Emerging Issues in Economic Development: A Contemporary Theoretical Perspective
Sugata Marjit and Meenakshi Rajeev

Print publication date: 2014
Print ISBN-13: 9780198099062
Published to Oxford Scholarship Online: January 2015
DOI: 10.1093/acprof:oso/9780198099062.001.0001

Inefficiency and the Golden Rule

Phelps–Koopmans Revisited

Tapan Mitra
Debraj Ray

DOI:10.1093/acprof:oso/9780198099062.003.0003

Abstract and Keywords

This chapter studies the celebrated Phelps–Koopmans theorem in environments with non-convex production technologies. The chapter argues that a robust failure of the theorem occurs in such environments. Specifically, it is proved that the Phelps–Koopmans theorem must fail whenever the net output of the aggregate production function f(x), given by f(x) – x, is increasing in any region between the golden rule and the maximum sustainable capital stock.

Keywords: capital over accumulation, inefficiency, Phelps–Koopmans theorem, non-convex production set

A planned path of consumptions is *efficient* if there is no other feasible planned path that generates just as much consumption at every date, with strictly more consumption at some date. This innocuous-looking definition contains one of the most interesting problems in classical growth theory: what criteria must one invoke to determine the efficiency of a planned path?

It isn't surprising that such a question was born in the early second half of the twentieth century. With the end of the Second World War, the newly-won independence of colonial nations, and the rising influence of socialist politics in Europe and elsewhere, planners and academics placed growing reliance on planned growth: on the deliberate allocation of resources, both across sectors and over time, to achieve economic development.

Of course, this relatively narrow definition of efficiency comes nowhere close to addressing the manifold complexities of such a **(p.44)** development. But it is a *necessary* requirement, and it is a fundamental consideration. It is also a subtle criterion, as the work of Edmond Malinvaud, Edmund Phelps, Tjalling Koopmans, David Cass, and others was to reveal. When time (and the number of commodities) is finite, efficiency is no more involved than an old-fashioned maximization problem; indeed, in the aggregative growth model with a single malleable commodity, efficiency is identical to the simple absence of waste. However, when the time-horizon is open-ended, and therefore infinite, new considerations appear. It is entirely possible for a path to not involve any waste at any particular point in time, and yet be inefficient.

It is in this context that the so-called *Phelps–Koopmans theorem* provides a celebrated necessary condition for efficiency. As a historical note, Phelps (1962) actually conjectured the necessity of the condition, while Koopmans proved that conjecture; the resulting theorem appears in Phelps (1965). The work merited a Nobel citation. In awarding the 2006 Prize to Edmund Phelps, the Royal Swedish Academy of Sciences observed that:

> Phelps…showed that all generations may, under certain conditions, gain from changes in the savings rate.

Briefly, the Phelps–Koopmans theorem lays the blame for inefficiency at the doorstep of capital *over*-accumulation. The extreme cases are easy enough: if *all* capital is forever accumulated, then the outcome must perforce be inefficient, and if all capital is instantly consumed in the first period, the outcome must be efficient (after all, all other paths must yield lower consumption in the first period).[1] But the theorem throws light on the intermediate cases as well. Define a *golden rule* capital stock to be one at which output net of capital is maximized.[2] The Phelps–Koopmans theorem states that: **(p.45)**

> If the capital stock of a path is above and bounded away from the golden rule stock, from a certain time onward, then the path is inefficient.

Later characterizations that seek a complete description of inefficiency (not just a sufficient condition), such as the work of Cass (1972), rely fundamentally on the Phelps–Koopmans insight. It is not our intention to survey the sizeable literature that works towards a complete characterization of efficiency. Rather we seek to investigate the original theorem in a more general context, one that allows for non-convexity of the production technology. The motivation behind such an investigation should be obvious. The vast bulk of literature assumes diminishing returns in production. This flies squarely in the face of empirical reality, in which minimum scales of operation (and the resulting non-convexities) are the rule rather than the exception. It is of some interest that this case has received little attention as far as the Phelps–Koopmans theorem is concerned. It is of even greater interest that without substantial qualification, the theorem actually fails to extend to this context.

