, x') &= \{y \in \mathbb{R}_+ : y \leq x \text{ and } y \leq 1 + (1/3)(x - 2x')\}\\
&= \{y \in \mathbb{R}_+ : y \leq \min[1 + (1/3)(x - 2x'), x]\},\\
u(x, x') &= \max\{w(by) : y \in \Lambda(x, x'), (x, x') \in \Omega\}\\
&= \min[1 + (1/3)(x - 2x'), x].
\end{aligned}$$

Consider the program $\{x(t), y(t)\}$ given by $x(t) = y(t) = 3/4$, with a gross investment of $z(t+1) = x(t+1) - (1-d)x(t) = 3/4 - (1/2)(3/4) = 3/8$, for all $t \in N$. We claim that this is a stationary optimal program from $x(0) = 3/4$. Towards this end, we show that (3/4, 3/4) is the unique solution to the problem delineated in Definition 5 and, hence, that 3/4 is the unique golden-rule stock.

First observe that $u(3/4, 3/4) = \min[1 - (1/3)(3/4), 3/4] = 3/4$, and that $u(x, x') = \min[1 - (1/3)x - (2/3)(x' - x), x]$. Now if $0 \leq x < 3/4$, $u(x, x') < (3/4)$. And if $x > 3/4$, then $x' \geq x$ implies $1 - (1/3)x - (2/3)(x' - x) \leq 1 - (1/3)x < 3/4$. Hence, $u(x, x') < (3/4) = u(3/4, 3/4)$, and the argument is complete.

Next, consider a program such that $y(t) = x(t)$ for all $t \in \mathbb{N}$, $x(t) = 1/2$ for all even $t \in \mathbb{N}$, and $x(t) = 1$ for all odd $t \in \mathbb{N}$. It is easy to check that this is a program that starts from 1/2 and oscillates around 3/4. All that we need to show is that it is an optimal program starting from 1/2. Towards this end, we note that $\hat{p} = \hat{q} = 1/2$, and that this program makes a zero value-loss in each period at these prices:

$$\delta(t) = \begin{cases} (1/2) + (1/2) - (1/2)(1/2) - 3/4 = 0 & \text{for } t = 0, 2, \cdots \\ 1 + (1/2)(1/2) - (1/2) - 3/4 = 0 & \text{for } t = 1, 3, \cdots. \end{cases}$$

An appeal to Proposition 9 then completes the argument.

6.3 Non-optimality of a Stiglitz program

It can be easily checked in the example of the previous subsection that the cyclic optimal program is a Stiglitz program. In this subsection, we ask whether the set of optimal programs is identical to the set of Stiglitz programs and, perhaps surprisingly, discover this to be decisively not the case. We present an example of a simple economy with a linear felicity function in which at a particular initial stock, the unique Stiglitz program is bad, leave alone optimal.

The economy has available to it only one type of machine whose (labor) input and output coefficients (a_1, b_1) are given by (2/5,1), the depreciation rate d by 1/2, and the felicity function by $w(by) = y$. The reduced form of the economy is given by:

$$\begin{aligned} \Omega &= \left\{(x, x') \in \mathbb{R}_+^2 : (1/2)x + (5/2) \geq x' \geq (1/2)x\right\}, \\ \Lambda(x, x') &= \{y \in \mathbb{R}_+ : y \leq x \text{ and } y \leq 1 + (1/5)(x - 2x')\} \\ &= \{y \in \mathbb{R}_+ : y \leq \min[1 + (1/5)(x - 2x'), x]\}, \\ u(x, x') &= \max\{w(by) : y \in \Lambda(x, x'), (x, x') \in \Omega\} \\ &= \min[1 + (1/5)(x - 2x'), x]. \end{aligned}$$

Consider the program $\{x(t), y(t)\}$ given by $x(t) = y(t) = 5/6$, with a gross investment of $z(t+1) = x(t+1) - (1-d)x(t) = 5/6 - (1/2)(5/6) = 5/12$, for all $t \in N$. We claim that this is a stationary optimal program from $x(0) = 5/6$. Towards this end, we show that (5/6, 5/6) is the unique solution to the problem delineated in Definition 5 and, hence, that 5/6 is the unique golden-rule stock.

First, observe that $u(5/6, 5/6) = \min[1 - (1/5)(5/6), 5/6] = 5/6$, and that $u(x, x') = \min[1 - (1/5)x - (2/5)(x' - x), x]$. Now if $0 \leq x < 5/6$, $u(x, x') < (5/6)$.

And if > 5/6, then $x' \geq x$ implies $1-(1/5)x-(2/5)(x'-x) \leq 1-(1/5)x < 5/6$. Hence, $u(x, x') < (5/6) = u(5/6, 5/6)$, and the argument is complete.

Next, consider a program such that for all $t \in \mathbb{N}$, $x(4t) = 1 = y(4t)$, $x(4t+1) = 1/2 = y(4t+1)$, $x(4t+2) = 3/2$, $y(4t+2) = 1$, $x(4t+3) = 3/4 = y(4t+3)$. It is easy to check that this is a program that starts from 1 and returns to it after 4 periods. It is also easy to see that it is a unique Stiglitz program starting from 1. In terms of Definition 8, $D = \{1\}$, and in 3 of the 4 periods of the 4-period cycle, condition (ii) applies and usage and production levels are uniquely set to maintain full employment and no excess capacity. In other words, in these periods, all of the desirable machines are utilized, and all of the remaining labor (none in 1 of the 3 periods) is allocated to the production of new machines. In the 1 remaining period, $4t+2$, there is full employment, but also excess capacity.

It is easy to check that $u(1, 1/2) = 1$, $u(1/2, 3/2) = 1/2$, $u(3/2, 3/4) = 1$, and $u(3/4, 1) = 3/4$. Hence, for all $n \in \mathbb{N}_+$, $\sum_{t=0}^{4n}[u(x(t), x(t+1)) - (5/6)] = -(1/12)n$, so that the Stiglitz program is bad. From Proposition 10, we can then conclude that the Stiglitz program is not optimal.

Because there is a unique Stiglitz program starting from $x(0) = 1$, the optimal program from $x(0) = 1$, which exists by virtue of Theorem 2, is not a Stiglitz program.

6.4 Non-optimality of a Stiglitz production program

We know from the example presented in the previous subsection that an optimal program is not in general a Stiglitz program. In this section, we ask whether every optimal program is at least a Stiglitz production program. Note that in an economy with only one type of machine (as in the examples of the previous two subsections) an optimal program is trivially a Stiglitz production program. Therefore, we need to consider economies with at least two types of machines for the question to be non-trivial. We present an example of an economy with two types of machines and a piece-wise linear felicity function in which the machine different from the golden-rule machine is produced in transition along an optimal program.

