Optimal Growth in the Robinson-Shinkai-Leontief Model: The Case of Capital-Intensive Consumption Goods*

Liuchun Deng,† Minako Fujio,‡, and M. Ali Khan§

October 11, 2018

Abstract: We study the two-sector Robinson-Shinkai-Leontief (RSL) model of discrete-time optimal economic growth for the case of capital-intensive consumption goods. We frame the model in the context of Nishimura's *oeuvre,* and more specifically, relate it to its neoclassical cousin: the Uzawa-Srinivasan continuous-time version studied in Haque (1970). We take Fujio's (2006) identification of the marginal rate of transformation ζ of capital goods today into capital goods tomorrow, and using a mix of the Bellman-Blackwell methods of dynamic programming and the value-loss approach of Brock-Mitra, pin down the optimal policy for a specific subset of the parameter space. We discern a bifurcation pattern with respect to the discount factor that echoes the results of Benhabib-Nishimura (1979) for the neoclassical model, and Khan-Mitra (2005) for the RSS model. We also study the optimal policy correspondence for a general parameter set.

(140 words)

Journal of Economic Literature Classification Numbers: C60, D90, O21.

Key Words: RSL model, RSS model, optimal policy correspondence, marginal rate of transformation, golden-rule stock, convergence, 2-cycles, value function, value-loss

Running Title: Optimal Growth in the RSL-model

*An earlier version of the results reported here was presented in the special session in honor of Professor Kazuo Nishimura at the 26th Annual Symposium of the *Society for Nonlinear Dynamics and Econometrics* in Keio University on March 19, 2018. The authors gratefully acknowledge the encouragement and stimulus into our thinking provided by Jess Benhabib, Hülya Eraslan, Tapan Mitra, Kazuo Nishimura and Makoto Yano.

†Halle Institute for Economic Research (IWH) and Friedrich Schiller University Jena, Germany.

‡Department of Economics, Yokohama National University, Japan

§Department of Economics, The Johns Hopkins University, Baltimore, MD 21218.

Contents

> My first objective is to describe in plain English how the model works, because I think it is an interesting extension of macroeconomics. It seems paradoxical to me that such an important characteristic of the equilibrium path should depend on such a casual property of the technology.[1] Solow (1962)

> I am going to proceed on the assumption that the desirability of a path depends only on the consumption that it gives at every instant of time. Therefore, to steer the economy, the government need only control the allocation of output between consumption and investment. (In a more detailed model, it would have to look after the allocation of investment between the consumption goods industry and the capital-goods industry.)[2] Solow (1970)

> Another outgrowth of the Ramsey work is the development of descriptive models of economic growth where individuals take into consideration explicitly intertemporal considerations. [T]he Ramsey model can be thought of as how a competitive Robinson-Crusoe economy, or a competitive economy with identical individuals living infinitely long, might behave.[3] Stiglitz-Uzawa (1969)

1 Introduction and Antecedent Literature

Among his influential contributions to the theory of optimal economic growth, Kazuo Nishimura has always reserved a pride of place for the two-sector model – a model to which Japanese economists from Shinkai, Uzawa, Morishima, Inada, Amano and Takayama downwards have contributed so much, thereby creating an immensely rich tradition in economic theory. It is only fitting that he does so, and through a substantial body of work, rightfully takes his own place in this distinguished line of authorship.[4]

[1]The two sentences are taken from the first two paragraphs of Solow's 1962 11-paragraphed comment on Uzawa's two-sector model of economic growth, reprinted as Chapter 24 in Stiglitz-Uzawa [45]. It is one of the two papers that constitute their Part IV titled *Two-Sector Models* and introduced with the assertion that it, along with Part V titled *Foundations of Optimal Economics,* "represent perhaps the two fastest growing areas of research in growth theory."

[2]See Chapter 5 of Solow's 1970 exposition, reprinted in [44], and commented on in [40, p. xii]. It testifies to the somewhat conflicted relationship of Solow to the two-sector model: his enthusiastic 1962 reception, its bracketing in 1970, and its complete elimination in the survey [43] and the updated [44].

[3]See the introduction to Part V of the Stiglitz-Uzawa [45] anthology. We remind the reader that it consists of only four papers: in addition to Ramsey's 1928 and Samuelson's 1958 classics, Malinvaud's 1961 analogy between atemporal and intertemporal resource allocation and Solow's 1962 comment on the golden rule. All, still eminently worth reading, testify to the profession's understanding of optimal growth theory at that time. The omission of Uzawa's 1964 paper on optimal growth is especially interesting from the viewpoint of the results reported in this paper.

[4]To be sure, the literature does not have an exclusive Japanese accent: as is well-acknowledged, important contributions have been made by Joan Robinson, James Meade, Ronald Jones, Mordecai Kurz, TN Srinivasan, John Chipman, Joseph Stiglitz and other non-Japanese economists. Once trade theory and political economy is admitted into the picture, the literature simply mushrooms; see Magee [27], Oswald [36] and Magee-Brock-Young [28], with the bibliographies of the first two references consisting

In terms of Nishimura's *oeuvre,* already in his 1979 paper on the application of the Hopf bifurcation to optimal multi-sectoral growth theory, he and Jess Benhabib remind the reader how their basic equations draw on the "duality between the Rybczyski and Stolper-Samuelson effects well known in trade theory."[5] Six years later, in their 1985 analysis of competitive equilibrium cycles, the principal theorems on optimal oscillatory cycles are set in the context of the two-sector model. In what is referred to as a setting where "production is characterized by a two-sector neoclassical nonjoint technology," Nishimura and Benhabib write

> Persistent cycles require a further restriction of the capital-intensity differences between consumption and investment goods. For cycles to be sustained, the oscillations in relative prices must not present intertemporal arbitrage opportunities. Thus possible gains from postponing consumption from periods when the marginal rate of transformation between consumption and investment is high to periods when it is low must not be worth it. Whether this is the case or not depends on the discount rate as well as the slope of the production possibility frontier.

Four years later, in their 1989 work on stochastic oscillatory cycles, Benhahib-Nishimura carry forward their representation of the production-possibility curve in [4, Figure 1] to two "suggestive" figures in a reserved section on diagrammatic exposition.[6]

However, it is in the late-eighties that the Li-Yorke theorem and topologically chaotic trajectories of perfect-foresight paths enter the picture, and the two-sector model is used in a different exemplary way. Writing in 1993, Nishimura-Yano note that

> two types of models are known to have chaotic optimal paths. The examples of Deneckere-Pelikan (1986), Boldrin-Montruccchio (1986) and Boldrin-Deneckere(1990) are based on two-sector models, while that of Majumdar-Mitra (1994) are based on the shape of the utility function.[7]

In what can be seen as the *piece de resistance,* Nishimura-Yano (1995) conclusively

> demonstrate the possibility of ergodically chaotic optimal accumulation in the case [of] a two-sector model with Leontief production functions. [They] construct a condition under which the optimal transition function is unimodal and expansive

of 162 and 166 items respectively, and that of the third, running to 49 pages. For a *rapprochement* of trade and growth theory, see Fujio-Khan [14]; and for a synthesis stemming from the Oniki-Uzawa model, Chen-Nishimura-Shimomura [7].

[5] See Chapter 9 reprinted in [3], and Figure 1 below. Also see Chapter 13 of [3], and the recurrence of these theorems in [5, Section 6.2]. Even though our emphasis in this work is on optimal growth in the two-sector model, we refer the reader to [6, Chapters 4-5] and [9, Chapter 6] for equilibrium growth. For the impact of these classical theorems on modern political economy, see [28, subject index].

[6] See Figures 13.1 and 13.2 in Section 2 in Benhabib-Nishimura (1989) reprinted in [3, pp. 291-293].

[7] See Footnote 2 in of Nishimura-Yano (1993) reprinted in [29, p. 149]. To save on references, we send the reader to the bibliography of the anthology.

[and show] that the set of parameter values satisfying that condition is nonempty no matter how weakly the future utilities are discounted.[8]

The direction that this (our) paper takes is, however, best summarized in the words of Boldrin-Deneckere (1990).

> What distinguishes the current paper from the [earlier] research is that we do not construct "artificial" economies that exhibit a pre-chosen dynamics in equilibrium. Rather, we start with a specification of technology and preferences, and *derive* the implied dynamics. While this was also the object of Boldrin's (1989) study, the focus there is on providing criteria for the existence of chaotic paths in abstract two-sector models [see also Deneckere (1988)]. Here we specify two analytically tractable functional forms for the production technologies, and carry out a complete parametric analysis of the resulting dynamics.[9]

This change in viewpoint emphasizing the desirability of a full working-out of a specific model is important and in keeping with May's 1976 [30] emphasis on showing complicated dynamics from simple models. It has been carried forward in the work of Khan-Mitra and Fujio.[10] Indeed, various permutations of the specific versions of the two-sector model are available in the antecedent literature: Leontief-Leontief, Cobb-Douglas-Cobb-Douglas, and Leontief-Cobb-Douglas. In this paper, we name and study the two-sector version of the Robinson-Shinkai-Leontief RSL model.

Leontief's name is pervasive in the work on the two-sector model that we refer to in the paragraphs above, but why Shinkai? why Robinson? As far as the first is concerned, in a 1969 book-chapter titled "A neo-classical passage to growth equilibrium" Morishima [33] already refers to the "Shinkai-Uzawa finding" and to the "Shinkai-Uzawa capital intensity condition." He writes:

> In the conventional discussion of the stability of growth equilibrium in a two-sector economy, the relative capital intensities of the two industries have served as a kind of litmus paper with which to test whether the Silvery Equilibrium is stable or not. Shinkai observed, for the first time, that the growth equilibrium is stable if and only if the consumption-good industry is more capital-intensive than the capital-good industry."[11]

---

[8] For other examples concerning both topological and ergodic chaos based on the two-sector model with Leontief production functions, (what we are calling the two-sector RSL model here), see Chapters 4, 8 and 9 in [29].

[9] See the fourth paragraph of Boldrin-Deneckere (1990) reprinted in [3].

[10] For the first, see [17, 22] and their references; and for the second, see [10, 11, 12, 13].

[11] See [33, p. 45], and his book more generally for the definitions of his terms; also [38, pp. 51-63], and his Lectures 1 and 2 in [32]. In [33], Morishima continues, "Uzawa replaced Shinkai's production functions of the Leontief type by the neo-classical ones which allow continuous substitution between labour and capital, to find that the relative capital-intensity criterion is a sufficient condition for stability but no longer a necessary condition — though Furuno later saw that the Shinkai-Uzawa finding should be subject to a proviso that the introduction of a production lag narrows the stability region."

This is the reason that Fujio in her 2006 dissertation, and in subsequent work,[12] refers to the model as the Leontief-Shinkai (LS) model. However, what seemed to have been missed is Robinson's 1956 contribution. Already, in his 1960 article, Shinkai added a "historical note."

> As the reader will find, our model has much in common with Mrs. Robinson's "simple model" of capital accumulation and Professor Morishima's interpretation of her model, and our equilibrium growth corresponds to her "golden age" (under the hypothesis of no technical progress).[13]

The presentation of the two-sector technology in Robinson's 1956 book can be found in the extensive reviews of that work, the clearest being Barna's [2], Lancaster's [24] and Worswick's [46].[14] However, it is important to bear in mind that we include her name in the naming of both the RSS and RSL models, primarily because of the technological specifications that can be found in her 1956 work, and not simply to use her name in an empty valorization of our enquiry. It is clear that her interests were not in optimal and equilibrium growth, as we take the terms to mean here, but rather in the functioning of a capitalist economy in the macroeconomic short-run and in its leading to a disequilibrium growth process in the long-run in which innovation, gestation lags and expectations play a crucial role.[15]

But this focus on the two-sector technology for the justification of the inclusion of Robinson in the naming of the RSL model blurs the distinction between descriptive and optimal growth. In terms of optimal growth in the smooth neoclassical setting in continuous time, the state of the art seems to be Haque's 1970 [16] analysis.

> In his pioneering essay [3], Srinivasan studied the problem of optimal growth in the two-sector model under the assumption that the capital-intensity k_C in the consumption sector is larger at every wage-rental ratio than the capital intensity k_I in the investment sector. He proved the following interesting theorem: If the initial capital-labour ratio is smaller (larger) than a certain critical ratio, then along the optimal path specialization in investment (consumption) takes place until the critical ratio is attained and then golden age is approached through non-specialization. The optimal path has the special [Srinivasan] property that once specialization is discarded in favour of non-specialization, the former does not re-appear.

[12] See the references to Fujio in Footnote 10.

[13] There are only five references in [42]: one of these is to Robinson's 1956 book, and the other to a 1958 unpublished note of Morishima.

[14] See [pp. 46-47][38] and [40], and in addition to Robinson herself, see [47], [15] and [26].

[15] See, for example, her evaluation of Meade's book on two-sector growth in the Conclusion in [40], the footnote to Uzawa in [40, p. 132], more generally, her reminiscences in [41], and the Harcourt-Kerr introduction to the second edition of [37]. Also see [39] and her exchange with Stiglitz referred to in Khan-Mitra [17].

It is this factor-intensity assumption, and the Srinivasan property drawing on it, that will constitute the background subtext to the results reported in the sequel.

So much for the framing of the two-sector RSL model in terms of the antecedent literature. The question then arises as to how we contribute substantively and technically to economic growth and to optimal economic dynamics through it. This can best be answered in terms of the plan of the remaining part of the paper. Section 2 presents the basics of the model. It reformulates the model in terms of the Gale-McKenzie reduced form articulated in terms of capital stocks today, and of those tomorrow, and identifies the marginal rate of transformation ζ between them as a crucial sufficiency statistic. This marginal rate had not been identified earlier.[16] Section 3 uses the methods of discounted dynamic programming and offers the preliminary characterizations of the optimal policy correspondence. It is by now well-understood that methods of dynamic programming, while useful for characterization of the optimal trajectory, do not enable the path itself to be fully characterized: one can only guess and verify. Section 4 uses the value-loss approach of Brock-Mitra to pin down the optimal trajectory for the case $\zeta \leq 1$. For the case $\zeta = 1$, we obtain two-period cycles, whereas for the case, $\zeta < 1$, we obtain convergence to the golden-rule stock. This is of consequence in that, in contrast to the antecedent literature, it is the case of $\zeta > 1$ that is responsible for complicated and chaotic dynamics. We collect our findings and conclude the paper in Section 5. The detailed proofs are relegated as supplementary material collected in an Appendix in Section 6.