To begin with, the setting is still the same: non-convexity is no impediment to the existence of a golden rule stock provided that suitable end-point conditions hold. Indeed, there may now be several such stocks; refer to the smallest of them as the *minimal* golden rule. Our recent paper (Mitra and Ray 2012) breaks up the Phelps–Koopmans assertion into three progressively stronger formats:

1. Every stationary path with capital stock in excess of the minimal golden rule is inefficient.
2. If a (possibly non-stationary) path converges to a limit capital stock in excess of the minimal golden rule, then it is inefficient.
3. If a (possibly non-stationary) path lies above, and bounded away from the minimal golden rule from a certain time onwards, then it is inefficient.

Obviously, version 3 nests version 2, which in turn nests version 1.

It is very easy to see that the weakest version, 1, of the Phelps–Koopmans theorem must be true. But version 2 of the theorem is false. In Mitra and Ray (2012), we present an example of an *efficient* **(p.46)** path that converges to a limit stock that exceeds the minimal golden rule. The circumstances under which version 2 is true is completely characterized in that paper—in terms of the curvature of the production function at the golden rules.

In short, while a variant of the Phelps–Koopmans theorem does hold when technology is non-convex, the 'over-accumulation of capital', as defined by Phelps, need not always imply inefficiency.

A corollary of our characterization is that version 2 is indeed true provided that the golden rule stock is *unique*. It turns out; however, that version 3 of the theorem is not true even if the golden rule is unique. Proposition 3 in Mitra and Ray (2012) provides a stringent condition on the production function under which version 3 is guaranteed to fail (see condition F.4 in that paper). However, the stringency of the condition precludes its necessity. Our paper is silent on the possibility of *completely* characterizing an economic environment for which version 3 stands or falls. Indeed, we ended our introduction to our paper thus:

> An interesting research question is to describe conditions under which version III is valid. We suspect that such conditions will involve strong restrictions on the production technology. Whether those conditions usefully expand the subset of convex technologies remains an open question.

The goal of the present chapter is to address this question. Under some mild restrictions on the allowable family of production technologies, we provide a complete characterization of what one might call the *Phelps–Koopmans property*, one that allows for all non-stationary paths, as in version 3. The property may be stated as:

> A path is inefficient if its capital stock sequence lies above and bounded away from the minimal golden rule capital stock from a certain time onwards.

We prove that the Phelps–Koopmans theorem must fail whenever the net output of the aggregate production function $f(x)$, given by $f(x) - x$, is increasing in any region between the golden rule and the maximum sustainable capital stock. As a corollary of our result, suppose that the production function $f(x)$ is continuously differentiable, with a strictly positive derivative and admits a unique **(p.47)** golden rule. Then if f is concave, the Phelps–Koopmans assertion holds, but if we perturb the function *ever so slightly* to the right of the golden rule (but below the maximum sustainable stock), so that it now admits a region over which $f'(x) > 1$, the Phelps–Koopmans property must fail. We return to this discussion after the statement of the main theorem.

Preliminaries

Consider an aggregative model of economic growth. At every date, capital x_t produces output $f(x_t)$, where $f : \mathbb{R}_+ \to \mathbb{R}_+$ is the production function. We assume throughout that:

> [F] The production function $f : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous and increasing on $\mathbb{R}_+$ with $f(0) = 0$, and there is $B \in (0, \infty)$ such that $f(x) > x$ for all $x \in (0, B)$ and $f(x) < x$ for all $x > B$. Further, the left hand derivative of f, denoted by f^-, exists and is positive for all $x > 0$.

We can think of B as the maximum sustainable stock.

Notice that [F] includes the standard convex technology, as well as technologies in which there are one or more regions of non-convexity. The somewhat awkward assumption that the left-hand derivative of f is always well-defined (but not necessarily the full derivative) allows us to accommodate cases in which f is the upper envelope of two or more neoclassical production functions, as described in Mitra and Ray (2012) and in the discussion later in this chapter.[3]

A *programme* from $\kappa > 0$ is a sequence of *capital stocks* $\mathbf{x} = \{x_t\}$ with:

$$x_0 = \kappa \text{ and } 0 \leq x_{t+1} \leq f(x_t)$$

for all $t \geq 0$. Let $c_{t+1} = f(x_t) - x_{t+1}$ be the associated consumption programme. With no real loss of generality, we presume that $\kappa \in [0, B]$.

A programme $\boldsymbol{x}'$ from κ *dominates* a programme $\boldsymbol{x}$ from κ if the associated consumption sequences satisfy:

$$c'_{t+1} - f(x'_t) - x'_{t+1} \geq f(x_t) - x_{t+1} = c_{t+1}$$

(p.48) for every t, with strict inequality for some t. A programme $\boldsymbol{x}$ from κ is *inefficient* if there is a programme $\boldsymbol{x}'$ from κ which dominates it. It is *efficient* if it is not inefficient.