Consider an economy in which there are two types of machines ($n = 2$) with input coefficients vector given by $a = (2, 3)$, output coefficients vector by $b = (4, 5)$ and the depreciation rate, d, by 0.45 ($m \equiv (1-d) = 0.55$). Note that

$$\frac{b_1}{a_1} = 2 > (5/3) = \frac{b_2}{a_2} \text{ and } \frac{b_1}{(1+da_1)} \approx 2.1052 < 2.1276 \approx \frac{b_2}{(1+da_2)},$$

and, therefore, $\sigma = 2$: machines of type 2 constitute the golden-rule stock. The social welfare function, w, is defined as follows:

$$w(y) = \begin{cases} y-2 & \text{for } y \geq 2 \\ 1000(y-2) & \text{for } 0 \leq y < 2. \end{cases}$$

The initial stock of machines is specified as $x_o = (0.5, 0)$.

We know from the analysis of Section 5 that in the long run only machines of type 2 will be produced and used along an optimal program. We are interested in demonstrating that machines of type 1 will nevertheless be produced in some time period along an optimal program from x_o. Our method of demonstrating this is to suppose, in contrast, that machines of type 1 are not produced in period 1 along an optimal program. We show that a consequence of this is that an optimal program will suffer a large disutility (negative utility, large in absolute value) in either the first or the second period of consumption, which results in a large disutility even in the long run. We construct a program from x_o that produces machines of type 1 initially, and reaches the golden-rule stock in a finite number (specifically, 8) of periods; it has non-negative utility in all periods and, of course, (positive) golden-rule utility from period 9 onwards. This shows that our hypothesis that machines of type 1 are not produced on the optimal program in period 1 must be false, and completes the demonstration. The computational details are available from the authors on request.

In conclusion to this section, note that the felicity w used in the above example is non-linear, but not strictly concave. We can check that all the calculations shown in the appendix below remain valid with a strictly concave w defined as follows:

$$w(y) = \begin{cases} (53/47)2(y-2)/(y-1) & \text{for } y \geq 2 \\ 1000(y-2) - 0.5(y-2)^2 & \text{for } 0 \leq y < 2. \end{cases}$$

7 Sufficient conditions for the optimality of the Stiglitz policy

The three examples presented in Section 6 show that an optimal program in our framework does not always follow the policy prescriptions of Stiglitz (1968). However, we can provide sufficient conditions under which it does. This section is devoted to presenting results along this line. In the first subsection, we consider a linear felicity function, and in a subsequent section, turn to the general case.

7.1 Case of a linear felicity function

The point to be noted about the two examples presented above (in subsections 6.2 and 6.3) is the particular value of the parameter $1 - d - (1/a_1)$. We have already referred to this parameter in Section 5 as ξ_1, and it takes the value -1 in the example in Section 6.2, and the value -2 in the example in Section 6.3. The sufficient condition for the optimal choice of technique that we present in this section then[45] requires that $\xi_\sigma \geq -1$. (Note that $\xi_\sigma < 1$.)

Theorem 4 *With $1 > \xi_\sigma \geq -1$ and w a linear function, any Stiglitz program is an optimal program.*

This theorem is a consequence of the following lemma.

Lemma 3 *With a linear felicity function w, and with $1 > \xi_\sigma \geq -1$, the aggregate value-losses of any Stiglitz program starting from $x(0)$ equal $\Delta(x(0))$.*

[45] Note that by default $D = \{1\} = \{\sigma\}$ in each of the one-machine examples considered above.

The proof of Lemma 3 relies crucially on the sources of value-loss already identified in the proof of Lemma 1. On rewriting (2), we obtain:

$$\begin{aligned}\delta(t) &= b\hat{y} - by(t) - \hat{q}(x(t+1) - x(t)) \\ &= c_\sigma(1 - ey(t) - az(t+1)) + \sum_{i=1}^{n}(c_\sigma - b_i)y_i(t) \\ &\quad + \sum_{i=1}^{n}(c_\sigma - c_i)a_i z_i(t+1) + dqx(t) \\ &= \alpha(t) + \sum_{i\in D}(c_\sigma - b_i)y_i(t) + \sum_{i\notin D}(c_\sigma - b_i)y_i(t) \\ &\quad + \sum_{i=1}^{n}(c_\sigma - c_i)a_i z_i(t+1) + dqx(t), \end{aligned} \tag{7}$$

where $\alpha(t) = c_\sigma(1 - ey(t) - az(t+1))$ is the value-loss from unemployment.[46] This is a five-fold decomposition of the value-loss at any time-period: the other four terms concern value losses from incorrect usage and incorrect investment. The proof can now be executed by the comparison, period by period, of the magnitudes of the value-losses of the Stiglitz program and those of any other candidate program. We relegate the details to the appendix, and turn to a

Proof of Theorem 4 Let $\{x^s(t), y^s(t)\}$, with an associated value-loss sequence $\{\delta^s(t)\}$, be a Stiglitz program. Because only one type of machine σ is constructed under a Stiglitz' policy, and because $\xi_\sigma \geq -1$, we can appeal to Lemma 3 to assert that the Stiglitz program is a good program, and (by Proposition 5) exhibits the average turnpike property. To complete the proof, we need to establish that the hypotheses of Proposition 9 are satisfied, which is to say that $\{\delta^s(t)\}$ satisfies $\sum_{t=0}^{\infty}\delta^s(t) = \Delta(x(0))$. An appeal to Lemma 3 then completes the proof. □

Next we ask whether, under the conditions identified in Theorem 4, a Stiglitz program is uniquely optimal. Towards this end, we can present

Theorem 5 *With a linear felicity function w, and with* $-1 < \xi_\sigma < 1$, *any optimal program* $\{x(t), y(t)\}$ *is a Stiglitz program.*

Before considering the proof of this theorem, we draw the reader's attention to the fact that unlike Theorem 4, Theorem 5 does not cover the case $\xi_\sigma = -1$. Indeed, Theorem 5 is false for this case. For the economy discussed in Section 6.2, there exists an optimal program (not discussed in subsection 6.2) which is not a Stiglitz program.[47] Therefore, we need to rule out the case when the optimal program stays in the von Neumann facet but does not converge to the golden-rule values. Towards this end, we can present a result that strengthens the conclusions of Theorem 3 in the case of a linear felicity function and also shows them to hold for a Stiglitz program. (The computational details are available on request.)