2 The RSL Model of Choice of Technique

Even if we forego short- and intermediate-term macroeconomics, and focus on growth and development, Joan Robinson can be seen to be motivated by the question of the "choice of technique." This was also the primary initial motivation of Khan-Mitra [17]. And so it stands to reason that we do not preclude the RSL model as not being able to say anything regarding this question: it ought to be, as the RSS model, an important special case of multi-sectoral capital theory. In this paper, however, we abstain from these considerations and limit ourselves to the two-sector variant.

2.1 The two-sector RSL Model

We consider the two-sector Robinson-Shinkai-Leontief model of optimal economic growth which employs the precise technological specification used in Nishimura-Yano ([34, 35]), Fujio ([11, 12, 10, 13]) and the literature surveyed in Table 1. There are two production sectors. One unit of consumption good is produced by one unit of labor and a_C units

[16] See Table 1. This sufficiency statistic is due to Fujio [11] for the case of undiscounted optimal growth, and builds on earlier work of Khan-Mitra [17] for the RSS model. Also see [31] and his references for the Gale-McKenzie reduced form.

of capital, and b units of investment goods are produced by one unit of labor and a_I units of capital. We only consider the case that the consumption good sector is more capital-intensive than the investment good sector, that is,[17]

$$a_C > a_I. \tag{1}$$

Capital cannot be consumed and depreciates at the rate $d \in (0, 1)$. The case of durable capital nests the RSS model and the Nishimura-Yano model as special or limiting case. This consideration makes the dynamic optimization problem doubly difficult. The amount of capital available at the beginning of next period is equal to the sum of the current production of investment goods and the left-over capital after depreciation. A constant amount of labor, normalized to be unity, is available in each time period $t \in \mathbb{N}$, where $\mathbb{N}$ is the set of non-negative integers. Thus, in the canonical formulation surveyed in McKenzie (1986, 2002), the collection of production plans (x, x'), the amount x' of capital in the next period (tomorrow) from the amount x of capital available in the current period (today), is given by the *transition possibility set*

$$\Omega = \{(x, x') \in \mathbb{R}_+ \times \mathbb{R}_+ : x' - (1-d)x \geq 0, x' - (1-d)x \leq b \min\{1, x/a_I\}\},$$

where $\mathbb{R}_+$ is the set of non-negative real numbers. For any $(x, x') \in \Omega$, one can consider the amount y of production of consumption goods, leading to a correspondence $\Lambda : \Omega \longrightarrow \mathbb{R}_+$ with

$$\begin{aligned}\Lambda(x, x') = \{y \in \mathbb{R}_+ : 0 \leq y \leq (1/a_C)(x - (a_I/b)(x' - (1-d)x))& \\ \text{and } \ 0 \leq y \leq 1 - (1/b)(x' - (1-d)x)\}.&\end{aligned}$$

A felicity function, $w : \mathbb{R}_+ \longrightarrow \mathbb{R}$, is linear and normalized to be equal to the amount of consumption, i.e., $w(y) = y$. Finally, the *reduced form utility function,* $u : \Omega \longrightarrow \mathbb{R}_+$, is defined on Ω such that

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

We assume that the future welfare is discounted with a discount factor $\rho \in (0, 1)$.[18]

An *economy* E consists of a triple (Ω, u, ρ) and the following concepts apply to it. A *program from* x_0 is a sequence $\{x_t, y_t\}$ such that for all $t \in \mathbb{N}$, $(x_t, x_{t+1}) \in \Omega$ and $y_t = \max \Lambda(x_t, x_{t+1})$. A program $\{x_t, y_t\}$ is called *stationary* if for all $t \in \mathbb{N}, (x_t, y_t) = (x_{t+1}, y_{t+1})$. For all $0 < \rho < 1$, a program $\{x_t^*, y_t^*\}$ from x_0 is said to be *optimal* if

$$\sum_{t=0}^{\infty} \rho^t [u(x_t, x_{t+1}) - u(x_t^*, x_{t+1}^*)] \leq 0$$

for every program $\{x_t, y_t\}$ from x_0.

[17] The same assumption is imposed in [34] and [11]. We pursue the case for $a_C \leq a_I$ in a different paper.

[18] For a comprehensive treatment of the undiscounted case ($\rho = 1$), see [12].

2.2 The Modified Golden Rule

Consider the today-tomorrow diagram furnished as Figure 2. Note that the transition possibility set Ω and the reduced-form felicity function $u(\cdot,\cdot)$ have nothing to do with a discount factor. So do the indifference maps. The transition possibility set, Ω, is a set $LVOD$ and the indifference maps are kinked lines along with MV line; when it is identical to LVO its utility level is zero.

The modified golden stock $\hat{x}$ is shown as a solution to the following problem;

$$u(\hat{x},\hat{x}) \geq u(x,x') \text{ for all } (x,x') \in \Omega \text{ such that } x \leq (1-\rho)\hat{x} + \rho x'. \tag{2}$$

. The equation of the line RG is $x' = \rho^{-1}x + \rho^{-1}(\rho - 1)\hat{x}$ specifies a constraint set. It is easily seen that the modified golden-rule stock is exactly equal to the golden-rule stock. It is due to the kinked indifference maps that the modified golden-rule stock is invariant to changes in a discount factor.

In order to guarantee the existence of a nontrivial golden-rule stock, the line OV must be steeper than RG, which is to say,

$$1/\rho < b/a_I + (1-d) \equiv \theta. \tag{3}$$

We will impose the condition above in what follows.[19] This condition ensures that our technology is *productive*, i.e., there exists a production plan that can produce a positive net amount of investment goods.

The *modified golden rule* is formally defined as a pair $(\hat{x},\hat{p}) \in \mathbb{R}^2_+$ such that $(\hat{x},\hat{x}) \in \Omega$ and

$$u(\hat{x},\hat{x}) + (\rho-1)\hat{p}\hat{x} \geq u(x,x') + \hat{p}(\rho x' - x) \text{ for all } (x,x') \in \Omega. \tag{4}$$

It is easily seen that, in the RSL model, free disposal, bounded paths, and the existence of a stock expansible by the factor ρ^{-1} are satisfied. According to [31], then $\hat{x}$ is the stationary optimal capital stock.

We further introduce a key parameter

$$\zeta \equiv b/(a_C - a_I) - (1-d)$$

which features prominently in the analysis of the two-sector RSL model. Under the assumption that the consumption good section is more capital intensive, $\zeta > -1$. ζ can be interpreted as the marginal rate of transformation of capital between today and tomorrow under full utilization of capital.[20]

The following proposition formalizes the observations from the simple geometry of the today-tomorrow's diagram.

---

[19]The RSS model can be viewed as a special case of the RSL model with $a_C = 1$, $a_I = 0$, and $b = 1/a$. Since $a_I = 0$, Condition 3 is always satisfied in the RSS model. Also see Condition 2.10 in [34] for the same condition with $d = 1$.

[20]Note that the counterpart of ζ in the RSS model is ξ as in, for example, [21]. It is assumed that $\xi > 1$ in [21]. For discussion about $\xi \leq 1$, see [18].

Proposition 1. *There exists a modified golden rule, which is given by*

$$(\hat{x},\hat{p}) = \left(\frac{a_C(\zeta+1-d)}{\zeta+1}, \frac{1}{(a_C-a_I)(1+\rho\zeta)}\right).$$

2.3 The Zero Value-loss Line and the Iso Value-loss lines

The *value-loss*[21] $\delta^{\rho}_{(\hat{p},\hat{x})}(x,x')$ at the golden-rule price system $\hat{p}$ associated with one-period plan (x,x') is defined as the difference between the left-hand side and the right-hand side of the inequality 4

$$\delta^{\rho}_{(\hat{p},\hat{x})}(x,x') \equiv u(\hat{x},\hat{x}) + (\rho-1)\hat{p}\hat{x} - u(x,x') - \hat{p}(\rho x' - x) \text{ for all } (x,x') \in \Omega. \quad (5)$$

Lemma 1. *There is no value loss when labor and capital are fully utilized.*

The line MV in Figure 2 represents a locus of plans with full employment and full capacity utilization. In the triangle VOM, there exits labor not utilized for production of either sector and in the region $LVMD$, there exists excess capacity of capital. As such, the line MV yields zero value-loss at the golden-rule price system $\hat{p}$. This is to say that MV is the von-Neumann facet.

Next, we shall show that the lines parallel to MV are iso-value-loss lines; any two plans on the line have the same amount of value-losses. As an iso-value-loss line moves away from MV line, the associated value-loss increases. With this property, we can see that along the lines OV, MV and MD one-period value-loss is minimized for each given x.

For a plan (x,x') in the triangle MOV which lies below the MV line, we have $(1/a_C)(x-(a_I/b)(x'-(1-d)x)) < 1-(1/b)(x'-(1-d)x)$, so the level of utility is written as $u(x,x') = x(b+a_I(1-d))/(a_Cb) - a_Ix'/(a_Cb)$. Then, the value-loss is:

$$\delta(x,x') = \frac{b-a_Id}{a_Cb}\hat{x} + \hat{p}\hat{x}(\rho-1) - x(b+a_I(1-d))/(a_Cb) + a_Ix'/(a_Cb) - \hat{p}(\rho x' - x). \quad (6)$$

The marginal change of value-loss with respect to x' is $\partial\delta(x.x')/\partial x' = a_I/(a_Cb) - \hat{p}\rho < 0$.

For a plan (x,x') in the region $LVMD$ which lies above the MV line, we have $(1/a_C)(x-(a_I/b)(x'-(1-d)x)) > 1-(1/b)(x'-(1-d)x)$, so the level of utility is written as $u(x,x') = 1 - x'/b + (1-d)x/b$. Then, the value-loss is:

$$\delta(x,x') = -d\hat{x}/b + \hat{p}\hat{x}(\rho-1) + x'/b - (1-d)x/b - \hat{p}(\rho x' - x). \quad (7)$$

The marginal change of value-loss with respect to x' is $\partial\delta(x.x')/\partial x' = 1/b - \hat{p}\rho > 0$. The next result formally establishes this geometric insight, which will be frequently used in our partial characterization of the optimal policy correspondence.

Lemma 2. $a_I/(a_Cb) < \hat{p}\rho < 1/b$.

[21] We shall abbreviate $\delta^{\rho}_{(\hat{p},\hat{x})}(x,x')$ by $\delta^{\rho}(x,x')$.

3 The Dynamic Programming Approach

3.1 Preliminaries

In this subsection, we describe the dynamic programming approach with a value function and a policy correspondence to delineate the optimal policy. We define a value function, $V : X \to \mathbb{R}$ by:

$$V(x) = \sum_{t=0}^{\infty} \rho^t u(x(t), x(t+1)) \tag{8}$$

where $\{x(t), y(t)\}$ is an optimal program starting from $x(0) \in X = (0, \infty)$. The value function is continuous. For each $x \in X$, the Bellman equation:

$$V(x) = \max_{x' \in \Gamma(x)} \{u(x, x') + \rho V(x')\} \tag{9}$$

holds where $\Gamma(x) = \{x' : (x, x') \in \Omega\}$. For each $x \in X$, we denote by $h(x)$ the set of $x' \in \Gamma(x)$ which maximizes $\{u(x, x') + \rho V(x')\}$ among all $x' \in \Gamma(x)$. That is, the optimal policy correspondence is described as,

$$h(x) = \arg \max_{x' \in \Gamma(x)} \{u(x, x') + \rho V(x')\} \tag{10}$$

for each $x \in X$. A program $\{x(t), y(t)\}$ from $x \in X$ is an optimal program from x if and only if it satisfies the equation: $V(x(t)) = u(x(t), x(t+1)) + \rho V(x(t+1))$ for $t \geq 0$.[22]

We now provide some useful properties of the value function which help us partially characterize the optimal dynamics.

Lemma 3. *The value function V is (i) concave and (ii) strictly increasing with x.*

Lemma 4. *The value function, V, satisfies the following properties: i) $V(x) - V(\hat{x}) \leq \hat{p}(x - \hat{x})$ and ii) $V'_+(\hat{x}) \leq \hat{p} \leq V'_-(\hat{x})$.*

Since $\hat{x}$ is the stationary optimal stock, we have

$$V(\hat{x}) = \frac{1 - d\hat{x}/b}{1 - \rho} = \frac{b - da_I}{(b + da_C - da_I)(1 - \rho)}.$$

Before solving for the optimal policy correspondence, we provide the following lemma to guide us to partition the set $(0, \infty)$ into sub-regions over which the optimal dynamics tends to behave differently and will be characterized separately.

Lemma 5. *If $\zeta > 0$, then $\hat{x}/\theta < a_I < \hat{x} < a_C < \hat{x}/(1 - d)$. If $\zeta = 0$, then $\hat{x}/\theta = a_I < \hat{x} < a_C = \hat{x}/(1 - d)$. If $\zeta < 0$, then $a_I < \hat{x}/\theta < \hat{x} < \hat{x}/(1 - d) < a_C$.*

We now turn to characterizing the optimal policy correspondence and the resulting optimal dynamics for different values of ζ.

[22]We refer interested readers to chapter 4 of [25] for a general treatment and rigorous proofs of the value function being well-defined, the existence of optimal program, and the value function being the unique continuous solution to the Bellman equation.

3.2 A Partial Characterization

In this subsection, we provide necessary conditions for the optimal policy correspondence for the general parameter space and discuss the convergence property of the system when $\zeta \leq 1$.