Define $s(x) \equiv f(x) - x$ for all $x \geq 0$. Under [F], s is continuous on $[0, B]$ with $s(0) = s(B) = 0 \geq s(x)$ for all $x \geq B$, so there is $x^* \in (0, B)$ such that:

$$s(x^*) \geq s(x) \text{ for all } x \geq 0.$$

Call x^* a *golden rule* stock, or simply a *golden rule*. Clearly, the set of golden rules lies in $(0, B)$ and is compact, so there is a smallest or *minimal* golden rule; denote it by k.

The Main Theorem

THEOREM 1 The Phelps–Koopmans property holds if and only if:

(3.1)

$$s(x) \textit{ is non-increasing for all } x \in [k, B],$$

where k is the minimal golden rule.

Proof [If] Suppose $\{x_t\}$ is a programme from $\kappa \geq 0$, and there is $\alpha > 0$ and $T \in \mathbb{N}$ such that $x_t \geq k + \alpha$ for all $t \geq T$. Define $\{x'_t\}$ by $x'_t = x_t$ for $t = 0,\ldots, T-1$, and $x'(t) = x(t)$ for $t \geq T$. Then $x'_t = x_t \geq 0$ for all $t \in \{0,\ldots, T-1\}$, and $x'_t \geq k$ for all $t \geq T$. Further, $c'_t = c_t$ for $t \in \{1,\ldots, T-1\}$ if any, and $c'_T = c_T + \alpha > c_T$. For $t \geq T$, we have:

$$\begin{aligned} c'_{t+1} = \; & f(x'_t) - x'_{t+1} = f(x_t - \alpha) - f(x_t) + f(x_t) - x_{t+1} + \alpha \\ = \; & c_{t+1} + f(x_t - \alpha) - f(x_t) + \alpha \\ = \; & c_{t+1} + [f(x_t - \alpha) - (x_t - \alpha)] - [f(x_t) - x_t] \\ \geq \; & c_{t+1}, \end{aligned}$$

by virtue of the fact that (3.1) holds. Thus, $\{x_t\}$ is inefficient.[4]

(p.49) [Only If] Suppose that (3.1) is violated. Then, there exist numbers b and b' such that $b' > b > k$ and $s(b') > s(b)$.[5] Furthermore, there is $w \in (b, b')$ such that $\eta \equiv f^{-}(w)$ exceeds 1.[6] Since, η is the left-hand derivative of f at w, we can find $0 < e < w - b < w - k$ such that whenever $x \in (w - e, w)$, we have:

$$\left| \frac{f(w) - f(x)}{w - x} - \eta \right| < (\eta - 1)/2.$$

In particular, for all $x \in (w - e, w)$:

$$\frac{f(w)-f(x)}{w-x} \geq \eta - [(\eta-1)/2] = [(\eta+1)/2] \equiv h > 1,$$

so that:
(3.2)

$$f(w)-f(x) \geq h(w-x) \text{ for all } x \in (w-e,\ w].$$

Next, pick $z \in (k, w - e)$, with z sufficiently close to k, so that if we define $y \equiv z - e$, then:

(3.3)

$$f(x)-x < f(z)-z \ for\ all\ x \leq y.$$

To see that this can be done, suppose by way of contradiction that such a construction is impossible. Then there exists a sequence $z^n \downarrow k$ such that for every n, there is $x^n \leq z^n - e$ with $f(x^n) - x^n \geq f(z^n) - z^n$. By passing to the limit (and taking a subsequence of $\{x^n\}$ if necessary), we contradict the fact that k is the minimal golden rule.

To complete the preliminaries, define:

$$\delta \equiv [f(z)-z] - \max_{x \leq y}[f(x)-x],$$

(p.50) and note that δ must be strictly positive. Choose a positive integer M such that:
(3.4)

$$M\delta > k.$$

Now define a cyclical programme $\boldsymbol{x}$ as follows. The programme starts at z and stays there for M periods, where M is defined by (3.4). The programme then steadily accumulates to reach w (to be concrete, think of this as pure accumulation with some adjustment in consumption in at most one period so as to hit w exactly). Say this takes N periods, thereby passing through $(N + 1)$ distinct values of capital.