---

[46] Introduced only for the typographical reason of reducing the length of the expression below.

[47] A detailed verification of this claim would lead us outside the scope of an already long paper; see Khan and Mitra (2002) for details.

Proposition 11 *With a linear felicity function w, and with* $-1 < \xi_\sigma < 1$, *for a program* $\{x(t), y(t)\}$ *that is either an optimal or a Stiglitz program,* $\lim_{t\to\infty} y_i(t) = \lim_{t\to\infty} z_i(t) = 0$ *for all* $i \neq \sigma$, *and* $\lim_{t\to\infty} y_\sigma(t) = \lim_{t\to\infty} x_\sigma(t) = \hat{x}_\sigma = 1/(1 + da_\sigma)$, $\lim_{t\to\infty} z_\sigma(t) = d/(1 + da_\sigma)$.

Next, we turn to a sharpening of Lemma 3 whose proof is a straightforward modification of the computations presented in the proof of Lemma 3 (and available on request).

Lemma 4 *Let* $\{\delta(t)\}$ *be the value-loss sequence of a program that is not a Stiglitz program and* $\{\delta^s(t)\}$ *the value-loss sequence of a Stiglitz program starting from the same initial stock. With* $-1 \leq \xi_\sigma < 1$ *and w a linear function, there exists* $\varepsilon > 0$ *such that* $\sum_{t=0}^{\infty} \delta(t) - \sum_{t=0}^{\infty} \delta^s(t) > \varepsilon$.

We can now present

PROOF OF THEOREM 5 Suppose, in contrast, that there exists an optimal program $\{x(t), y(t)\}$ with an associated value-loss sequence $\{\delta(t)\}$ that is not a Stiglitz' program. Let $\{x^s(t), y^s(t)\}$ be a Stiglitz program starting from $x(0)$ and with an associated value-loss sequence $\{\delta^s(t)\}$. An appeal to Lemma 4 and to Proposition 6 yields for all $T \in \mathbb{N}_+$,

$$\begin{aligned}\sum_{t=0}^{T}(by(t) - by^s(t)) &= \hat{p}(x^s(T+1) - x(T+1)) + \sum_{t=0}^{T} \delta^s(t) - \sum_{t=0}^{T} \delta(t) \\ &< \hat{p}(x^s(T+1) - x(T+1)) - \varepsilon.\end{aligned}$$

Using Proposition 11, we can assert that $\limsup_{T\to\infty} \sum_{t=0}^{T}(by(t) - by^s(t)) \leq (-\varepsilon)$. Therefore, we obtain $\liminf_{T\to\infty} \sum_{t=0}^{T}(by^s(t) - by(t)) = -\limsup_{T\to\infty} \sum_{t=0}^{T}(by(t) - by^s(t)) \geq \varepsilon$, a contradiction to the optimality of $\{x(t), y(t)\}$. This verifies the truth of the initial claim, and completes the argument that any optimal program is a Stiglitz program. □

7.2 Case of a general felicity function

In the previous subsection, we considered optimal programs in the context of both the long and the short run when the felicity function is linear; which is to say a situation when a *ceteris paribus* transfer of a unit of consumption from a lower consumption level period to the higher one will definitely reduce social welfare. In the first subsection, we present a price-support property and some of its implications that have independent interest, and in a subsequent subsection, use these results to offer a sufficient condition under which an optimal program is a Stiglitz production program.

7.2.1 Price-support property and its implications

So far we have worked only with the golden-rule price system, and in this subsection we present McKenzie's price support property as Theorem 6 below. Because we do not exclude the situation where the economy has no stock of machines, $x_o = 0$, the result is a direct consequence of methods available in McKenzie (1986; proof of lemma 1), rather than a corollary. The (straightforward) details of how McKenzie's interiority assumptions are fulfilled in our context, and allow his proof to work, are available on request.

Theorem 6 *Let $\{x(t), y(t)\}$ be an optimal program starting from an arbitrary initial stock, $x_o \in \mathbb{R}^n_+$. Then there exists a sequence $\{p(t)\}_{t=0}^{\infty}$, $p(t) \in \mathbb{R}^n_+$, such that for all $(x, x') \in \Omega$ and $y \in \Lambda(x, x')$,*

$$w(by(t)) + p(t+1)x(t+1) - p(t)x(t) \geq w(by) + p(t+1)x' - p(t)x.$$

Next, for the convenience of the reader, we simply state in words the consequences of Theorem 6, leaving their precise statement and proof to the appendix. We can show that along an optimal program, investment in the machine of type σ never ceases (Proposition 12), that there is investment in machines of this type only if they are valuable today (Proposition 14) and, as a consequence, they are always valuable (Proposition 15), that machines types that are valuable today were valuable in the past (Proposition 13) and that the prices are bounded (Proposition 17). We can use these results to establish expressions for the relative prices of produced machines (Proposition 18), and under the sufficient condition to be discussed below, an expression for the evolution of relative prices (Proposition 19). These results do not rely on linearity or strict concavity of the felicity function, and simply exploit the average-turnpike property of good programs and, as such, rely on the uniqueness of the golden-rule stock. Therefore, the standing hypothesis presented as (1) above continues to be the driving force.

7.2.2 A sufficient condition

The scenario in which there can be a difference in the choice of techniques is one where the short-run consumption requirements are quite different from the long-run consumption requirements on an optimal program. As we have seen in Section 5, the unique golden-rule type of machine, σ, is the best machine to use for meeting the long-run consumption requirements, regardless of whether the social welfare function is linear or strictly concave. In the short run, however, the important question is which machine built today will provide the most consumption tomorrow, given an available amount of labor for new machine production today, and without taking into account the fact that the machine depreciates. This is clearly qualitatively different from the long-run (golden-rule) problem and points to b_i/a_i rather than to $b_i/(1+da_i)$. When the orderings of these two magnitudes differ, as they do in the example of Section 6.4, one machine is best for the short-run problem and another machine is best for the long-run problem. This seems to suggest that if the orderings coincide, then the machine that is best for the long run remains best for the short run, and only the golden-rule machine is produced and used. We assume that a unique type of machine σ is best, irrespective of the time horizon under which the planning exercise is being conducted and, furthermore, that it requires for its production more labor than a machine of any other type.

Assumption 1 $(b_\sigma/a_\sigma) > \max_{i\neq\sigma}(b_i/a_i)$ and $a_\sigma > \max_{i\neq\sigma} a_i$.