3.2.1 The Case of Positive ζ

In order to derive the optimal policy correspondence, we divide the state space into the following three regions:[23] $(0, \hat{x}/\theta]$, $(\hat{x}/\theta, \hat{x}/(1-d))$, $[\hat{x}/(1-d), \infty)$. We first fully characterize the optimal policy correspondence h, which is single-valued, for the first and third region.

Lemma 6. *Let $\zeta > 0$. For $x \in (0, \hat{x}/\theta]$, $h(x) = \{\theta x\}$.*

Lemma 7. *Let $\zeta > 0$. For $x \in [\hat{x}/(1-d), \infty)$, $h(x) = \{(1-d)x\}$.*

The intuition behind Lemmas 6 and 7 is straightforward. If the initial capital stock is sufficiently low, much lower than the golden rule stock, then it is optimal to accumulate capital as quickly as possible and consequently, the economy fully specializes in producing investment goods. On the other hand, if the initial capital stock is sufficiently high, much higher than the golden rule stock, then it is optimal to reduce the capital stock as quickly as possible and as a result, the economy fully specializes in producing consumption goods.

What remains to characterize is the middle region $(\hat{x}/\theta, \hat{x}/(1-d))$. According to Lemma 5, we can further divide the middle region into four subregions: $(\hat{x}/\theta, a_I]$, $(a_I, \hat{x}]$, $(\hat{x}, a_C]$, $(a_C, \hat{x}/(1-d))$.

Proposition 2. *Let $\zeta > 0$. For $x \in (\hat{x}/\theta, \hat{x}/(1-d))$, the optimal policy correspondence h satisfies*

$$h(x) \subset \begin{cases} [\hat{x}, \theta x] & \text{for } x \in (\hat{x}/\theta, a_I] \\ [\hat{x}, \zeta(\hat{x} - x) + \hat{x}] & \text{for } x \in (a_I, \hat{x}] \\ [\zeta(\hat{x} - x) + \hat{x}, \hat{x}] & \text{for } x \in (\hat{x}, a_C] \\ [(1-d)x, \hat{x}] & \text{for } x \in (a_C, \hat{x}/(1-d)). \end{cases}$$

.

Collecting the results above, we obtain the following characterization of the optimal policy correspondence for $\zeta > 0$.

Theorem 1. *Let $\zeta > 0$. The optimal policy correspondence h satisfies*

$$h(x) \subset G(x) \equiv \begin{cases} \{\theta x\} & \text{for } x \in (0, \hat{x}/\theta] \\ [\hat{x}, \theta x] & \text{for } x \in (\hat{x}/\theta, a_I] \\ [\hat{x}, \zeta(\hat{x} - x) + \hat{x}] & \text{for } x \in (a_I, \hat{x}] \\ [\zeta(\hat{x} - x) + \hat{x}, \hat{x}] & \text{for } x \in (\hat{x}, a_C] \\ [(1-d)x, \hat{x}] & \text{for } x \in (a_C, \hat{x}/(1-d)). \\ \{(1-d)x\} & \text{for } x \in [\hat{x}/(1-d), \infty) \end{cases}$$

[23] By definition, $\theta > (1-d)$, so the three intervals are pairwise disjoint.

.

The proposition is illustrated in Figure 3. For the middle region, we know that the optimal policy correspondence has to be a selection from the two shaded triangles.

It should be noticed that the optimal policy correspondence h does not coincide with the optimal policy correspondence of the RSS model as in [20] when $a_I \to 0$. The sharp characterization of the optimal policy for x in $(a_I, \hat{x}]$ in the RSS model hinges on the fact that $a_I = 0$.[24]

Similar to the RSS model, when $0 < \zeta < 1$, the optimal policy leads to global convergence.

Corollary 1. *If $\zeta \in (0,1)$, the stock always converges to the golden rule stock $\hat{x}$.*

Moreover, we can show that, if $\zeta = 1$, the stock converges to either the golden rule stock $\hat{x}$ or a two-period cycle. The proof follows from the proof of Corollary 1 with slight modification. If $\theta a_I \leq a_C$, for $x \in [a_I, \hat{x})$, or if $a_I \leq (1-d)a_C$, for $x \in (\hat{x}, a_C]$, $G(x) = [\hat{x}, \zeta(\hat{x} - x) + \hat{x}] = [\hat{x}, 2\hat{x} - x]$ and $G^2(x) = [(1-\zeta^2)\hat{x} + \zeta^2 x, \hat{x}] = [x, \hat{x}]$, which could give rise to a two-period cycle with periodic points $(\hat{x}, 2\hat{x} - x)$.

To further compare a policy that yields a two-period cycle with a straight-down-to-turnpike policy, consider the case that $\theta a_I \leq a_C$. Let $x \in [a_I, \hat{x}]$. Consider a program under which the stock alternates between x and $2\hat{x} - x$. The utility of this program is given by

$$U_1(x) = \frac{u(x, 2\hat{x} - x) + \rho u(2\hat{x} - x, x)}{1 - \rho^2} = \frac{(2-d)x}{b(1+\rho)} - \frac{2(1-\rho+\rho d)\hat{x}}{b(1-\rho^2)} + \frac{1}{1-\rho}.$$

Consider another program under which the stock starts from x and jumps to $\hat{x}$ the next period and stay at the golden rule stock thereafter. The utility of this program is given by

$$U_2(x) = u(x, \hat{x}) + \frac{\rho u(\hat{x}, \hat{x})}{1-\rho} = \left(\frac{1}{a_C} + \frac{a_I(1-d)}{a_C b}\right) x - \left(\frac{a_I}{a_C b} + \frac{\rho d}{(1-\rho)b}\right)\hat{x} + \frac{\rho}{1-\rho}.$$

By construction, $U_1(\hat{x}) = U_2(\hat{x}) = (1 - d\hat{x}/b)/(1-\rho)$. Moreover,

$$\frac{\partial U_1(x)}{\partial x} = \frac{2-d}{b(1+\rho)} = \frac{b + (2-d)a_I}{b a_C (1+\rho)} = \frac{a_I(1+\theta)}{b a_C(1+\rho)} < \frac{a_I \theta}{b a_C} = \frac{1}{a_C} + \frac{a_I(1-d)}{a_C b} = \frac{\partial U_2(x)}{\partial x},$$

where the second equality follows from $\zeta = 1$ and the inequality follows from Condition 3. Since $\partial U_1(x)/\partial x < \partial U_2(x)/\partial x$ and $U_1(\hat{x}) = U_2(\hat{x})$, we must have $U_1(x) > U_2(x)$ for x in $[a_I, \hat{x})$, which suggests that the two-period policy dominates the straight-down-to-turnpike policy for $\zeta = 1$ and x in $[a_I, \hat{x})$.

[24] See the proof of their Lemma 2. For a sharp characterization of the optimal dynamics of the RSS model, also see [21] and especially their Figure 1.

3.2.2 The Case of Non-Positive ζ

For $\zeta \le 0$, the characterization of high and low initial stocks is the same, while characterization of the middle range becomes simpler.

Proposition 3. *Let $\zeta \le 0$. The optimal policy correspondence h satisfies*

$$h(x) \subset \begin{cases} \{\theta x\} & \text{for } x \in (0, a_I] \\ [\zeta(\hat{x} - x) + \hat{x}, \hat{x}] & \text{for } x \in (a_I, \hat{x}] \\ [\hat{x}, \zeta(\hat{x} - x) + \hat{x}] & \text{for } x \in (\hat{x}, a_C) \\ \{(1-d)x\} & \text{for } x \in [a_C, \infty) \end{cases} \tag{11}$$

The proposition is illustrated in Figure 4, with the two shaded triangles indicating the indeterminate region for the optimal policy. Taking into account the transition possibility set Ω, the optimal policy correspondence for the middle range can further be sharpened as

$$h(x) \subset \begin{cases} [\zeta(\hat{x} - x) + \hat{x}, \min\{\hat{x}, (1-d)x + b\}] & \text{for } x \in (a_I, \hat{x}] \\ [\max\{\hat{x}, (1-d)x\}, \zeta(\hat{x} - x) + \hat{x}] & \text{for } x \in (\hat{x}, a_C) \end{cases} \tag{12}$$

Following the similar argument for Corollary 1, the following corollary follows from Proposition 3 and the fact that $\zeta > d - 1 > -1$.

Corollary 2. *If $\zeta \le 0$, the stock converges to the golden rule stock $\hat{x}$ monotonically.*

4 The Value-Loss Approach

4.1 Preliminaries

In this section, based on the value-loss approach as in [18], we sharpen our characterization of the optimal dynamics. In particular, we completely pin down the optimal policy correspondence for $\zeta \le 1$ and discuss the bifurcation of the optimal policy correspondence with respect to the discount factor. It is worth emphasizing that this type of bifurcation analysis, as in [21], is conceptually different from the bifurcation analysis in the literature of the non-linear dynamics which focuses on primarily how the property of the dynamics of a *given* map changes with the parameter of interest.

We first restate the result described in Section 5 of [18].[25]

[25] The proof of the result see their appendix. The main idea is to establish the following equation

$$\sum_{t=0}^{\infty} \rho^t \left[u(x_t^*, x_{t+1}^*) - u(x_t, x_{t+1})\right] = \sum_{t=0}^{\infty} \rho^t \left[\delta^\rho(x_t, x_{t+1}) - \delta^\rho(x_t^*, x_{t+1}^*)\right],$$

for $\{x_t, y_t\}$ and $\{x_t^*, y_t^*\}$ that start from the same initial stock.

Lemma 8. *For all* $0 < \rho < 1$, *a program* $\{x_t^*, y_t^*\}$ *from* x_0 *is optimal if and only if*

$$\sum_{t=0}^{\infty} \rho^t [\delta^\rho(x_t, x_{t+1}) - \delta^\rho(x_t^*, x_{t+1}^*)] \geq 0$$

for every program $\{x_t, y_t\}$ *from* x_0.

The lemma says that, in order to find the optimal program that maximizes the total discounted utility, it is equivalent to finding a program that minimizes the total discounted value loss.

4.2 A Complete Characterization

4.2.1 The Case of Non-Positive ζ

According to Lemma 1, we know that if (x, x') satisfies $x' = \zeta(\hat{x} - x) + \hat{x}$ for $x \in [a_I, a_C]$, there is full utilization of resources and thus, the one-period value loss is zero. Further suggested by Lemma 8, we know that if a program yields zero total value loss, it must be optimal. In light of this insight, we can now characterize the optimal policy *function* for $\zeta \leq 0$.

Proposition 4. *Let* $\zeta \leq 0$. *The optimal policy correspondence is single valued, given by*

$$h(x) = \begin{cases} \{\theta x\} & \text{for } x \in (0, a_I] \\ \{\zeta(\hat{x} - x) + \hat{x}\} & \text{for } x \in (a_I, a_C) \\ \{(1 - d)x\} & \text{for } x \in [a_C, \infty) \end{cases} \tag{13}$$

The proof of this result is relatively straightforward. We first show that for the initial stock x in the middle range (a_I, a_C), if the next period stock follows $\zeta(\hat{x} - x) + \hat{x}$, then the stock will always stay on the full utilization arm, thus leading to zero total value loss. Any deviation from this policy inevitably leads to positive value loss and therefore is not optimal. For the initial stock that is outside the middle range, we show that the unique optimal policy is to converge to the middle range as fast as possible. Any delay leads to additional value loss.

We turn to sharpening the optimal policy correspondence for $0 < \zeta \leq 1$, which turns out to be much more complicated.

4.2.2 The Case of $0 < \zeta \leq 1$

Proposition 5. *Let* $\zeta \in (0, 1]$. *There are three possible cases:*

(i) If $a_C \geq a_I\theta$ *and* $a_C(1 - d) \geq a_I$, *then*

$$h(x) = \begin{cases} \{\theta x\} & \text{for } x \in (0, a_I) \\ \{\zeta(\hat{x} - x) + \hat{x}\} & \text{for } x \in [a_I, a_C] \\ \{(1 - d)x\} & \text{for } x \in (a_C, \infty) \end{cases} .$$

(ii) If $a_C < a_I\theta$ and $a_C(1-d) \geq a_I$, then

$$h(x) \subset \begin{cases} \{\theta x\} & \text{for } x \in (0, a_C/\theta] \\ [a_C, \min\{\theta x, \zeta(\hat{x}-x)+\hat{x}\}] & \text{for } x \in (a_C/\theta, a_C(\zeta-d)/\zeta) \\ \{\zeta(\hat{x}-x)+\hat{x}\} & \text{for } x \in [a_C(\zeta-d)/\zeta, a_C] \\ \{(1-d)x\} & \text{for } x \in (a_C, \infty) \end{cases}.$$

(iii) If $a_C \geq a_I\theta$ and $a_C(1-d) < a_I$, then

$$h(x) \subset \begin{cases} \{\theta x\} & \text{for } x \in (0, a_I) \\ \{\zeta(\hat{x}-x)+\hat{x}\} & \text{for } x \in [a_I, a_C(1+(1-d)/\zeta)-a_I/\zeta] \\ [\max\{\zeta(\hat{x}-x)+\hat{x}, (1-d)x\}, a_I] & \text{for } x \in (a_C(1+(1-d)/\zeta)-a_I/\zeta, a_I/(1-d)) \\ \{(1-d)x\} & \text{for } x \in [a_I/(1-d), \infty) \end{cases}.$$

Remark 1. In the RSS model, $a_I = 0$, so $a_C(1-d) \geq a_I$ always holds. For the special case of our characterization result in the RSS model, see Section 5 of [18]. It should be noted that under this extreme assumption that $a_I = 0$, sharper characterization could be obtained for $\zeta > 1$.

Remark 2. Given this refined characterization for $\zeta \leq 1$, the optimal dynamics as in Corollaries 1 can be easily obtained. In fact, when $\zeta = 1$, the result can be strengthened: the system almost always converges to a two-period cycle.