To describe the remainder of the programme, we need some more notation. Denote the distinct values of the stock by $(z_0, z_1,\ldots, z_N)$, with $z_0 = z$, and $z_N = w$. The left-hand derivative of f exists and is positive at each of these points, and so $\mu \equiv \min\{f^-(z_0),\ldots, f^-(z_N)\}$ is strictly positive. For each $j \in \{0, 1,\ldots, N\}$, there is $0 < \theta_j < z$ such that for all $x \in (z_j - \theta_j, z_j)$:

$$\left|\frac{f(z_j)-f(x)}{z_j-x} - f^-(z_j)\right| < \mu/2.$$

so that for all $x \in (z_j - \theta_j, z_j)$:

$$\frac{f(z_j)-f(x)}{z_j-x} \geq f^-(z_j) - \mu/2 \geq \mu/2$$

and therefore:

$$f(z_j) - f(x) \geq (\mu/2)(z_j - x) \text{ for all } x \in (z_j - \theta_j, z_j).$$

Define $\theta \equiv \min\{\theta_0,\ldots,\theta_N\}$. Then, for all $j \in \{0, 1,\ldots, N\}$:

(3.5)

$$f(z_j) - f(x) \geq (\mu/2)(z_j - x) \text{ for all } x \in (z_j - \theta, z_j)$$

Consider the function $L(x) \equiv f(x) - f(x - \theta)$ for all $x \in [z, B]$. Then $L(x)$ is a positive continuous function on $[z, B]$, and has a minimum **(p.51)** value, which we call $\overline{L}$. Define $\beta = \overline{L}/B$, and $\ell = \min\{\beta,(\mu/2)\}$. Note that $\ell > 0$. We now claim that for each $j \in \{0, 1,\ldots, N\}$,

(3.6)

$$f(z_j) - f(x) \geq \ell(z_j - x) \text{ for all } x \in [0, z_j]$$

For $x = z_j$, this is trivially true. So, consider $x \in [0, z_j)$. Either (i) $x \in (z_j - \theta, z_j)$, or (ii) $0 \leq x \leq z_j - \theta$. In case (i), using (3.5) we have $[f(z_j) - f(x)] \geq (\mu/2)(z_j - x) \geq \ell(z_j - x)$. In case (ii), we have:

$$\begin{aligned} f(z_j) - f(x) &\geq f(z_j) - f(z_j - \theta) \\ &\geq \overline{L} = \beta B \geq \beta(z_j - x) \geq \ell(z_j - x) \end{aligned}$$

This establishes our claim (3.6).

Choose a positive integer Q so that:

(3.7)

$$h^Q \ell^{M+N} \equiv \lambda > 1.$$

Because $h > 1$ and $\ell > 0$, Q can always be chosen to satisfy (3.7).

We now complete the description of the programme. After accumulating to w, it stays there for Q periods, and then returns to z, whereupon the cycle is indefinitely repeated.

We claim that $\boldsymbol{x}$ is efficient.

If not, there is a dominating programme $\boldsymbol{x}'$. Let $t(i)$ be the date at which a fresh round i starts. We claim that there is i such that $\varepsilon_i \equiv x_{t(i)} - x'_{t(i)} \geq e$.

To establish the claim, notice that at every date t:

$$f(x'_t) - x'_{t+1} \geq f(x_t) - x_{t+1},$$

so that:

(3.8)

$$x_{t+1} - x'_{t+1} \geq f(x_t) - f(x'_t) \text{ for all } t \geq 0.$$

Since $x_0 = x'_0$, we have $x_t \geq x'_t$ for all $t \geq 0$ by using (3.8).

(p.52) Now, during the first $M + N$ periods of any round, we know from (3.6) that:

$$f(x_t) - f(x'_t) \geq \ell [x_t - x'_t],$$

so that combining this information with (3.8), we see that during the first $M + N$ periods of any round:

(3.9)

$$x_{t+1} - x'_{t+1} \geq \ell [x_t - x'_t].$$

Let us refer to the phase, in which the stock is kept stationary at w in the original programme, as the 'upper phase'. There are two possibilities to consider: (i) there is some date t in the upper phase, in some round, for which $x'_t \leq w - e$, and (ii) $x'_t > w - e$ at every date in the upper phase for every round.