We can now show that under this congruence, the golden-rule machine σ is the only type that is produced.

Theorem 7 *Under Assumption 1, an optimal program is a Stiglitz production program.*

The proof of Theorem 7 is relegated to the appendix; although the basic intuition is clear, it requires the use of all of the consequences of the price-support property that we mentioned above, and which constitute Propositions 12–19 of the appendix.

8 Concluding remarks

If we leave aside the methodological reformulation of the RSS model in the vocabulary of the Gale–McKenzie reduced form, we see the principal contribution of this work: a complete resolution of the choice of technique problem in the long run, and the identification and formalization of the Stiglitz policy as a cornerstone for the theory of transition dynamics. As regards the latter, three simple examples are of decisive importance, and they might also be of independent interest for future investigations of related issues that remain open. In conclusion, we briefly mention four of these.

Throughout this paper, we have emphasized the sharp and surprising differences that arise between our results and those of Stiglitz: in particular, the parameters ξ_i do not appear in his paper. It is of some importance to settle the issue as to whether this is a consequence of the different treatment of time in the two papers, discrete versus continuous, or to the methods that Stiglitz had to work with in 1968.[48]

In his retrospective, Stiglitz (1990, p. 61) observes the "greatest challenge facing growth theory":

> We now need to understand better the relationship between the short-run behavior of the economy – in which imperfect information and imperfect competition in financial, labor, and product markets will play a central role – and its long-run dynamics.

It is interesting that this remains a challenge even for a planning framework without uncertainty and the stark simplicity of the specifications of the RSS model, technological and otherwise. The complete characterization of the optimal path in the short run remains an open problem when the planners' felicity function is linear but $\xi_\sigma < -1$, and when it is strictly concave.

We have drawn attention (in Footnotes 3 and 21 above) to the conceptual similarities between our work and that of Mitra and Wan (1986) on the economics of forestry; it would be of interest if the analogy is analytically explored in a synthesis based on the multi-sectoral setting of Koopmans (1971) and Koopmans and Hansen (1972). This work also gives a singular prominence to Kuhn–Tucker theory.

Finally, the results reported in this paper are a testimony to the strength of the standing hypothesis that there is a unique type of machine that minimizes effective labor costs and simultaneously maximizes the steady-state consumption; see (1) above. It would be of interest to examine how the results are modified without this hypothesis.

9 Appendix

We begin with the proofs of three propositions in Section 4.

PROOF OF LEMMA 3 Let $\{x^s(t), y^s(t)\}$ be a Stiglitz program with an associated gross investment sequence $\{z^s(t+1)\}$ and an associated value-loss sequence $\{\delta^s(t)\}$. We shall denote corresponding values of any other (candidate) program starting from $x^s(0)$ by $\{x(t), y(t)\}$, $\{z(t+1)\}$, and $\{\delta(t)\}$. We shall consider three different ranges for the value of ξ_σ and make repeated use of (7) and of Definition 8.

[48] See Footnotes 9 and 41 above, and for preliminary work on this question, Khan and Mitra (2003).

Case (i) $0 < \xi_\sigma < 1$

Suppose that for any $t \in \mathbb{N}$, $0 \leq \sum_{i \in D} x_i^s(t) \leq 1$. In this case, we see from (i) of Definition 8 that $y_i^s(t) = x_i^s(t)$ for all $i \in D$, $y_i^s(t) = 0$ for all $\notin D$, and $z^s(t+1) = (1/a_\sigma)\left(1 - \sum_{i \in D} x_i^s(t)\right) e(\sigma)$. We leave it to the reader to check that $(x^s(t), x^s(t+1)) \in \Omega$ and that $y^s(t) \in \Lambda(x^s(t), x^s(t+1))$. On substituting these values in (7), we obtain that:[49]

$$\begin{aligned}\delta^s(t) &= \sum_{i \in D}(c_\sigma - b_i)x_i^s(t) + dqx^s(t) = \sum_{i \in D}(c_\sigma - b_i + dq_i)x_i^s(t) + d\sum_{i \notin D} q_i x_i^s(t)\\ &= \sum_{i \in D}(c_\sigma - c_i)x_i^s(t) + d\sum_{i \notin D} q_i x_i^s(t).\end{aligned}$$

Again from (7) we obtain:[50]

$$\begin{aligned}\delta(t) &\geq \sum_{i \in D}(c_\sigma - b_i)y_i(t) + dqx(t) \geq \sum_{i \in D}(c_\sigma - b_i)x_i(t) + dqx(t)\\ &= \sum_{i \in D}(c_\sigma - b_i + dq_i)x_i(t) + d\sum_{i \notin D} q_i x_i(t) = \sum_{i \in D}(c_\sigma - c_i)x_i(t) + d\sum_{i \notin D} q_i x_i(t).\end{aligned}$$

Next, we claim that for all $t \in \mathbb{N}$, $x_i(t) \geq x_i^s(t)$ for all $i \neq \sigma$. Because the candidate program starts from the same initial stock as the Stiglitz program, the claim holds for $t = 0$. Suppose it to be true for any $t \in \mathbb{N}$, in keeping with the induction hypothesis. Then

$$x_i^s(t+1) = (1-d)x_i^s(t) \leq (1-d)x_i(t) \leq (1-d)x_i(t) + z_i(t+1) = x_i(t+1).$$

Given the standing hypothesis, it is clear that for all $t \in \mathbb{N}$, $\delta^s(t) \leq \delta(t)$. Therefore, we need only to verify that the Stiglitz program is feasible in the sense that once in the range $0 \leq \sum_{i \in D} x_i^s(t) \leq 1$, the program always remains in it. We proceed by induction. For any $t \in \mathbb{N}$, note that

$$x_i^s(t+1) = \begin{cases}(1-d)x_i^s(t) & \text{for all } i \neq \sigma\\ (1-d)x_\sigma^s(t) + (1/a_\sigma)\left(1 - \sum_{i \in D} x_i^s(t)\right) & \text{for } i = \sigma.\end{cases} \tag{8}$$

Because $0 \leq \sum_{i \in D} x_i^s(t) \leq 1$, we obtain from (8) that $z_\sigma(t+1) = x_\sigma^s(t+1) - (1-d)x_\sigma^s(t) \geq 0$, and that

$$\begin{aligned}\sum_{i \in D} x_i^s(t+1) &= (1-d)\sum_{i \in D} x_i^s(t) + \frac{1}{a_\sigma}\left(1 - \sum_{i \in D} x_i^s(t)\right)\\ &= \left(1 - d - \frac{1}{a_\sigma}\right)\sum_{i \in D} x_i^s(t) + \frac{1}{a_\sigma} = \xi_\sigma \sum_{i \in D} x_i^s(t) + \frac{1}{a_\sigma}.\end{aligned} \tag{9}$$

Given the possible values of ξ_σ, we obtain $0 < \sum_{i \in D} x_i^s(t+1) < 1$.