Case (i) is the simplest case of Proposition 5. The optimal policy is illustrated as the bold line in Figure 5. Similar to the case of $\zeta \leq 0$, the map consists of two arms of specialization and the middle arm of full resource utilization. The proof idea is also similar. Under the condition of Case (i), we show that full utilization of resource for any x starting from the middle range leads to zero value loss. Case (ii), as illustrated in Figure 6, is symmetric to Case (iii), as illustrated in Figure 7. The optimal policy has been pinned down except for the small shaded triangle. The intuition is that if the initial stock starts from the range that gives rise to the triangle, full utilization of resource, despite having a zero value loss for the current period, pushes the next period stock out of the middle range, and therefore, the value loss starting from the next period must be positive. There is a tradeoff between suffering from the value loss this period in order to stay on the middle range and delaying the value loss to the next period, which naturally hinges on the discount factor. The next two propositions fully characterize the optimal policy correspondence for case (ii) and (iii) in Proposition 5, demonstrating an intriguing bifurcation pattern of how optimal policy changes with the discount factor.

Theorem 2. *Suppose $\zeta \in (0,1]$, $a_C < a_I\theta$, and $a_C(1-d) \geq a_I$. Let t_0 be the smallest integer such that $a_I\theta(1-d)^{t_0} < a_C$. Let $\bar{\rho}_t$ be the unique positive root to the following equation*

$$a_C b(1-d)^t \rho^{t+1} - (a_C - a_I)a_I\zeta\rho - a_I(a_C - a_I) = 0,$$

for t being the integer satisfying $1 \leq t \leq t_0$. Then $\bar{\rho}_t$ strictly increases with t and $\bar{\rho}_1 > 1/\theta$. Moreover, the optimal policy correspondence for $x \in (a_C/\theta, a_C(\zeta - d)/\zeta)$ depends on ρ, which is given by

$$h(x) = \begin{cases} \{a_C\} & \text{for } \rho < \bar{\rho}_1 \\ [a_C, \min\{\zeta(\hat{x} - x) + \hat{x}, \theta x\}] & \text{for } \rho = \bar{\rho}_1 \\ \min\{\zeta(\hat{x} - x) + \hat{x}, \theta x\} & \text{for } \rho > \bar{\rho}_1 \end{cases},$$

for $t_0 = 1$, and

$$h(x) = \begin{cases} \{a_C\} & \text{for } \rho < \bar{\rho}_1 \\ [\min\{\zeta(\hat{x} - x) + \hat{x}, \theta x, a_C/(1-d)^{t-1}\}, & \\ \min\{\zeta(\hat{x} - x) + \hat{x}, \theta x, a_C/(1-d)^{t}\}] & \text{for } \rho = \bar{\rho}_t,\ t = 1, ..., t_0 \\ \min\{\zeta(\hat{x} - x) + \hat{x}, \theta x, a_C/(1-d)^{t}\} & \text{for } \bar{\rho}_{t+1} > \rho > \bar{\rho}_t,\ t = 1, ..., t_0 - 1 \\ \min\{\zeta(\hat{x} - x) + \hat{x}, \theta x\} & \text{for } \rho > \bar{\rho}_{t_0} \end{cases},$$

for $t_0 > 1$.

This proposition obtains precisely the optimal policy correspondence for $x \in (a_C/\theta, a_C(\zeta - d)/\zeta)$ of case (ii) in Proposition 5. Generically, the optimal policy correspondence is a function. To understand the characterization, we illustrate an example of $t_0 = 3$ in Figure 8. When $\rho < \bar{\rho}_1$, the optimal policy for the small interval $(a_C/\theta, a_C(\zeta - d)/\zeta)$ is represented by the segment AB. When $\rho = \bar{\rho}_1$, the optimal policy is represented by the trapezoid ADEB. When $\bar{\rho}_1 < \rho < \bar{\rho}_2$, the optimal policy is represented by three segments: AD, DE, and EB. When $\rho = \bar{\rho}_2$, the optimal policy is represented by two segments, AD and EB, and a trapezoid DFGE. When $\bar{\rho}_2 < \rho < \bar{\rho}_3$, the optimal policy is represented by three segments again: AF, FG, and GB. When $\rho = \bar{\rho}_3$, the optimal policy is represented by two segments, AF and GB, and a triangle, FCG. Last, when $\rho > \bar{\rho}_3$, the optimal policy is represented by two segments, AC and CB, which coincides with Case (i) in Proposition 5. Even though it looks complicated, this bifurcation pattern is consistent with our intuition: When the discount factor increases, the planner is more willing to sacrifice today's consumption for tomorrow's gain and the optimal policy gets closer to the case that the economy converges as soon as possible to the full utilization arm.

The main proof idea of this proposition is again based on the value loss approach. We first note that Proposition 5 implies that the economy only suffers from positive value loss for finite periods for any initial stock. Then the key step is to calculate the total discounted value loss for up to $t_0 + 1$ periods for an initial stock in $(a_C/\theta, a_C(\zeta - d)/\zeta)$.

The technical complication arises from the fact that the number of periods with positive value loss depends crucially on the policy chosen for the initial stock. We utilize the relationship between the value loss for the finite periods and the discount factor to finally establish the optimality.

Since $\bar{\rho}_1 > 1/\theta$, there always exists a ρ sufficiently close to $1/\theta$ with $\bar{\rho}_1 > \rho > 1/\theta$ such that the optimal policy function has a flat top. Moreover, notice that it is possible to have $\bar{\rho}_1 \geq 1$. For example, consider $a_I = 1$, $a_C = 1.5$, $d = 0.25$, $b = 0.8$. It can be verified that all the conditions in Proposition 2 are satisfied and we have $t_0 = 1$ with $\bar{\rho}_1 \approx 1.018 > 1$. According to Proposition 2, the optimal policy correspondence has a flat top with $h(x) = \{a_C\}$ for any $x \in (a_C/\theta, a_C(\zeta - d)/\zeta)$. For a numerical example that the optimal policy correspondence bifurcates with respect to the discount factor, consider $a_I = 1$, $a_C = 2$, $d = 0.05$, $b = 1.3$. Again, we can verify that all the conditions in Proposition 2 are satisfied and we have $t_0 = 3$ with $\bar{\rho}_1 \approx 0.711$, $\bar{\rho}_2 \approx 0.818$, and $\bar{\rho}_3 \approx 0.875$.

We now refine the characterization for the optimal dynamics for case (iii) in Proposition 5.

Theorem 3. *Suppose $\zeta \in (0,1]$, $a_C \geq a_I\theta$, and $a_C(1-d) < a_I$. Let t_0 be the smallest integer such that $a_C(1-d)\theta^{t_0} > a_I$. Let $\tilde{\rho}_t$ be the unique positive root to the following equation*

$$b\theta^t \rho^{t+1} - (a_C - a_I)\zeta\rho - (a_C - a_I) = 0,$$

for t being the integer satisfying $1 \leq t \leq t_0$. Then $\tilde{\rho}_t$ strictly decreases with t and $\tilde{\rho}_{t_0} > 1/\theta$. Moreover, the optimal policy correspondence for $x \in (a_C(1 + (1-d)/\zeta) - a_I/\zeta, a_I/(1-d))$ depends on ρ, which is given by

$$h(x) = \begin{cases} \{a_I\} & \text{for } \rho > \tilde{\rho}_1 \\ [\max\{\zeta(\hat{x} - x) + \hat{x}, (1-d)x\}, a_I] & \text{for } \rho = \tilde{\rho}_1 \\ \max\{\zeta(\hat{x} - x) + \hat{x}, (1-d)x\} & \text{for } \rho < \tilde{\rho}_1 \end{cases},$$

for $t_0 = 1$, and

$$h(x) = \begin{cases} \{a_I\} & \text{for } \rho > \tilde{\rho}_1 \\ [\max\{\zeta(\hat{x} - x) + \hat{x}, (1-d)x, a_I/\theta^t\}, & \\ \max\{\zeta(\hat{x} - x) + \hat{x}, (1-d)x, a_I/\theta^{t-1}\}] & \text{for } \rho = \tilde{\rho}_t,\ t = 1, ..., t_0 \\ \max\{\zeta(\hat{x} - x) + \hat{x}, (1-d)x, a_I/\theta^t\} & \text{for } \tilde{\rho}_{t+1} < \rho < \tilde{\rho}_t,\ t = 1, ..., t_0 - 1 \\ \max\{\zeta(\hat{x} - x) + \hat{x}, (1-d)x\} & \text{for } \rho < \tilde{\rho}_{t_0} \end{cases},$$

for $t_0 > 1$.

The proof idea is similar to the previous one and the characterization result also looks symmetric. An example for $t_0 = 3$ is illustrated in Figure 9.

Since $\tilde{\rho}_{t_0} > 1/\theta$, there always exists a ρ sufficiently close to $1/\theta$ with $\tilde{\rho}_{t_0} > \rho > 1/\theta$ such that the optimal policy has a V-shaped bottom. Again, notice that it is possible to have $\tilde{\rho}_{t_0} \geq 1$. For example, consider $a_I = 1$, $a_C = 1.3$, $d = 0.25$, $b = 0.3$. It can be verified that all the conditions in Proposition 3 are satisfied and we have $t_0 = 1$ with $\tilde{\rho}_1 \approx 1.102 > 1$. According to Proposition 3, the optimal policy correspondence has a V-shaped bottom with $h(x) = \max\{\zeta(\hat{x} - x) + \hat{x}, (1-d)x\}$ for any $x \in (a_C(1 + (1-d)/\zeta) - a_I/\zeta, a_I/(1-d))$. For a numerical example that the optimal policy correspondence bifurcates with respect to the discount factor, consider $a_I = 1$, $a_C = 3$, $d = 0.95$, $b = 2$. Again, we can verify that all the conditions in Proposition 3 are satisfied and we have $t_0 = 3$ with $\tilde{\rho}_1 \approx 0.968$, $\tilde{\rho}_2 \approx 0.740$, and $\tilde{\rho}_3 \approx 0.659$.

Even though we have shown that there is global convergence for the optimal policy when $\zeta < 1$, our analysis above suggests that the optimal policy itself might experience cascade of changes with the discount factor. This stands in contrast with the bifurcation result of the RSS model as in [21]. As a special case of our model, they demonstrate that the bifurcation only arises when $\zeta > 1$.

5 Discussion and Open Questions

In this paper, we provide a partial characterization of the optimal policy correspondence of the RSL model, and based on the value-loss approach, we fully characterize the optimal policy for $\zeta \leq 1$. The characterization results bring out how the optimal policy bifurcates with respect to the discount factor. And so a natural question arises as to the nature of the optimal policy for the case $\zeta > 1$ even for the case of a capital-intensive consumption good sector. We hope to turn to this in a future investigation.

6 Appendix: Proofs of the Results

Proof of Proposition 1:

Since $a_C > a_I$, $\zeta > d - 1$, which implies $(\hat{x}, \hat{p}) \in \mathbb{R}^2_+$. Given Condition 3, $b/a_I > d + 1/\rho - 1 > d$, where the second inequality stems from $0 < \rho < 1$. Since $b/a_I > d$, $d\hat{x} = dba_C/(da_C + b - da_I) < b$ and $d\hat{x} < b\hat{x}/a_I$. Therefore, $d\hat{x} < b \min\{1, \hat{x}/a_I\}$. Then we must have $(\hat{x}, \hat{x}) \in \Omega$.

Consider y in $\Lambda(x, x')$. Define

$$\begin{aligned}
\alpha(x, x', y) &= (1/a_C)(x - (a_I/b)(x' - (1-d)x)) - y, \\
\beta(x, x', y) &= 1 - (1/b)(x' - (1-d)x) - y.
\end{aligned}$$

We have

$$y + \hat{p}(\rho x' - x) = (1 - A) - A\alpha(x, x', y) - (1 - A)\beta(x, x', y),$$

where $A \equiv \frac{a_C(1-\rho(1-d))}{(a_C - a_I)(1+\rho\zeta)}$. Since $a_C > a_I$ (Condition 1), $\zeta > -1$, and we know $0 < \rho < 1$ and $0 < d < 1$, we have $A > 0$. Given Condition 3, $A < 1$. By construction, $\alpha(x, x', y) \geq 0$ and $\beta(x, x', y) \geq 0$, which implies

$$y + \hat{p}(\rho x' - x) \leq 1 - A.$$

Let $\hat{y} = 1-(d/b)\hat{x} = (1/a_C)(1-a_I d/b)\hat{x} = (b-da_I)/(b-da_I+da_C) > 0$. We have $\hat{y} \in \Lambda(\hat{x},\hat{x})$ and $\alpha(\hat{x},\hat{x},\hat{y}) = \beta(\hat{x},\hat{x},\hat{y}) = 0$. Therefore, $\hat{y} = u(\hat{x},\hat{x})$ and $\hat{y} + (\rho-1)\hat{p}\hat{x} = 1 - A$, which implies

$$u(\hat{x},\hat{x}) + (\rho-1)\hat{p}\hat{x} \geq y + \hat{p}(\rho x' - x) \text{ for all } (x,x') \in \Omega \text{ and } y \in \Lambda(x,x'). \tag{14}$$

Since $u(x,x') = \max \Lambda(x,x')$, we then obtain the desired inequality

$$u(\hat{x},\hat{x}) + (\rho-1)\hat{p}\hat{x} \geq u(x,x') + \hat{p}(\rho x' - x) \text{ for all } (x,x') \in \Omega. \tag{15}$$

Proof of Lemma 1:

When labor and capital are fully utilized, we must have

$$(1/a_C)(x - (a_I/b)(x' - (1-d)x)) = 1 - (1/b)(x' - (1-d)x).$$

Let $y = (1/a_C)(x - (a_I/b)(x' - (1-d)x)) = 1 - (1/b)(x' - (1-d)x)$. Following the argument in the proof of Proposition 1, we then have $\alpha(x,x',y) = \beta(x,x',y) = 0$, and we have shown $\alpha(\hat{x},\hat{x},\hat{y}) = \beta(\hat{x},\hat{x},\hat{y}) = 0$, so $u(\hat{x},\hat{x}) + (\rho-1)\hat{p}\hat{x} = u(x,x') + \hat{p}(\rho x' - x)$, or equivalently, $\delta^\rho(x,x') = 0$.