In case (i), using (3.8) again:

$$x_{t+1} - x'_{t+1} \geq f(x_t) - f(x'_t) \geq f(w) - f(w - e) \geq e,$$

because $f(x) - x$ is increasing over the range $[w - e, w]$. Thus, $x'_{t+1} \leq w - e$, and this step can be repeated for all subsequent dates of the upper phase of that round to obtain $x_\tau - x'_\tau \geq e$ where τ is the first period of the next round. This establishes our claim in case (i).

In case (ii), using (3.8) yet again, along with (3.2):

(3.10)

$$x_{t+1} - x'_{t+1} \geq h [x_t - x'_t]$$

for every date t of the upper phase of every round. Combining (3.9) and (3.10), we must conclude that:

$$\varepsilon_{i+1} \geq \ell^{M+N} h^Q \varepsilon_i = \lambda \varepsilon_i,$$

where $\lambda > 1$; see (3.7). So ε_i expands geometrically across rounds once it turns positive. But it must turn positive at some round, because $\mathbf{x'}$ is a dominating programme.

So in this case too, the claim is proved.

Consider any round i, then, at which $\varepsilon_i \geq e$. For the next M periods, we have $x'_t < y$ by induction, using (3.3), and moreover: **(p.53)**

$$\begin{aligned} x'_{t+1} &\leq f(x'_t) - [f(z) - z] \\ &\leq [f(x'_t) - x'_t] - [f(z) - z] + x'_t \\ &\leq \max_{x \leq y} [f(x) - x] - [f(z) - z] + x'_t \\ &\leq x'_t - \delta \end{aligned}$$

But, this means that by the end of M more periods, we must have $x'_t < y - M\delta < k - M\delta < 0$, a contradiction. So no dominating programme can exist, and $\boldsymbol{x}$ must be efficient.

Discussion

We can illustrate the main theorem as follows. Call a production function F *neoclassical* if it satisfies condition [F] and in addition:

[C] F is strictly concave on $\mathbb{R}_+$ and twice continuously differentiable on $\mathbb{R}_{++}$, with $F''(x) < 0$ for all $x > 0$.

Consider a production function f that can be written as the point-wise maximum of two neoclassical functions; call them h and g:

$$f(x) = \max\{h(x), g(x)\} \text{ for all } x \geq 0.$$

Suppose that first h, then g, occupies the envelope, that is, there exists u such that:

$$f(x) = h(x) > g(x) \text{ for } x < u.$$
$$f(x) = g(x) > h(x) \text{ for } x > u.$$

Suppose, moreover, that:

$$f(u) = h(u) = g(u) > u.$$

Then f satisfies [F].

Denote the (unique) golden rule of h by k_h, and the golden rule of g by k_g. Assume that these two values lie on either side of u:

$$k_h < u < k_g.$$

(p.54) Clearly, the technology set defined by f is non-convex (note that f is not, in general, differentiable). Observe that the minimal golden rule of f is k_h if:

(3.11)

$$h(k_h) - k_h \geq g(k_g) - k_g$$

with k_h the unique golden rule of f if and only if strict inequality holds in (3.11). Similarly, the minimal (and only) golden rule of f is k_g if (3.11) fails. In the latter case, since f is concave on [k_g, B], the standard Phelps–Koopmans theory applies to paths which are above and bounded away from the golden rule stock k_g. So the Phelps–Koopmans property clearly holds in this case.

Our main theorem implies, however, that this is the *only* situation in which the Phelps–Koopmans theorem is valid. As soon as (3.11) holds, the Phelps–Koopmans theorem fails, no matter how briefly g occupies the outer envelope that comprises f. This is a remarkable fact that requires some explanation. Figure 3.1 illustrates a situation in which g occupies the frontier for a relatively short stretch. The proof of the theorem constructs a cycle that starts from a point z close to the minimal golden rule k_h, stays there for M periods, and then goes up into the zone where g occupies the envelope, but to the left of k_g; see the point w in Figure 3.1. The programme stays there for Q periods, and then drops back to z again, whereupon the cycle starts all over again.