We can now collect these steps to assert that for all $t \in \mathbb{N}$, $\delta^s(t) \leq \delta(t)$ and, hence, that $\sum_{t=0}^{\infty} \delta^s(t) = \Delta(x^s(0))$.

Next, we turn to the case when for any $t \in \mathbb{N}$, $\sum_{i \in D} x_i^s(t) > 1$, $x_1^s(t) < 1$. In this case, we see from (iii) of Definition 8 that $y_i^s(t) = x_i^s(t)$ for all $i < i_o$, $y_{i_o}^s(t) = 1 - \sum_{i=1}^{i_o - 1} x_i^s(t)$, $y_i^s(t) = 0$ for all $i > i_o$, and that $z_i^s(t+1) = 0$ for all i. We leave it to the reader to check that $(x^s(t), x^s(t+1)) \in \Omega$ and that $y^s(t) \in \Lambda(x^s(t), x^s(t+1))$. On

[49] Note that in the third equality we use the identity referred to in Footnote 36 above. We shall not draw attention to this in the sequel.

[50] We rely on the standing hypothesis (1) and on the definition of desirable machines, in addition to the feasibility of the program.

substituting these values in (7), we obtain that

$$\begin{aligned}\delta^s(t) &= \sum_{i=1}^{i_o-1}(c_\sigma - b_i)x_i^s(t) + (c_\sigma - b_{i_o})(1 - \sum_{i=1}^{i_o-1} x_i^s(t)) + dqx^s(t)\\ &= \sum_{i=1}^{i_o-1}(c_\sigma - c_i)x_i^s(t) + (c_\sigma - b_{i_o})(1 - \sum_{i=1}^{i_o-1} x_i^s(t)) + d\sum_{i\geq i_o} q_i x_i^s(t),\end{aligned}$$

and for any other (candidate) program with $D_o = D/\{1, \dots, i_o\}$ that

$$\begin{aligned}\delta(t) &\geq \sum_{i=1}^{i_o-1}(c_\sigma - b_i)y_i(t) + (c_\sigma - b_{i_o})y_{i_o}(t) + \sum_{i\in D_o}(c_\sigma - b_i)y_i(t) + dqx(t)\\ &\geq \sum_{i=1}^{i_o-1}(c_\sigma - b_i)x_i(t) + (c_\sigma - b_{i_o})y_{i_o}(t) + \sum_{i\in D_o}(c_\sigma - b_i)y_i(t) + dqx(t)\\ &= \sum_{i=1}^{i_o-1}(c_\sigma - c_i)x_i(t) + (c_\sigma - b_{i_o})y_{i_o}(t) + \sum_{i\in D_o}(c_\sigma - b_i)y_i(t) + d\sum_{i\geq i_o} q_i x_i(t).\end{aligned}$$

Now by hypothesis, for all $i \in D_o$, $b_i \leq b_{i_o}$ so that: $\sum_{i\in D_o}(c_\sigma - b_i)y_i(t) \geq (c_\sigma - b_{i_o})\sum_{i\in D_o} y_i(t)$. Hence, we obtain

$$(c_\sigma - b_{i_o})y_{i_o}(t) + \sum_{i\in D_o}(c_\sigma - b_i)y_i(t)) \geq (c_\sigma - b_{i_o})\sum_{i\in D_o\cup i_o} y_i(t) \geq (c_\sigma - b_{i_0})\sum_{i\in D_o\cup i_o} x_i(t).$$

Also, by definition of $\{x^s(t), y^s(t)\}$, we have $\sum_{i\in D_o\cup i_o} x_i(t) \geq \sum_{i\in D_o\cup i_o} x_i^s(t) \geq (1 - \sum_{i=1}^{i_o-1} x_i^s(t))$. Now $\sum_{i\in D} x_i^s(0) > 1$ implies that there exists a first $t_1 \in T$ such that $\sum_{i\in D} x_i^s(t_1) \leq 1$. For the Stiglitz program, we know that for all $t \in \mathbb{N}$, $t < t_1$, $z_i^s(t+1) = 0$ for all i and, hence, that $x^s(t) \leq x(t)$. In particular, $x_1^s(t) < 1$ for all $t \in \mathbb{N}$, $t < t_1$. This implies that for all $t \in \mathbb{N}$, $t < t_1$, $\delta^s(t) \leq \delta(t)$. [Note that i_o may vary with t, but given our period-by-period verification, it is of no consequence.] But for $t \geq t_1$, we are in the case considered earlier and, hence, we can assert that for all $t \in \mathbb{N}$, $\delta^s(t) \leq \delta(t)$ and, hence, that $\sum_{t=0}^{\infty} \delta^s(t) = \Delta(x^s(0))$.

Next, we turn to the case when for any $t \in \mathbb{N}$, $\sum_{i\in D} x_i^s(t) > 1$, $x_1^s(t) > 1$. In this case, we see from (ii) of Definition 8 that $y_1^s(t) = 1$, $y_i^s(t) = 0$ for all $i \neq 1$, and that $z_i^s(t+1) = 0$ for all i. We leave it to the reader to check that $(x^s(t), x^s(t+1)) \in \Omega$ and that $y^s(t) \in \Lambda(x^s(t), x^s(t+1))$. On substituting these values in (7), we obtain that $\delta^s(t) = (c_\sigma - b_1) + dqx^s(t)$, and for any other (candidate) program that $\delta(t) \geq \sum_{i\in D}(c_\sigma - b_i)y_i(t) + dqx(t)$. Now by hypothesis, $b_i \leq b_1$, so that: $\sum_{i\in D}(c_\sigma - b_i)y_i(t) \geq (c_\sigma - b_1)\sum_{i\in D} y_i(t)$. Because $\sum_{i\in D} y_i(t) \leq 1$ by definition of a program, and because $(c_\sigma - b_1) \leq 0$, $\sum_{i\in D}(c_\sigma - b_i)y_i(t) \geq (c_\sigma - b_1)$. Because both programs start from the same initial stock and $x_1^s(t) > 1$ implies $x_i^s(t-r) > 1$ for all $r = 0, \dots, t-1$, $x_i(t) \geq x_i^s(t)$ for all i. Therefore, $\delta^s(t) \leq \delta(t)$.