Proof of Lemma 2:

Given Condition 3, we have $\rho b > a_I - \rho(1-d)a_I$, and since we know $a_C > a_I$, we must have

$$\frac{\rho}{(a_C - a_I) + \rho b - \rho(1-d)(a_C - a_I)} > \frac{a_I}{a_C b},$$

where the left hand side is equal to $\hat{p}\rho$.

Moreover, since $a_C > a_I$, $\rho \in (0,1)$ and $d \in (0,1)$, $(a_C - a_I)(1 - \rho(1-d)) > 0$, which implies

$$\hat{p}\rho = \frac{\rho}{(a_C - a_I) + \rho b - \rho(1-d)(a_C - a_I)} < \frac{1}{b}.$$

Proof of Lemma 3:

For (i), consider x_1 and x_2 in X and λ in $(0,1)$. Let $x_3 = \lambda x_1 + (1-\lambda)x_2$. Let $\{x_i(t), y_i(t)\}$ be an optimal program starting from x_i for $i = 1,2,3$. By construction, Ω is convex. To see u being concave, letting $x = \lambda x_1(t) + (1-\lambda)x_2(t)$ and $x' = \lambda x_1(t+1) + (1-\lambda)x_2(t+1)$, we have

$$\begin{aligned}
&u(x,x') = \min\{(1/a_C)(x - (a_I/b)(x' - (1-d)x)), 1 - (1/b)(x' - (1-d)x)\} \\
&\geq \lambda \min\{(1/a_C)(x_1(t) - (a_I/b)(x_1(t+1) - (1-d)x_1(t))), 1 - (1/b)(x_1(t+1) - (1-d)x_1(t))\} \\
&+(1-\lambda)\min\{(1/a_C)(x_2(t) - (a_I/b)(x_2(t+1) - (1-d)x_2(t))), 1 - (1/b)(x_2(t+1) - (1-d)x_2(t))\} \\
&= \lambda u(x_1(t), x_1(t+1)) + (1-\lambda)u(x_2(t), x_2(t+1)).
\end{aligned}$$

Since u is concave, we have

$$\begin{aligned}
\lambda V(x_1) + (1-\lambda)V(x_2) &\leq \sum_{t=0}^{\infty} \rho^t u(\lambda x_1(t) + (1-\lambda)x_2(t), \lambda x_1(t+1) + (1-\lambda)x_2(t+1)) \\
&\leq V(\lambda x_1 + (1-\lambda)x_2),
\end{aligned}$$

where the second inequality follows from the fact that $\{\lambda x_1(t)+(1-\lambda)x_2(t)\}_{t=0}^{\infty}$ generates a program starting from $x_3 = \lambda x_1 + (1-\lambda)x_2$.

For (ii), consider two initial stocks, x_1 and x_2 with $x_1 < x_2$. Let $x_1' \in \arg\max_{x' \in \Gamma(x_1)}\{u(x_1, x') + \rho V(x')\}$. We then have $V(x_1) = u(x_1, x_1') + \rho V(x_1')$. By the optimality of V, $V(x_2) \geq u(x_2, x_1') + \rho V(x_1') > u(x_1, x_1') + \rho V(x_1') = V(x_1)$, where the second inequality follows from $u(x, x')$ being strictly increasing with x.

Proof of Lemma 4:

For i), based on the definition of the value function,

$$\begin{aligned} V(x) - V(\hat{x}) &= \sum_{t=0}^{\infty} \rho^t [u(x(t), x(t+1)) - u(\hat{x}, \hat{x})] \\ &\leq \sum_{t=0}^{\infty} \rho^t [(\rho - 1)\hat{p}\hat{x} - \hat{p}(\rho x(t+1) - x(t))] \\ &= (\rho - 1)\hat{p}\hat{x}/(1-\rho) + \hat{p}x(0) = \hat{p}(x - \hat{x}) \end{aligned}$$

where $x = x(0)$ and the inequality follows from Equation 4.

For ii), take x such as $x = \hat{x} + \epsilon$ for $\epsilon > 0$. Letting $\epsilon \to 0$, we obtain $V_+'(\hat{x}) \leq \hat{p}$. Similarly, take x such as $x = \hat{x} - \epsilon$ for $\epsilon > 0$. Letting $\epsilon \to 0$, we obtain $\hat{p} \leq V_-'(\hat{x})$.

Proof of Lemma 5:

Condition 1 implies $\zeta > d - 1$, which guarantees $\hat{x} = a_C(\zeta + 1 - d)/(\zeta + 1) < a_C$.

Condition 3 implies $b > da_I$ and we know $a_C > a_I$, so we have

$$\hat{x} - a_I = \frac{a_C b}{a_C d - a_I d + b} - a_I = \frac{(a_C - a_I)(b - da_I)}{a_C d - a_I d + b} > 0.$$

Therefore, $\hat{x} > a_I$.

So far we have shown that $a_I < \hat{x} < a_C$ holds for any ζ.

We now consider the relationship between $\hat{x}/(1-d)$ and a_C. Since $\hat{x}/(1-d) = a_C(\zeta+1-d)/(\zeta+1-d-d\zeta)$, $\hat{x}/(1-d) > a_C$. if and only if $\zeta > 0$. If $\zeta = 0$, $\hat{x}/(1-d) = a_C$. If $\zeta < 0$, $\hat{x}/(1-d) < a_C$, but $\hat{x}/(1-d) > \hat{x}$ still holds because $d > 0$.

Last, we consider the relationship between $\hat{x}/\theta$ and a_I.

$$\begin{aligned} \theta a_I - \hat{x} &= b + (1-d)a_I - \frac{a_C(\zeta + 1 - d)}{\zeta + 1} \\ &= \zeta(a_C - a_I) + (1-d)a_C - \frac{a_C(\zeta + 1 - d)}{\zeta + 1} \\ &= \zeta\left(a_C - a_I - \frac{a_C d}{\zeta + 1}\right) \\ &= \frac{\zeta}{\zeta + 1}\left(b + d(a_C - a_I) - a_C d\right) \\ &= \frac{\zeta}{\zeta + 1}\left(b - da_I\right). \end{aligned}$$

Hence, $a_I > \hat{x}/\theta$ if and only if $\zeta > 0$. If $\zeta = 0$, $a_I = \hat{x}/\theta$. If $\zeta < 0$, $a_I < \hat{x}/\theta$, but $\hat{x}/\theta < \hat{x}$ still holds because $\theta > 1$ according to Condition 3.

Proof of Lemma 6:

Suppose on the contrary there exists $x \in (0, \hat{x}/\theta]$ and $z \in h(x)$ such that $z \neq \theta x$. Since $\zeta > 0$, according to Lemma 5, $x \leq \hat{x}/\theta < a_I$. Then we must have $z < \theta x$, and

$$V(x) = u(x, z) + \rho V(z) \geq u(x, \theta x) + \rho V(\theta x).$$

Rearranging the equation, we have

$$u(x, z) - u(x, \theta x) \geq \rho(V(\theta x) - V(z)) \geq \rho V'_-(\theta x)(\theta x - z) \geq \rho V'_-(\hat{x})(\theta x - z) \geq \rho\hat{p}(\theta x - z),$$

where the second inequality follows from concavity of V, the third inequality follows from concavity of V and the fact that $\theta x \leq \hat{x}$ for $x \in (0, \hat{x}/\theta]$, and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (0, \hat{x}/\theta]$, we have

$$u(x, z) - u(x, \theta x) = (1/a_C)[x - (a_I/b)(z - (1 - d)x)] = \frac{a_I}{ba_C}(\theta x - z) < \hat{p}\rho(\theta x - z),$$

where the last inequality follows from Lemma 2. Since we have shown that $u(x, z) - u(x, \theta x) \geq \rho\hat{p}(\theta x - z)$. This leads to a contradiction and establishes the desired result.

Proof of Lemma 7:

Suppose on the contrary there exists $x \in [\hat{x}/(1 - d), \infty)$ and $z \in h(x)$ such that $z \neq (1 - d)x$. Then we must have $z > (1 - d)x$, and

$$V(x) = u(x, z) + \rho V(z) \geq u(x, (1 - d)x) + \rho V((1 - d)x).$$

Rearranging the equation, we have

$$\begin{aligned} u(x, (1 - d)x) - u(x, z) &\leq \rho(V(z) - V((1 - d)x)) \leq \rho V'_+((1 - d)x)(z - (1 - d)x) \\ &\leq \rho V'_+(\hat{x})(z - (1 - d)x) \leq \rho\hat{p}(z - (1 - d)x), \end{aligned}$$

where the second inequality follows from concavity of V, the third inequality follows from concavity of V and the fact that $\hat{x} \leq (1 - d)x$, and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in [\hat{x}/(1 - d), \infty)$, we have

$$u(x, (1 - d)x) - u(x, z) = (z - (1 - d)x)/b > \hat{p}\rho(z - (1 - d)x),$$

where the last inequality follows from Lemma 2. Since we have shown that $u(x, (1 - d)x) - u(x, z) \leq \rho\hat{p}(z - (1 - d)x)$. This leads to a contradiction and establishes the desired result.

Proof of Lemma 2:

We proceed by going over each subregion.

Subregion $(\hat{x}/\theta, a_I]$**:**

Suppose on the contrary there exists $x \in (\hat{x}/\theta, a_I]$ and $z \in h(x)$ such that $z \notin [\hat{x}, \theta x]$. Then we must have $z < \hat{x}$. By the optimality of z,

$$V(x) = u(x,z) + \rho V(z) \geq u(x,\hat{x}) + \rho V(\hat{x}).$$

Rearranging the equation, we have

$$u(x,z) - u(x,\hat{x}) \geq \rho(V(\hat{x}) - V(z)) \geq \rho V'_-(\hat{x})(\hat{x} - z) \geq \rho\hat{p}(\hat{x} - z),$$

where the second inequality follows from concavity of V and the third inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (\hat{x}/\theta, a_I]$, we have

$$u(x,z) - u(x,\hat{x}) \leq \frac{a_I}{ba_C}(\hat{x} - z) < \hat{p}\rho(\hat{x} - z),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x,z) - u(x,\hat{x}) \geq \rho\hat{p}(\hat{x} - z)$. This leads to a contradiction and establishes the desired result.

Subregion $(a_I, \hat{x}]$**:**

Consider $(a_I, \hat{x}]$.[26] Suppose on the contrary there exists $x \in (a_I, \hat{x}]$ and $z \in h(x)$ such that $z \notin [\hat{x}, \zeta(\hat{x} - x) + \hat{x}]$. There are two possible cases: (i) $z < \hat{x}$; (ii) $z > \zeta(\hat{x} - x) + \hat{x}$.

Consider (i) $z < \hat{x}$. We have

$$V(x) = u(x,z) + \rho V(z) \geq u(x,\hat{x}) + \rho V(\hat{x}).$$

Rearranging the equation, we have

$$u(x,z) - u(x,\hat{x}) \geq \rho(V(\hat{x}) - V(z)) \geq \rho V'_-(\hat{x})(\hat{x} - z) \geq \rho\hat{p}(\hat{x} - z),$$

where the second inequality follows from concavity of V and the third inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (a_I, \hat{x}]$, we have

$$u(x,z) - u(x,\hat{x}) \leq \frac{a_I}{ba_C}(\hat{x} - z) < \hat{p}\rho(\hat{x} - z),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x,z) - u(x,\hat{x}) \geq \rho\hat{p}(\hat{x} - z)$. This leads to a contradiction and establishes the desired result.

---

[26] This is the range of $(0, \hat{x}]$ in the RSS model. When $a_I = 0$, the optimal policy correspondence can be further reduced to a function. For a complete characterization for $x \in (a_I, \hat{x}]$ in the RSS model, see Lemma 2 in [20].

Consider (ii) $z > \zeta(\hat{x} - x) + \hat{x}$. We have

$$V(x) = u(x, z) + \rho V(z) \geq u(x, \zeta(\hat{x} - x) + \hat{x}) + \rho V(\zeta(\hat{x} - x) + \hat{x}).$$

Rearranging the equation, we have

$$\begin{aligned} u(x, \zeta(\hat{x} - x) + \hat{x}) - u(x, z) &\leq \rho(V(z) - V(\zeta(\hat{x} - x) + \hat{x})) \\ &\leq \rho V'_+(\zeta(\hat{x} - x) + \hat{x})(z - (\zeta(\hat{x} - x) + \hat{x})) \\ &\leq \rho V'_+(\hat{x})(z - (\zeta(\hat{x} - x) + \hat{x})) \leq \rho\hat{p}(z - (\zeta(\hat{x} - x) + \hat{x})) \end{aligned}$$

where the second inequality follows from concavity of V, the third inequality follows from concavity of V and the fact that $\zeta(\hat{x} - x) + \hat{x} \geq \hat{x}$ for $x \in (a_I, \hat{x}]$, and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (a_I, \hat{x}]$, we have

$$u(x, \zeta(\hat{x} - x) + \hat{x}) - u(x, z) \geq (1/b)(z - (\zeta(\hat{x} - x) + \hat{x})) > \hat{p}\rho(z - (\zeta(\hat{x} - x) + \hat{x})),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x, (\zeta(\hat{x} - x) + \hat{x})) - u(x, z) \leq \rho\hat{p}(z - (\zeta(\hat{x} - x) + \hat{x}))$. This leads to a contradiction and establishes the desired result.

Subregion $(\hat{x}, a_C]$**:**

Suppose on the contrary there exists $x \in (\hat{x}, a_C]$ and $z \in h(x)$ such that $z \notin [\zeta(\hat{x} - x) + \hat{x}, \hat{x}]$. There are two possible cases: (i) $z > \hat{x}$; (ii) $z < \zeta(\hat{x} - x) + \hat{x}$.