A comparison programme that dominates this cycle must ultimately have lower stocks relative to the cycle at every date. If the cycle reaches its peak in a zone where h still occupies the envelope, this is not a problem; indeed, there is some surplus to be gained by lowering stocks by a tiny amount. However, if the peak is reached when g occupies the envelope, control over the comparison stocks is weaker: in the region in which w is being repeated for Q periods, the comparison programme must steadily drift further away from w, because the surplus $s(x)$ falls locally to the left of w. This drift magnifies over rounds until no matter how alike the comparison programme was to start with, the difference in the stocks is pronounced. At this point, the comparison programme drops below a point such as y to the *left* of the golden rule k_h, and generates *lower* surplus than z. Once here, though, it cannot recover. The M subsequent **(p.55)** repetitions of z force the comparison programme to fall ever lower in stocks, until feasibility is violated. So the cycle is efficient. It is in this way that a tiny 'intrusion' of the second neoclassical technique g destroys the Phelps–Koopmans property.

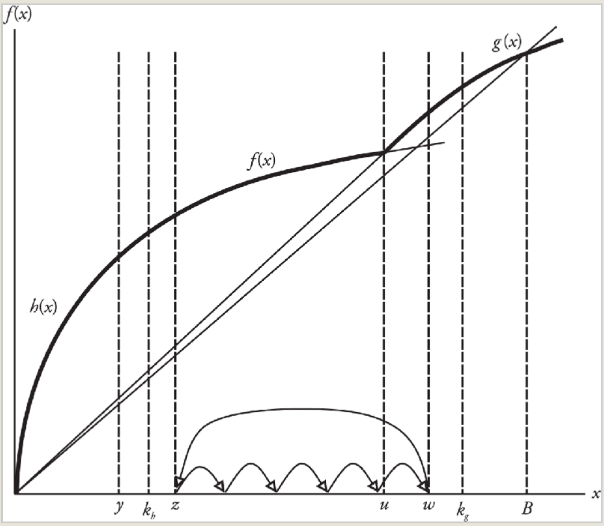

Figure 3.1 An Illustration of Theorem 1

* * *

This paper studies the well-known Phelps–Koopmans theorem in an environment with a non-convex production technology. We argue that in such a setting, the Phelps–Koopmans result generally fails to hold, and that this failure is quite robust. Specifically, we prove that the theorem fails whenever the net output of the aggregate production function $f(x)$, given by $f(x) - x$, is increasing in *any* region between the golden rule and the maximum sustainable capital stock. That is, in such cases, it is always possible to find an *efficient* program in which capital stocks ultimately lie above (and stay bounded away from) the golden rule. Thus the 'overaccumulation of capital', as captured by **(p.56)** a positive excess of capital over the golden rule, is no longer related to efficiency, once the convexity of the production technology is dispensed with.

References

Bibliography references:

Cass, D. 1972. 'On Capital Overaccumulation in the Aggregative Neo-classical Model of Economic Growth: A Complete Characterization', *Journal of Economic Theory*, 4: 200–23.

Mitra, T., and D. Ray. 2012. 'On the Phelps–Koopmans Theorem', *Journal of Economic Theory*, 147: 833–49.

Phelps, E.S. 1962. 'The End of the Golden Age in Solovia: Comment', *American Economic Review*, 52: 1097–9.

————. 1965. 'Second Essay on the Golden Rule of Accumulation', *American Economic Review*, 55: 793–814.

Ray, D. 2010. 'The Phelps–Koopmans Theorem and Potential Optimality', *International Journal of Economic Theory*, 6: 11–28.

Royden, H.L. 1988. *Real Analysis*. Third Edition. New York: Macmillan.

Notes:

([1]) The latter example underscores the fact that efficiency is a weak requirement: stronger optimality criteria such as the maximization of time-separable utility would generally rule out such paths. See Ray (2010) for a discussion of this case.

([2]) With a standard production function $f(k)$ satisfying the usual curvature and end-point restrictions, such a stock is characterized by the condition $f'(k) = 1$.

([3]) We conjecture that the differentiability restriction on f can be dropped at no cost.

([4]) The 'if' part is standard, and is inspired by the original proof of the Phelps–Koopmans theorem, as suggested by Koopmans. We include it here for a self-contained treatment.

([5]) That $b > k$ is guaranteed by virtue of the fact that k is a golden rule.

(6) Suppose, on the contrary, that $f^{-}(x) \leq 1$ for all $x \in (b, b')$. Define $S(x) = -s(x)$ for all $x \in I \equiv [b, b']$. Then, S is continuous on I, and its left hand derivative is non-negative for all $x \in (b, b')$. By Proposition 2 of Royden (1988: 99), we must then have $S(b') \geq S(b)$, so that $s(b') \leq s(b)$, a contradiction.