It is clear that if $x_1^s(0) > 1$, there exists $t_1 \in T$ such that $x_1^s(t_1) \leq 1$. We have already seen that $\delta^s(t) \leq \delta(t)$ for all $t < t_1$. Now either $\sum_{i\in D} x_i^s(t_1) \leq 1$, in which case we appeal to the first case, or $\sum_{i\in D} x_i^s(t_1) > 1$, in which case we appeal to the second case and complete the demonstration that for all $t \in \mathbb{N}$, $\delta^s(t) \leq \delta(t)$. Hence, $\sum_{t=0}^{\infty} \delta^s(t) = \Delta(x^s(0))$.

Case (ii) $0 > \xi_\sigma \geq -1$

Suppose that for any $t \in \mathbb{N}$, $(1 - d) \leq \sum_{i\in D} x_i^s(t) \leq 1$. On examining the argument for this subcase within case (i) above, we see that the value of ξ_σ is used only to verify the feasibility of the Stiglitz program in equation (9).

However, with $0 > \xi_\sigma \geq -1$, $\sum_{i \in D} x_i^s(t) \leq 1$ implies $\xi_\sigma \sum_{i \in D} x_i^s(t) \geq \xi_\sigma$ and, therefore,

$$\sum_{i \in D} x_i^s(t+1) = \xi_\sigma \sum_{i \in D} x_i^s(t) + (1/a_\sigma) \geq \xi_\sigma + (1/a_\sigma) = (1-d).$$

Furthermore, $\sum_{i \in D} x_i^s(t) \geq (1-d)$ implies $\xi_\sigma \sum_{i \in D} x_i^s(t) \leq \xi_\sigma(1-d)$ and, therefore,

$$\sum_{i \in D} x_i^s(t+1) = \xi_\sigma \sum_{i \in D} x_i^s(t) + (1/a_\sigma) \leq \xi_\sigma(1-d) + (1/a_\sigma) = (1-d)^2 + (d/a_\sigma).$$

Because $\xi_\sigma = 1 - d - (1/a_\sigma) \geq -1$, $(1/a_\sigma) \leq 2 - d$, which implies that $(d/a_\sigma) \leq (2-d)d$ and, hence, that $(1-d)^2 + (d/a_\sigma) \leq (1-d)^2 + (2-d)d = 1$. We have, therefore, shown that once in the range $(1-d) \leq \sum_{i \in D} x_i^s(t) \leq 1$, the program always remains in it.

For the other two subcases in the argument within case (i) above, we note that ξ_σ plays no role and that everything hinges on the value of d. Therefore, the only remaining case to be considered is when $0 \leq \sum_{i \in D} x_i^s(t) < (1-d)$. Here there are two possibilities: either $\sum_{i \in D} x_i^s(t+1) \leq 1$ or $\sum_{i \in D} x_i^s(t+1) > 1$. Because we have already shown that $\sum_{i \in D} x_i^s(t+1) \geq (1-d)$, there is nothing further to be shown under the first possibility. Under the second, there exists a first $t_1 \in \mathbb{N}$, $t_1 > t$ such that $\sum_{i \in D} x_i^s(t_1) \leq 1$. Because $\sum_{i \in D} x_i^s(t_1) = (1-d) \sum_{i \in D} x_i^s(t_1 - 1) > (1-d)$, we are in the case analyzed above, and the demonstration is complete.

Case (iii) $\xi_\sigma = 0$

This is a trivial case where $(1/a_\sigma) = 1/(1 + da_\sigma) = \hat{x}$. Suppose that for any $t \in \mathbb{N}$, $0 \leq \sum_{i \in D} x_i^s(t) \leq 1$. Then we see from (9) that $\sum_{i \in D} x_i^s(t) = 1/a_\sigma$. For the other subcases, the argument is identical to that presented under case (i).

We have now covered all possible cases, and the proof of the lemma is complete. □

The proof of Theorem 7 can be constructed on the basis of the following Propositions 12–19, informally described in Subsection 7.2.1., and for which the hypotheses of Theorem 7 are in force. This is to say that $\{x(t), y(t)\}$ is an optimal program, and $\{p(t)\}$ its associated price-support.

Proposition 12 *There exists a sequence $\{t_i\}_{i \in \mathbb{N}_+}$ such that $z_\sigma(t_i) > 0$ for all $i \in \mathbb{N}_+$.*

PROOF: Suppose this not to be the case. Then there exists $T \in \mathbb{N}$ such that for all $t \geq T$, $z_\sigma(t) = 0$. Because all machines depreciate at the rate $d \in (0, 1)$, this implies that $x_\sigma(t) \to 0$ as $t \to \infty$ and, therefore, that the time-average of $x_\sigma(t)$, $\bar{x}_\sigma(t) \to 0$ as $t \to \infty$. An appeal to Propositions 5 and 10 furnishes a contradiction and completes the argument. □

Proposition 13 *For any $t \in \mathbb{N}$, and any $i = 1, \ldots, n$, $(i)(1-d)p_i(t+1) \leq p_i(t)$, and $(ii) p_i(t+1) > 0 \Longrightarrow p_i(t) > 0$.*

PROOF: For any time-period t and any machine of type i, let $x = x(t) + e(i)\varepsilon$, $z = z(t+1)$, $x' = (1-d)x + z$, $y = y(t)$ where $\varepsilon > 0$. Then, $(x, x') \in \Omega$, and $y \in \Lambda(x, x')$. Using Theorem 6, we have $p_i(t+1)(1-d)\varepsilon - p_i(t)\varepsilon \leq 0$, which yields result (i). In particular, if for some $t \in \mathbb{N}$, and $i = (1, \ldots, n)$, $p_i(t+1) > 0$, then we must have $p_i(t) > 0$. □

Proposition 14 *For any $t \in \mathbb{N}$, $z_\sigma(t+1) > 0$ implies $p_\sigma(t) > 0$.*

PROOF: Suppose that for some time-period $t, z_\sigma(t+1) > 0$ and $p_\sigma(t) = 0$. Then Proposition 13 implies that $p_\sigma(t+1) = 0$. Pick ε such that $0 < \varepsilon < a_\sigma z_\sigma(t+1)$, and define $x = x(t) + \varepsilon e(\sigma)$, $x' = x(t+1) + ((1-d)x_\sigma - x_\sigma(t+1))e(\sigma)$, and $y = y(t) + \varepsilon e(\sigma)$. Then $ey = ey(t) + \varepsilon$ and $a(x' - (1-d)x) = a(x(t+1) - (1-d)x(t)) - a_\sigma z_\sigma(t+1) < az(t+1) - \varepsilon$. Therefore,

$(x, x') \in \Omega$, and $y \in \Lambda(x, x')$, and from Theorem 6, $w(by(t)) + p(t+1)x(t+1) - p(t)x(t) \geq w(by(t) + b_\sigma \varepsilon) + p(t+1)x' - p(t)x$. This yields $w(by(t)) \geq w(by(t) + b_\sigma \varepsilon)$, a contradiction to the fact that w is strictly increasing. □

Proposition 15 *For any* $t \in \mathbb{N}$, $p_\sigma(t) > 0$.