Consider (i) $z > \hat{x}$. We have

$$V(x) = u(x, z) + \rho V(z) \geq u(x, \hat{x}) + \rho V(\hat{x}).$$

Rearranging the equation, we have

$$u(x, \hat{x}) - u(x, z) \leq \rho(V(z) - V(\hat{x})) \leq \rho V'_+(\hat{x})(z - \hat{x}) \leq \rho\hat{p}(z - \hat{x}),$$

where the second inequality follows from concavity of V and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (\hat{x}, a_C]$, we have

$$u(x, \hat{x}) - u(x, z) \geq (1/b)(z - \hat{x}) > \hat{p}\rho(z - \hat{x}),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x, \hat{x}) - u(x, z) \leq \rho\hat{p}(z - \hat{x})$. This leads to a contradiction and establishes the desired result.

Consider (ii) $z < \zeta(\hat{x} - x) + \hat{x}$. We have

$$V(x) = u(x, z) + \rho V(z) \geq u(x, \zeta(\hat{x} - x) + \hat{x}) + \rho V(\zeta(\hat{x} - x) + \hat{x}).$$

Rearranging the equation, we have

$$\begin{aligned}
u(x,z)-u(x,\zeta(\hat{x}-x)+\hat{x}) &\geq \rho(V(\zeta(\hat{x}-x)+\hat{x})-V(z)) \\
&\geq \rho V'_{-}(\zeta(\hat{x}-x)+\hat{x})((\zeta(\hat{x}-x)+\hat{x})-z) \\
&\geq \rho V'_{-}(\hat{x})((\zeta(\hat{x}-x)+\hat{x})-z) \geq \rho\hat{p}((\zeta(\hat{x}-x)+\hat{x})-z)
\end{aligned}$$

where the second inequality follows from concavity of V, the third inequality follows from concavity of V and the fact that $\zeta(\hat{x}-x)+\hat{x} \leq \hat{x}$ for $x \in (\hat{x}, a_C]$, and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (\hat{x}, a_C]$, we have

$$u(x,z)-u(x,\zeta(\hat{x}-x)+\hat{x}) \leq \frac{a_I}{a_C b}((\zeta(\hat{x}-x)+\hat{x})-z) < \hat{p}\rho((\zeta(\hat{x}-x)+\hat{x})-z),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x,z)-u(x,(\zeta(\hat{x}-x)+\hat{x})) \geq \rho\hat{p}((\zeta(\hat{x}-x)+\hat{x})-z)$. This leads to a contradiction and establishes the desired result.

Subregion $(a_C, \hat{x}/(1-d))$**:**

Last, we consider $(a_C, \hat{x}/(1-d))$. Suppose on the contrary there exists $x \in (a_C, \hat{x}/(1-d))$ and $z \in h(x)$ such that $z \notin [(1-d)x, \hat{x}]$. Then we must have $z > \hat{x}$. By optimality of z,

$$V(x) = u(x,z)+\rho V(z) \geq u(x,\hat{x})+\rho V(\hat{x}).$$

Rearranging the equation, we have

$$u(x,\hat{x})-u(x,z) \leq \rho(V(z)-V(\hat{x})) \leq \rho V'_{+}(\hat{x})(z-\hat{x}) \leq \rho\hat{p}(z-\hat{x}),$$

where the second inequality follows from concavity of V, and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (a_C, \hat{x}/(1-d))$, we have

$$u(x,\hat{x})-u(x,z) \geq (1/b)(z-\hat{x}) > \hat{p}\rho(z-\hat{x}),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x,\hat{x})-u(x,z) \leq \rho\hat{p}(z-\hat{x})$. This leads to a contradiction and establishes the desired result.

Proof of Corollary 1:

First, $h(\hat{x}) = \{\hat{x}\}$. If the initial stock is the golden rule stock, the system will stay at the golden rule stock. To see the dynamics for the initial stock $x \neq \hat{x}$, we rewrite $\zeta < 1$ more explicitly as

$$\frac{b}{a_C-a_I}-(1-d) < 1 \Leftrightarrow b-(1-d)(a_C-a_I) < a_C-a_I \Leftrightarrow [\theta a_I - a_C]+[a_I-(1-d)a_C] < 0,$$

where the last inequality suggests either (i) $\theta a_I < a_C$ or (ii) $a_I < (1-d)a_C$ (or both).

Consider (i) $\theta a_I < a_C$. Suppose $x \in [a_I, \hat{x})$. We know $G(x) = [\hat{x}, \zeta(\hat{x}-x)+\hat{x}]$. Since $\theta a_I < a_C$, $G(x) \subset [\hat{x}, \theta a_I] \subset [\hat{x}, a_C]$, and therefore, $G^2(x) = [(1-\zeta^2)\hat{x} + \zeta^2 x, \hat{x}]$. Since $\zeta \in (0,1)$ and $x < \hat{x}$, we have $(1-\zeta^2)\hat{x}+\zeta^2 x > x$. This implies that $\lim_{t\to\infty} G^{2t}(x) = \{\hat{x}\}$. Since $\lim_{x\to\hat{x}} \zeta(\hat{x}-x)+\hat{x} = \hat{x}$, we must have $\lim_{t\to\infty} G^{2t+1}(x) = \{\hat{x}\}$. This leads to the desired conclusion that x converges to $\hat{x}$ for $x \in [a_I, \hat{x})$. Since $\lim_{t\to\infty} G^t(x) = \{\hat{x}\}$ for any $x \in [a_I, \hat{x})$, $\lim_{t\to\infty} G^t(x) = \{\hat{x}\}$ for any $x \in G([a_I, \hat{x})) = [\hat{x}, \theta a_I]$. Further, since $G([\hat{x}/\theta, a_I]) = [\hat{x}, \theta a_I]$, the system must converge for any x in $[\hat{x}/\theta, a_I]$. Since we know that $h(x) = \{\theta x\}$ for x in $(0, \hat{x}/\theta)$, for any x in $(0, \hat{x}/\theta)$, after finite periods, the stock must enter the region $[\hat{x}/\theta, \hat{x})$, thus leading to convergence. So far we have shown that the system converges for any x in $(0, \theta a_I]$. According to Proposition 1, for any x greater than $\hat{x}$, after finite periods, the stock must be below $\hat{x}$, again according to what we have shown, leading to convergence.

Consider (ii) $a_I < (1-d)a_C$. Suppose $x \in (\hat{x}, a_C]$. We know $G(x) = [\zeta(\hat{x}-x)+\hat{x}, \hat{x}]$. Since $a_I < (1-d)a_C$, $G(x) \subset [(1-d)a_C, \hat{x}] \subset [a_I, \hat{x}]$, and therefore, $G^2(x) = [\hat{x}, (1-\zeta^2)\hat{x} + \zeta^2 x]$. Since $\zeta \in (0,1)$ and $x > \hat{x}$, we have $(1-\zeta^2)\hat{x} + \zeta^2 x < x$. This implies that $\lim_{t\to\infty} G^{2t}(x) = \{\hat{x}\}$. Since $\lim_{x\to\hat{x}} \zeta(\hat{x}-x)+\hat{x} = \hat{x}$, we must have $\lim_{t\to\infty} G^{2t+1}(x) = \{\hat{x}\}$. This leads to the desired conclusion that x converges to $\hat{x}$ for $x \in (\hat{x}, a_C]$. Since $\lim_{t\to\infty} G^t(x) = \{\hat{x}\}$ for any $x \in (\hat{x}, a_C]$, $\lim_{t\to\infty} G^t(x) = \{\hat{x}\}$ for any $x \in G((\hat{x}, a_C]) = [(1-d)a_C, \hat{x}]$. Further, since $G([a_C, \hat{x}/(1-d))) = [(1-d)a_C, \hat{x}]$, the system must converge for any x in $[a_C, \hat{x}/(1-d))$. Since we know that $h(x) = \{(1-d)x\}$ for x in $[\hat{x}/(1-d), \infty)$, for any x in $[\hat{x}/(1-d), \infty)$, after finite periods, the stock must enter the region $([\hat{x}, \hat{x}/(1-d))$, thus leading to convergence. So far we have shown that the system converges for any x in $[(1-d)a_C, \infty)$. According to Proposition 1, for any x less than $\hat{x}$, after finite periods, the stock must be above $\hat{x}$, again according to what we have shown, leading to convergence.

We now have shown for any x, the optimal policy leads to a convergence to the golden rule stock.

Proof of Proposition 3:

Consider the subregion $(0, a_I]$.

Suppose on the contrary there exists $x \in (0, a_I]$ and $z \in h(x)$ such that $z \neq \theta x$. Then we must have $z < \theta x$. By the optimality of z,

$$V(x) = u(x,z) + \rho V(z) \geq u(x, \theta x) + \rho V(\theta x).$$

Rearranging the equation, we have

$$u(x,z) - u(x,\theta x) \geq \rho(V(\theta x) - V(z)) \geq \rho V'_{-}(\theta x)(\theta x - z) \geq \rho V'_{-}(\hat{x})(\theta x - z) \geq \rho\hat{p}(\theta x - z),$$

where the second inequality follows from concavity of V, the third inequality follows from concavity of V and the fact that $\theta a_I \leq \hat{x}$ for $\zeta \leq 0$ (Lemma 5), and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \leq a_I$, we have

$$u(x,z) - u(x,\theta x) = (1/a_C)[x - (a_I/b)(z - (1-d)x)] = \frac{a_I}{ba_C}(\theta x - z) < \hat{p}\rho(\theta x - z),$$

where the last inequality follows from Lemma 2. Since we have shown that $u(x,z) - u(x,\theta x) \geq \rho\hat{p}(\theta x - z)$. This leads to a contradiction and establishes the desired result.

Consider the subregion $[a_C, \infty)$.

Suppose on the contrary there exists $x \geq a_C$ and $z \in h(x)$ such that $z \neq (1-d)x$. Then we must have $z > (1-d)x$. By the optimality of z,

$$V(x) = u(x,z) + \rho V(z) \geq u(x,(1-d)x) + \rho V((1-d)x).$$

Rearranging the equation, we have

$$\begin{aligned} u(x,(1-d)x) - u(x,z) &\leq \rho(V(z) - V((1-d)x)) \leq \rho V'_+((1-d)x)(z-(1-d)x) \\ &\leq \rho V'_+(\hat{x})(z-(1-d)x) \leq \rho\hat{p}(z-(1-d)x), \end{aligned}$$

where the second inequality follows from concavity of V, the third inequality follows from concavity of V and the fact that $\hat{x} \leq (1-d)a_C \leq (1-d)x$ (Lemma 5), and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \geq a_C$, we have

$$u(x,(1-d)x) - u(x,z) = (z-(1-d)x)/b > \hat{p}\rho(z-(1-d)x),$$

where the last inequality follows from Lemma 2. Since we have shown that $u(x,(1-d)x) - u(x,z) \leq \rho\hat{p}(z-(1-d)x)$. This leads to a contradiction and establishes the desired result.

Consider the subregion $(a_I, \hat{x}]$.

Suppose on the contrary there exists $x \in (a_I, \hat{x}]$ and $z \in h(x)$ such that $z \notin [\zeta(\hat{x}-x)+\hat{x}, \hat{x}]$. There are two possible cases: (i) $z > \hat{x}$; (ii) $z < \zeta(\hat{x}-x)+\hat{x}$.

Consider (i) $z > \hat{x}$. We have

$$V(x) = u(x,z) + \rho V(z) \geq u(x,\hat{x}) + \rho V(\hat{x}).$$

Rearranging the equation, we have

$$u(x,\hat{x}) - u(x,z) \leq \rho(V(z) - V(\hat{x})) \leq \rho V'_+(\hat{x})(z-\hat{x}) \leq \rho\hat{p}(z-\hat{x}),$$

where the second inequality follows from concavity of V and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (a_I, \hat{x}]$ and $\zeta \leq 0$, we have

$$u(x,\hat{x}) - u(x,z) \geq (1/b)(z-\hat{x}) > \hat{p}\rho(z-\hat{x}),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x,\hat{x}) - u(x,z) \leq \rho\hat{p}(z-\hat{x})$. This leads to a contradiction and establishes the desired result.

Consider (ii) $z < \zeta(\hat{x} - x) + \hat{x}$. We have

$$V(x) = u(x, z) + \rho V(z) \geq u(x, \zeta(\hat{x} - x) + \hat{x}) + \rho V(\zeta(\hat{x} - x) + \hat{x}).$$

Rearranging the equation, we have

$$\begin{aligned} u(x, z) - u(x, \zeta(\hat{x} - x) + \hat{x}) &\geq \rho(V(\zeta(\hat{x} - x) + \hat{x}) - V(z)) \\ &\geq \rho V'_-(\zeta(\hat{x} - x) + \hat{x})((\zeta(\hat{x} - x) + \hat{x}) - z) \\ &\geq \rho V'_-(\hat{x})((\zeta(\hat{x} - x) + \hat{x}) - z) \geq \rho\hat{p}((\zeta(\hat{x} - x) + \hat{x}) - z) \end{aligned}$$

where the second inequality follows from concavity of V, the third inequality follows from concavity of V and the fact that $\zeta(\hat{x} - x) + \hat{x} \leq \hat{x}$ for $x \in (a_I, \hat{x}]$, and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (a_I, \hat{x}]$, we have

$$u(x, z) - u(x, \zeta(\hat{x} - x) + \hat{x}) \leq \frac{a_I}{a_C b}((\zeta(\hat{x} - x) + \hat{x}) - z) < \hat{p}\rho((\zeta(\hat{x} - x) + \hat{x}) - z),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x, z) - u(x, (\zeta(\hat{x} - x) + \hat{x})) \geq \rho\hat{p}((\zeta(\hat{x} - x) + \hat{x}) - z)$. This leads to a contradiction and establishes the desired result.

Last, consider the subregion $(\hat{x}, a_C)$.