PROOF: Suppose that there exists $t \in \mathbb{N}$ such that $p_\sigma(t) = 0$. From Proposition 12, there exists a time-period $t_i > t$ such that $z_\sigma(t_i) > 0$. From Proposition 14, this implies that $p_\sigma(t_i - 1) > 0$. If $t_i = t + 1$, we obtain a contradiction. If $t_i > t + 1$, we make as many (finite) appeals to Proposition 13 as is necessary to obtain a contradiction. □

Proposition 16 *For any* $t \in \mathbb{N}$, *and any* $i \in 1, \ldots, n$, $x_i(t) > y_i(t) \Longrightarrow p_i(t+1)(1-d) = p_i(t)$.

PROOF: Let $\varepsilon = x_i(t) - y_i(t) > 0$, and define $x = x(t) - \varepsilon e_i$, $y = y(t)$, and $x' = x(t+1) - (1-d)\varepsilon e(i)$. Then, it can be easily checked that $(x, x') \in \Omega$, and $y \in \Lambda(x, x')$. Using Theorem 6, we obtain $(1-d)p_i(t+1)\varepsilon \geq p_i(t)\varepsilon$. We can now complete the proof of the claim by using Proposition 13. □

Proposition 17 $\liminf_{t\to\infty} ||p(t)|| < \infty$.

PROOF: For any $t \in \mathbb{N}$, and any $i = 1, \ldots, n$, define $x = x(t)$, $z = z(t+1) + (ey(t)/a_i)e(i)$, $x' = z + (1-d)x$, $y = 0$. Then $az + ey = az(t+1) + ey(t)$ and, hence, $(x, x') \in \Omega$ and $y \in \Lambda(x, x')$. We can now appeal to Theorem 6 to obtain $p_i(t+1)(ey(t)/a_i) \leq w(by(t)) - w(0) \leq w(be) - w(0)$. Therefore, there is $M > 0$ such that for all $t \in \mathbb{N}$, $||p(t+1)||(ey(t)) \leq M$. If $\liminf_{t\to\infty} ||p(t)|| = \infty$, then we must have $ey(t) \to 0$ as $t \to \infty$. But, then, the optimal program cannot be good, a contradiction to Proposition 10 that establishes the claim. □

Proposition 18 *For any* $i \in 1, \ldots, n$, $z_i(t+1) > 0 \Longrightarrow p_i(t+1)/p_\sigma(t+1) \geq a_i/a_\sigma$.

PROOF: Suppose that for any $t \in \mathbb{N}$, and any $i = 1, \ldots, n$, $z_i(t+1) > 0$. Define $x = x(t)$, $y = y(t)$, $z = z(t+1) - z_i(t+1)e(i) + (z_i(t+1)(a_i/a_\sigma))e(\sigma)$, $x' = (1-d)x + z$. Because $az + ey = az(t+1) + ey(t) \leq 1$, $(x, x') \in \Omega$ and $y \in \Lambda(x, x')$. We can now appeal to Theorem 6 to obtain $p_i(t+1)z_i(t+1) \geq p_\sigma(t+1)z_i(t+1)(a_i/a_\sigma)$. Because $z_i(t+1) > 0$, the proof of the claim is complete. □

Proposition 19 *Under Assumption 1, for any time-period* $t \in \mathbb{N}$, *and any* $i \in 1, \ldots, n$,

$$x_i(t) > 0 \Longrightarrow (1-d)[(b_i/b_\sigma)p_\sigma(t+1) - p_i(t+1)] \leq [(b_i/b_\sigma)p_\sigma(t) - p_i(t)].$$

PROOF: We consider two cases: (i) $y_i(t) < x_i(t)$, and (ii) $y_i(t) = x_i(t)$.

Under case (i), let $\varepsilon = x_i(t) - y_i(t)$, $x = x(t) + \varepsilon e(i)$, $z = z(t+1)$, $y = y(t)$ and $x' = (1-d)x + z$. An appeal to Proposition 16 yields $(1-d)p_i(t+1) = p_i(t)$. Also, by Proposition 13, we have $(1-d)p_\sigma(t+1) \leq p_\sigma(t)$, and so

$$(b_i/b_\sigma)[(1-d)p_\sigma(t+1) - p_\sigma(t)] \leq 0 = [(1-d)p_i(t+1) - p_i(t)].$$

Next, consider case (ii) where $y_i(t) = x_i(t) > 0$. Let $0 < \varepsilon < x_i(t)$, $\nu = (b_i/b_\sigma)\varepsilon$, and note that from Assumption 1 that $(b_i/b_\sigma) \leq (a_i/a_\sigma) \leq 1$, which implies that $\nu \leq \varepsilon$. Define $x = x(t) - \varepsilon e(i) + \nu e(\sigma)$, $y = y(t) - \varepsilon e(i) + \nu e(\sigma)$, $z = z(t+1)$, $x' = (1-d)x + z$, and note that $ey \leq ey(t)$, $az = az(t+1)$ and $0 \leq y \leq x$. Therefore, $(x, x') \in \Omega$, and $y \in \Lambda(x, x')$. Then from Theorem 6, we obtain $p_\sigma(t+1)(1-d)\nu - p_i(t+1)(1-d)\varepsilon - p_\sigma(t)\nu + p_i(t)\varepsilon \leq 0$. This establishes the claim after transposing terms. □

PROOF OF THEOREM 7 Suppose that for some time-period T, and any $i = 1, \ldots, n$, $z_i(T+1) > 0$. Then by Proposition 18 and Assumption 1, we obtain:

$$p_i(T+1) \geq p_\sigma(T+1)(a_i/a_\sigma) > p_\sigma(T+1)(b_i/b_\sigma).$$

Also, we must have $x_i(t+1) > 0$ for all $t \geq T$. Then, iterating on the result presented as Proposition 19, we obtain $p_i(t+1) \to \infty$ as $t \to \infty$, a contradiction to Proposition 17. □

References

Atsumi, H. (1965), "Neoclassical Growth and the efficient program of capital accumulation," *Review of Economic Studies* **32,** 127–36.