Suppose on the contrary there exists $x \in (\hat{x}, a_C)$ and $z \in h(x)$ such that $z \notin [\hat{x}, \zeta(\hat{x} - x) + \hat{x}]$. There are two possible cases: (i) $z < \hat{x}$; (ii) $z > \zeta(\hat{x} - x) + \hat{x}$.

Consider (i) $z < \hat{x}$. We have

$$V(x) = u(x, z) + \rho V(z) \geq u(x, \hat{x}) + \rho V(\hat{x}).$$

Rearranging the equation, we have

$$u(x, z) - u(x, \hat{x}) \geq \rho(V(\hat{x}) - V(z)) \geq \rho V'_-(\hat{x})(\hat{x} - z) \geq \rho\hat{p}(\hat{x} - z),$$

where the second inequality follows from concavity of V and the third inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (\hat{x}, a_C)$ and $\zeta \leq 0$, we have

$$u(x, z) - u(x, \hat{x}) \leq \frac{a_I}{b a_C}(\hat{x} - z) < \hat{p}\rho(\hat{x} - z),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x, z) - u(x, \hat{x}) \geq \rho\hat{p}(\hat{x} - z)$. This leads to a contradiction and establishes the desired result.

Consider (ii) $z > \zeta(\hat{x} - x) + \hat{x}$. We have

$$V(x) = u(x, z) + \rho V(z) \geq u(x, \zeta(\hat{x} - x) + \hat{x}) + \rho V(\zeta(\hat{x} - x) + \hat{x}).$$

Rearranging the equation, we have

$$\begin{aligned} u(x,\zeta(\hat{x}-x)+\hat{x})-u(x,z) &\leq \rho(V(z)-V(\zeta(\hat{x}-x)+\hat{x})) \\ &\leq \rho V'_+(\zeta(\hat{x}-x)+\hat{x})(z-(\zeta(\hat{x}-x)+\hat{x})) \\ &\leq \rho V'_+(\hat{x})(z-(\zeta(\hat{x}-x)+\hat{x})) \leq \rho\hat{p}(z-(\zeta(\hat{x}-x)+\hat{x})) \end{aligned}$$

where the second inequality follows from concavity of V, the third inequality follows from concavity of V and the fact that $\zeta(\hat{x}-x)+\hat{x} \geq \hat{x}$ for $x \in (\hat{x}, a_C)$, and the last inequality follows from Lemma 4 ii).

By the definition of u, and given that $x \in (\hat{x}, a_C)$ we have

$$u(x,\zeta(\hat{x}-x)+\hat{x})-u(x,z) \geq (1/b)(z-(\zeta(\hat{x}-x)+\hat{x})) > \hat{p}\rho(z-(\zeta(\hat{x}-x)+\hat{x})),$$

where the last inequality follows from Lemma 2.

Since we have shown that $u(x,(\zeta(\hat{x}-x)+\hat{x}))-u(x,z) \leq \rho\hat{p}(z-(\zeta(\hat{x}-x)+\hat{x}))$. This leads to a contradiction and establishes the desired result.

Proof of Proposition 4:

It has been covered in Proposition 3 for $x \in (0, a_I] \cup [a_C, \infty)$. We only need to consider $x \in (a_I, a_C)$. Suppose x is in $(a_I, \hat{x}]$. Since $\zeta > -1$ and $\zeta \leq 0$, we have $x < \zeta(\hat{x}-x)+\hat{x} \leq \hat{x}$, which implies $\zeta(\hat{x}-x)+\hat{x} \in (a_I, a_C)$. Similarly, if $x \in (\hat{x}, a_C)$, we must have $\zeta(\hat{x}-x)+\hat{x} \in (a_I, a_C)$. This suggests that for any $x \in (a_I, a_C)$, if we follow the policy such that $x' = \zeta(\hat{x}-x)+\hat{x}$, the stock of the next period is also in (a_I, a_C) and therefore, the policy that fully utilizes resources for $x \in (a_I, a_C)$ leads to zero total value loss. Any deviation from this policy leads to a positive value loss for $x \in (a_I, a_C)$, and therefore, it is not optimal. Hence, according to Lemma 8, $h(x) = \{\zeta(\hat{x}-x)+\hat{x}\}$ for $x \in (a_I, a_C)$.

Proof of Proposition 5:

Since ζ is in $(0,1]$, or more explicitly, $b(a_C - a_I) - (1-d) \leq 1$, rearranging the terms, we must have $(a_C - \theta a_I) + ((1-d)a_C - a_I) \geq 0$, which implies that at least one of the following two inequalities holds: (A) $a_C \geq \theta a_I$; (B) $(1-d)a_C \geq a_I$. Therefore, we consider three possible cases.

(i) Both (A) and (B) hold: $a_C \geq \theta a_I$ and $(1-d)a_C \geq a_I$.

This is the simplest case. Consider $x \in [a_I, a_C]$. Since $f(x) \equiv \zeta(\hat{x}-x)-\hat{x}) \in [a_C(1-d), a_I\theta] \subset [a_I, a_C]$, the sequence of the capital stock generated by f, $\{f^t(x)\}_{t=1}^{\infty}$, is bounded by $[a_I, a_C]$. Further, since we know from Lemma 1 that the value loss associated with $(x, f(x))$ is zero for $x \in [a_I, a_C]$, the sum of the discounted value losses associated with $\{f^t(x)\}_{t=1}^{\infty}$ is zero. Stating from x, any program that deviates from $\{f^t(x)\}_{t=1}^{\infty}$ yields a positive value loss. According to Lemma 8, $h(x) = \{\zeta(\hat{x}-x)-\hat{x})\}$ for $x \in [a_I, a_C]$.

Now consider $x \in (\hat{x}/\theta, a_I)$. According to Proposition 1, we know $h(x) \subset [\hat{x}, \theta x] \subset [\hat{x}, \theta a_I] \subset [\hat{x}, a_C]$. Since we know that the total value loss for the optimal program starting from $x \in [a_I, a_C]$ is always zero, we just need to check the one-period value loss for (x, x')

with $x \in (\hat{x}/\theta, a_I)$ and $x' \in [\hat{x}, \theta x]$:

$$\begin{aligned}\delta^\rho(x,x') &= u(\hat{x},\hat{x}) + (\rho-1)\hat{p}\hat{x} - u(x,x') - \hat{p}(\rho x' - x) \\ &= u(\hat{x},\hat{x}) + (\rho-1)\hat{p}\hat{x} - (1/a_C)(x - (a_I/b)(x' - (1-d)x)) - \hat{p}(\rho x' - x).\end{aligned}$$

Then we have

$$\frac{\partial \delta^\rho(x,x')}{\partial x'} = \frac{a_I}{a_C b} - \hat{p}\rho < 0,$$

where the inequality follows from Lemma 2. Since the one-period value loss strictly decreases with x', it attains its unique minimum and therefore the total value loss attains its unique minimum, when x' attains its unique maximum, which implies that $h(x) = \{\theta x\}$ for $x \in (\hat{x}/\theta, a_I)$.

Consider $x \in (a_C, \hat{x}/(1-d))$. According to Proposition 1, we know $h(x) \subset [(1-d)x, \hat{x}] \subset [(1-d)a_C, \hat{x}] \subset [a_I, \hat{x}]$. Since we know that the total value loss for the optimal program starting from $x \in [a_I, a_C]$ is always zero, we just need to check the one-period value loss for (x, x') with $x \in (a_C, \hat{x}/(1-d))$ and $x' \in [(1-d)x, \hat{x}]$:

$$\delta^\rho(x,x') = u(\hat{x},\hat{x}) + (\rho-1)\hat{p}\hat{x} - (1 - (1/b)(x' - (1-d)x)) - \hat{p}(\rho x' - x).$$

Then we have

$$\frac{\partial \delta^\rho(x,x')}{\partial x'} = \frac{1}{b} - \hat{p}\rho > 0,$$

where the inequality follows from Lemma 2. Since the one-period value loss strictly increases with x', it attains its unique minimum and therefore the total value loss attains its unique minimum, when x' attains its unique minimum, which implies $h(x) = \{(1-d)x\}$ for $x \in (a_C, \hat{x}/(1-d))$.

Combined with the characterization for $x \in (0, \hat{x}/\theta] \cup [\hat{x}/(1-d), \infty)$ as in Proposition 1, we have obtained the desired result for case (i).

(ii) Only (B) holds: $a_C < \theta a_I$ and $(1-d)a_C \geq a_I$.

The complication arises from the fact that $a_C < a_I\theta$. As $a_C < a_I\theta$, $f(a_I) = \zeta(\hat{x} - a_I) + \hat{x} = a_I\theta > a_C$, which means, $f(a_I) \notin [a_I, a_C]$. The total value loss could be strictly positive even if we follow the policy f with an initial stock starting from a_I.

Consider $x \in [\hat{x}, a_C]$. Since $a_C(1-d) \geq a_I$, $f(x) = \zeta(\hat{x} - x) + \hat{x} \in [(1-d)a_C, \hat{x}] \subset [a_I, \hat{x}]$. Since $f(x) \in [a_I, \hat{x}]$, $f^2(x) = \zeta^2 x + (1-\zeta^2)\hat{x} \in [\hat{x}, x] \subset [\hat{x}, a_C]$, where $\zeta^2 x + (1-\zeta^2)\hat{x} \leq x$ follows from $\zeta \in (0,1]$ and $x \geq \hat{x}$. Therefore, $\{f^t(x)\}_{t=1}^\infty$, is bounded by $[a_I, a_C]$. It follows from the argument for case (i) that $h(x) = \{\zeta(\hat{x} - x) + \hat{x}\}$ for $x \in [\hat{x}, a_C]$.

Consider $x \in [a_C(\zeta - d)/\zeta, \hat{x})$. Since $a_C < a_I\theta$, $a_C(\zeta - d)/\zeta > a_I$. Since $f(x) \in (\hat{x}, a_C]$ with $\delta^\rho(x, f(x)) = 0$ and we have shown that the optimal policy function leads to the total value loss being zero for any initial stock in $(\hat{x}, a_C]$, we must have $h(x) = \{\zeta(\hat{x} - x) + \hat{x}\}$ for $x \in [a_C(\zeta - d)/\zeta, \hat{x})$.

Consider $x \in (a_C, \hat{x}/(1-d))$. According to Proposition 1, we know $h(x) \subset [(1-d)x, \hat{x}] \subset [(1-d)a_C, \hat{x}] \subset [a_C(\zeta - d)/\zeta, \hat{x}]$, where $[(1-d)a_C, \hat{x}] \subset [a_C(\zeta - d)/\zeta, \hat{x}]$ follows from $\zeta \in (0,1]$. Then it follows from the argument for case (i) that $h(x) = \{(1-d)x\}$ for $x \in (a_C, \hat{x}/(1-d))$.

Consider $x \in (\hat{x}/\theta, a_C/\theta]$. Since $a_C < a_I\theta$, $a_C/\theta < a_I$. According to Proposition 1, we know $h(x) \subset [\hat{x}, \theta x] \subset [\hat{x}, a_C]$. Again, it follows from the argument for case (i) that $h(x) = \{\theta x\}$ for $x \in (\hat{x}/\theta, a_C/\theta]$.

Last, consider $x \in (a_C/\theta, a_C(\zeta - d)/\zeta)$.According to Proposition 1, $h(x) \subset [\hat{x}, \min\{\theta x, \zeta(\hat{x} - x) + \hat{x}\}]$. Since x is in $(a_C/\theta, a_C(\zeta - d)/\zeta)$, we have $[\hat{x}, a_C] \subset [\hat{x}, \min\{\theta x, \zeta(\hat{x} - x) + \hat{x}\}]$. Let $x' \in [\hat{x}, \min\{\theta x, \zeta(\hat{x} - x) + \hat{x}\}]$. If $x' \le a_C$, then the total value loss is simply the one period value loss $\delta^\rho(x, x')$. Following the argument for case (i), the one period value loss is minimized when x' attains its maximum, a_C. Hence, we must have $h(x) \subset [a_C, \min\{\theta x, \zeta(\hat{x} - x) + \hat{x}\}]$.

Combined with the characterization for $x \in (0, \hat{x}/\theta] \cup [\hat{x}/(1-d), \infty)$ in Proposition 1, we have obtained the desired result for case (ii).

(iii) Only (A) holds: $a_C \ge \theta a_I$ and $(1-d)a_C < a_I$.

The complication for this case arises from the fact that $a_C(1-d) < a_I$. As $a_C(1-d) < a_I$, $f(a_C) = \zeta(\hat{x} - a_C) + \hat{x} = (1-d)a_C < a_I$, which means $f(a_C) \notin [a_I, a_C]$. The total value loss could be strictly positive even if we follow the policy f with an initial stock starting from a_C.

Consider $x \in [a_I, \hat{x}]$. Since $a_C \ge a_I\theta$, it follows symmetrically from the argument for $[\hat{x}, a_C]$ in case (ii) that $h(x) = \{\zeta(\hat{x} - x) + \hat{x}\}$ for $x \in [a_I, \hat{x}]$.

Consider $x \in (\hat{x}, a_C(1+(1-d)/\zeta) - a_I/\zeta]$. Since $a_C(1-d) < a_I$, $a_C(1+(1-d)/\zeta) - a_I/\zeta < a_C$. Then it follows symmetrically from the argument for $[a_C(\zeta-d)/\zeta, \hat{x})$ in case (ii) that $h(x) = \{\zeta(\hat{x} - x) + \hat{x}\}$ for $x \in (\hat{x}, a_C(1 + (1-d)/\zeta) - a_I/\zeta]$.

Consider $x \in (\hat{x}/\theta, a_I)$. According to Proposition 1, $h(x) \subset [\hat{x}, \theta x] \subset [\hat{x}, \theta a_I] \subset [\hat{x}, a_C(1+(1-d)/\zeta) - a_I/\zeta]$, where the last $\subset$ holds because $\theta a_I \le a_C(1+(1-d)/\zeta) - a_I/\zeta$, which itself follows from $f(\theta a_I) \ge a_I$ (due to $\zeta \le 1$), $f(a_C(1 + (1-d)/\zeta) - a_I/\zeta) = a_I$, and f being decreasing. Then it follows from the argument for case (i) that $h(x) = \{\theta x\}$ for $x \in (\hat{x}/\theta, a_I)$.