Bardhan, P. K. (1971), "Optimum growth and allocation of foreign exchange," *Econometrica* **39,** 955–71.

Birner, J. (2002), *The Cambridge Controversies in Capital Theory*, London: Routledge.

Bliss, C. J. (1968), "On putty-clay," *Review of Economic Studies* **35,** 105–22.

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

Bruno, M. (1967), "Optimal accumulation in discrete capital models," K. Shell, ed., *Essays on the Theory of Optimal Economic Growth*, 181–218, Cambridge: MIT Press.

Cass, D., and J. E. Stiglitz (1969), "The implications of alternative savings and expectation hypotheses for choices of technique and patterns of growth," *Journal of Political Economy* **77,** 586–627.

Dixit, A. (1990), "Growth theory after thirty years," P. Diamond, ed., *Growth, Productivity and Unemployment*, Cambridge: The MIT Press.

Dobb, M. (1956), "Second thoughts on capital-intensity of investment," *Review of Economic Studies* **24,** 33–42.

Dobb, M. (1960), *An Essay on Economic Growth and Planning*, New York: Monthly Review Press.

Dobb, M. (1961), "Review of A. K. Sen's *Choice of Techniques*," *Kyklos* **14,** 1961.

Dobb, M. (1967), "The question of 'investment-priority for heavy industry'," *Capitalism, Development and Planning*, New York: International Publishers.

Dutta, P., and T. Mitra (1989), "Maximum theorems for convex structures with an application to the theory of optimal intertemporal allocation," *Journal of Mathematical Economics* **18,** 77–86.

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

Halevi, J. (1987), "Investment planning," J. Eatwell, P. K. Newman, and M. Milgate, eds, *The New Palgrave*, London: MacMillan.

Khan, M. Ali (2000), "Srinivasan on choice of technique again," Ranis-Raut, ed., *Trade, Growth and Development: Essays in Honor of Professor T. N. Srinivasan*, Chapter 2, Amsterdam: North Holland.

Khan, M. Ali, and T. Mitra (2002), "Optimal growth in the Robinson–Solow–Srinivasan model: The two-sector setting without discounting," mimeo, Cornell University, Ithaca, New York.

Khan, M. Ali, and T. Mitra (2003), "Optimal growth under irreversible investment: A continuous time analysis," mimeo, Cornell University, Ithaca, New York.

Koopmans, T. C. (1971), "A model of a continuing state with scarce capital," G. Bruckmann and W. Weber, eds, *Contributions to the von Neumann Growth Model*, 11–22, New York: Springer-Verlag.

Koopmans, T. C., and Hansen, T. (1972), "On the definition and computation of a capital stock invariant under optimization," *Journal of Economic Theory* **5,** 487–523.

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. 3, 1281–355, New York: North-Holland Publishing Company.

McKenzie, L. W. (1987), "Turnpike Theory," J. Eatwell, M. Milgate, and P. K. Newman, eds, *The New Palgrave*, vol. 4, 712–20, London: MacMillan.

Mirrlees, J. A. (1962), "Choice of techniques," *Indian Economic Review* **6,** 93–102.

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

Naqvi, K. A. (1963), "Machine tools and machines: A physical interpretation of marginal rate of saving," *Indian Economic Review* **7,** 19–28.

Okishio, N. (1966), "Technical choice under full employment in a socialist economy," *Economic Journal* **76,** 585–92.

Okishio, N. (1987), "Choice of technique and the rate of profit," J. Eatwell, P. K. Newman, and M. Milgate, eds, *The New Palgrave*, vol. 1, 418–21, London: MacMillan.

Raj, K. N., and A. K. Sen (1961), "Alternative patterns of growth under conditions of stagnant export earnings," *Oxford Economic Papers* **13,** 43–52.

Ramsey, F. (1928), "A mathematical theory of savings," *Economic Journal* **38,** 543–59.

Robinson, J. (1960), *Exercises in Economic Analysis*, London: MacMillan.

Robinson, J. (1969), "A model for accumulation proposed by J. E. Stiglitz," *Economic Journal* **79,** 412–13.

Sen, A. K. (1960), *Choice of Techniques*, Oxford: Basil Blackwell.

Sen, A. K. (1968), "Optimal savings, technical choice and the shadow price of labour," introduction to the third edition of Sen (1960), Oxford: Basil Blackwell.

Solow, R. M. (1962a), "Some problems in the theory and practice of economic planning," *Economic Development and Cultural Change* **10,** 216–22.

Solow, R. M. (1962b), "Substitution and fixed proportions in the theory of capital," *Review of Economic Studies* **29,** 207–18.

Solow, R. M. (2000), "Srinivasan on choice of technique," G. Ranis, and L. K. Raut, eds, *Trade, Growth and Development: Essays in Honor of Professor T. N. Srinivasan,* Chapter 1, Amsterdam: North Holland.

Srinivasan, T. N. (1962a), "Investment criteria and choice of techniques of production," *Yale Economic Essays* **1,** 58–115.

Srinivasan, T. N. (1962b), Review of Sen (1960) in *Journal of Political Economy* **70,** 306–7.

Stiglitz, J. E. (1968), "A note on technical choice under full employment in a socialist Economy," *Economic Journal* **78,** 603–9.

Stiglitz, J. E. (1970), "Reply to Mrs. Robinson on the choice of technique," *Economic Journal* **80,** 420–22.

Stiglitz, J. E. (1973), "Recurrence of techniques in a dynamic economy," J. Mirrlees and N. H. Stern, eds., *Models of Economic Growth*, 138–61, New York: John-Wiley and Sons.

Stiglitz, J. E. (1990), "Comments: Some retrospective views on growth theory," P. Diamond, ed., *Growth, Productivity and Unemployment*, 50–68, Cambridge: The MIT Press.

von Weizsäcker, C. C. (1965), "Existence of optimal programs of accumulation for an infinite time horizon," *Review of Economic Studies* **32,** 85–104.

von Neumann, J. (1935-36). "Über ein Ökonomisches Gleichungs-System und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes," K. Menger, ed., *Ergebnisse eines Mathematischen Kolloquiums*, No. 8. Translated as 'A model of general economic equilibrium," *Review of Economic Studies* **XIII,** 1–9.