Consider $x \in (a_I/(1-d), \hat{x}/(1-d))$. Since $a_C(1-d) < a_I$, $a_I/(1-d) > a_C$. According to Proposition 1, $h(x) \subset [(1-d)x, \hat{x}] \subset [a_I, \hat{x}]$. It then follows from the argument for case (i) that $h(x) = \{(1-d)x\}$ for $x \in (a_I/(1-d), \hat{x}/(1-d))$.

Last, consider $x \in (a_C(1+(1-d)/\zeta) - a_I/\zeta, a_I/(1-d))$. According to Proposition 1, $h(x) \subset [\max\{(1-d)x, \zeta(\hat{x}-x)+\hat{x}\}, \hat{x}]$. Since x is in $(a_C(1+(1-d)/\zeta) - a_I/\zeta, a_I/(1-d))$, we have $[a_I, \hat{x}] \subset [\max\{(1-d)x, \zeta(\hat{x} - x) + \hat{x}\}, \hat{x}]$. If $x' \ge a_I$, then total value loss is simply the one period value loss. Following the argument for case (i), the one period value loss is minimized when x' attains its minimum, a_I. Then we must have $h(x) \subset [\max\{(1-d)x, \zeta(\hat{x} - x) + \hat{x}\}, a_I]$.

Combined with the characterization for $x \in (0, \hat{x}/\theta] \cup [\hat{x}/(1-d), \infty)$ in Proposition 1, we have obtained the desired result for case (iii).

Proof of Proposition 2:

We first show the first part of the proposition concerning the definition and the order of $\bar{\rho}_t$. Let $f_t(\rho) \equiv a_C b(1-d)^t \rho^{t+1} - (a_C - a_I)a_I\zeta\rho - a_I(a_C - a_I)$. Since $f_t(0) < 0$ and $f_t(\rho) > 0$ for ρ sufficiently large, there must exist at least one positive root to the equation $f_t(\rho) = 0$. Suppose there are two different roots, denoted by ρ_1 and ρ_2. Without loss of generality, let $\rho_1 > \rho_2$. Then we have

$$\begin{aligned} a_C b(1-d)^t \rho_1^{t+1} - (a_C - a_I)a_I\zeta\rho_1 - a_I(a_C - a_I) &= 0 \\ a_C b(1-d)^t \rho_2^{t+1} - (a_C - a_I)a_I\zeta\rho_2 - a_I(a_C - a_I) &= 0, \end{aligned}$$

which implies

$$\begin{aligned} & a_C b(1-d)^t(\rho_1^{t+1} - \rho_2^{t+1}) = (a_C - a_I)a_I\zeta(\rho_1 - \rho_2) \\ \Leftrightarrow \quad & (a_C - a_I)a_I\zeta = \frac{a_C b(1-d)^t(\rho_1^{t+1} - \rho_2^{t+1})}{\rho_1 - \rho_2} > a_C b(1-d)^t \rho_2^t, \end{aligned}$$

where the last equality follows from $\rho_1 > \rho_2$. Since $(a_C - a_I)a_I\zeta > a_C b(1-d)^t\rho_2^t$, $a_C b(1-d)^t \rho_2^{t+1} - (a_C - a_I)a_I\zeta\rho_2 - a_I(a_C - a_I) < 0$, leading to the contradiction. Hence, $\bar{\rho}_t$ is the unique positive root, being well-defined. Further, since $f_1(1/\theta) = ba_C(1-d-\theta)/\theta^2 < 0$ and we know $f_1(\rho)$ is positive for ρ sufficiently large, $\bar{\rho}_1 > 1/\theta$. Since $f_t(1/(1-d)) = b(a_C - a_I)/(1-d) > 0$ and we know $f_t(0) < 0$, $\bar{\rho}_t < 1/(1-d)$ for any t.

By definition, we have

$$\begin{aligned} f_{t+1}(\bar{\rho}_{t+1}) = 0 \Leftrightarrow a_C b(1-d)^{t+1}\bar{\rho}_{t+1}^{t+2} - (a_C - a_I)a_I\zeta\bar{\rho}_{t+1} - a_I(a_C - a_I) = 0 \\ f_t(\bar{\rho}_t) = 0 \Leftrightarrow a_C b(1-d)^t \bar{\rho}_t^{t+1} - (a_C - a_I)a_I\zeta\bar{\rho}_t - a_I(a_C - a_I) = 0. \end{aligned}$$

Since $\bar{\rho}_{t+1} < 1/(1-d)$, or equivalently, $\bar{\rho}_{t+1}(1-d) < 1$,

$$f_{t+1}(\bar{\rho}_t) = a_C b(1-d)^{t+1}\bar{\rho}_t^{t+2} - (a_C - a_I)a_I\zeta\bar{\rho}_t - (a_C - a_I)a_I < f_t(\bar{\rho}_t) = 0.$$

Further, we know $f_t(\rho) > 0$ for ρ sufficiently large, so $\bar{\rho}_{t+1} > \bar{\rho}_t$.

Now we turn to characterizing the optimal policy correspondence for $x \in (a_C/\theta, a_C(\zeta - d)/\zeta)$.

Pick the smallest integer t_0 such that $a_I\theta(1-d)^{t_0} < a_C$. By construction, $t_0 \geq 1$ and $a_I\theta(1-d)^{t_0-1} \geq a_C$, so $a_I\theta(1-d)^{t_0} \geq (1-d)a_C \geq a_C(\zeta-d)/\zeta$, where the last inequality follows from $0 < \zeta \leq 1$.

Pick $x \in (a_C/\theta, a_C(\zeta-d)/\zeta)$. According to case (ii) in Proposition 5, the stock for the next period, x', has to be in $[a_C, \min\{\zeta(\hat{x}-x)+\hat{x}, \theta x\}]$, so $x' \leq a_I\theta$. Pick the smallest integer t_1 such that $(1-d)^{t_1}x' < a_C$. Since $x' \leq a_I\theta$, by construction, $1 \leq t_1 \leq t_0$ and $(1-d)^{t_1-1}x' \geq a_C$, so $(1-d)^{t_1}x' \geq (1-d)a_C \geq a_C(\zeta-d)/\zeta$.

For any stock above a_C, notice that the optimality mandates the stock in the following period to shirk by $(1-d)$ times. Following x', the stock for the next t_1 periods are given by $\{(1-d)^t x'\}_{t=1}^{t_1}$. Since $(1-d)^{t_1}x' \subset [a_C(\zeta-d)/\zeta, a_C)$, after t_1+1 periods, the total value loss of the remaining periods will be zero, so we focus on the total value loss for the first t_1+1 periods.

Consider the (t_1+1)-period value loss associated with (x, x') and $\{((1-d)^t x', (1-d)^{t+1}x')\}_{t=0}^{t=t_1-1}$.

$$\begin{aligned}\ell_{t_1}(x') &\equiv \delta^\rho(x,x') + \sum_{t=0}^{t_1-1}\rho^{t+1}\delta^\rho((1-d)^t x', (1-d)^{t+1}x') \\ &= \frac{1-\rho^{t_1}}{1-\rho}u(\hat{x},\hat{x}) + (\rho^{t_1}-1)\hat{p}\hat{x} - (1/a_C)(x-(a_I/b)(x'-(1-d)x)) \\ &\quad -\frac{\rho-\rho^{t_1+1}}{1-\rho} - \hat{p}(\rho^{t_1+1}(1-d)^{t_1}x' - x)\end{aligned}$$

Then we have

$$\frac{\partial \ell_{t_1}(x')}{\partial x'} = \frac{a_I}{ba_C} - \hat{p}\rho^{t_1+1}(1-d)^{t_1} = \frac{a_I}{ba_C} - \frac{\rho^{t_1+1}(1-d)^{t_1}}{(a_C-a_I)(1+\rho\zeta)} = \frac{-f_{t_1}(\rho)}{ba_C(a_C-a_I)(1+\rho\zeta)}.$$

By construction of $\bar{\rho}_{t_1}$, we know $\partial\ell_{t_1}(x')/\partial x' > 0$ if $\rho < \bar{\rho}_{t_1}$; $\partial\ell_{t_1}(x')/\partial x' = 0$ if $\rho = \bar{\rho}_{t_1}$; $\partial\ell_{t_1}(x')/\partial x' < 0$ if $\rho > \bar{\rho}_{t_1}$.

Consider two possible cases: (1) $t_0 = 1$; (2) $t_0 > 1$.

For (1), $t_0 = 1$, so we must have $t_1 = 1$. Hence, we only need to consider the two-period value loss. If $\rho > \bar{\rho}_1$, the total value loss attains its minimum when x' attains its maximum, suggesting that $h(x) = \min\{\zeta(\hat{x}-x)+\hat{x}, \theta x\}$. If $\rho = \bar{\rho}_1$, then the total value loss is constant with respect to x', so $h(x) = [a_C, \min\{\zeta(\hat{x}-x)+\hat{x}, \theta x\}]$. If $\rho < \bar{\rho}_1$, the total value loss attains its minimum when x' attains its minimum, which implies that $h(x) = \{a_C\}$.

For (2), $(1/\theta, \infty)$ is partitioned by $\{\bar{\rho}_t\}_{t=1}^{t_0}$: $(1/\theta, \bar{\rho}_1)$, $\{\bar{\rho}_1\}$, $(\bar{\rho}_1, \bar{\rho}_2)$,..., $(\bar{\rho}_{t_0-1}, \bar{\rho}_{t_0})$, $\{\bar{\rho}_{t_0}\}$, and $(\bar{\rho}_{t_0}, \infty)$.

Consider $\rho < \bar{\rho}_1$. The two-period value loss is minimized and total value loss is equal to the two-period value loss when $x' = a_C$, so $h(x) = \{a_C\}$.

Consider $\rho = \bar{\rho}_{t_1}$ for t_1 taking value from $\{1, 2, ..., t_0\}$. The (t_1+1)-period value loss and also the total value loss is constant with respect to x' for a fixed t_1. Since $\rho = \bar{\rho}_{t_1}$, $\rho > \bar{\rho}_{t'_1}$ for any $t'_1 < t_1$ and $\rho < \bar{\rho}_{t'_1}$ for any $t'_1 > t_1$. Since $\rho > \bar{\rho}_{t'_1}$ for any $t'_1 < t_1$, the total value loss decreases with x' for $t'_1 < t_1$, or equivalently, for $x'(1-d)^{t_1-1} < a_C$.[27] Since $\rho < \bar{\rho}_{t'_1}$ for any $t'_1 > t_1$, the total value loss increases with x' for $t'_1 > t_1$, or equivalently, for $x'(1-d)^{t_1} \geq a_C$. If $\min\{\zeta(\hat{x}-x)+\hat{x}, \theta x\} > a_C/(1-d)^{t_1}$, then $h(x) = [a_C/(1-d)^{t_1-1}, a_C/(1-d)^{t_1}]$. If $\min\{\zeta(\hat{x}-x)+\hat{x}, \theta x\} \in [a_C/(1-d)^{t_1-1}, a_C/(1-d)^{t_1}]$, then $h(x) = [a_C/(1-d)^{t_1-1}, \min\{\zeta(\hat{x}-x)+\hat{x}, \theta x\}]$. If $\min\{\zeta(\hat{x}-x)+\hat{x}, \theta x\} < a_C/(1-d)^{t_1}$, then $h(x) = \min\{\zeta(\hat{x}-x)+\hat{x}, \theta x\}$. In sum, for $\rho = \bar{\rho}_{t_1}$, $h(x) = [\min\{\zeta(\hat{x}-x)+\hat{x}, \theta x, a_C/(1-d)^{t-1}\}, \min\{\zeta(\hat{x}-x)+\hat{x}, \theta x, a_C/(1-d)^t\}]$.

Consider $\rho \in (\bar{\rho}_{t_1}, \bar{\rho}_{t_1+1})$ for t_1 taking value from $\{1, ..., t_0-1\}$. The (t_1+1)-period value loss and also the total value loss is minimized when x' attains its maximum

[27] Here we implicitly rely on the continuity of the value function.

for a fixed t_1. Since $\rho < \bar{\rho}_{t_1+1}$, $\rho < \bar{\rho}_{t'_1}$ for any $t'_1 > t_1$, which implies that the total value loss increases with x' for $t'_1 > t_1$, or equivalently, for $x'(1-d)^{t_1} \geq a_C$. Since $\rho > \bar{\rho}_{t_1}$, $\rho > \bar{\rho}_{t'_1}$ for any $t'_1 < t_1$, which implies that the total value loss decreases with x' for $t'_1 < t_1$, or equivalently, for $x'(1-d)^{t_1-1} < a_C$. Hence, we have $h(x) = \min\{\zeta(\hat{x}-x)+\hat{x}, \theta x, a_C/(1-d)^{t_1}\}$.

Last, consider $\rho > \bar{\rho}_{t_0}$. Since we know $\bar{\rho}_{t_0} \geq \bar{\rho}_t$ for any $t = 1, 2, ..., t_0$, $\rho > \bar{\rho}_{t_1}$ for any t_1. This suggests that the (t_1+1)-period value loss and also the total value loss decreases with x' for *any* given t_1. Then the total value loss is minimized when x' attains its maximum. Hence, $h(x) = \min\{\zeta(\hat{x}-x)+\hat{x}, \theta x\}$.

We have now obtained the desired conclusion.

Proof of Proposition 3:

We first show the first part of the proposition concerning the definition and the order of $\tilde{\rho}_t$. Let $f_t(\rho) \equiv b\theta^t\rho^{t+1} - (a_C - a_I)\zeta\rho - (a_C - a_I)$. Since $f_t(0) < 0$ and $f_t(\rho) > 0$ for ρ sufficiently large, there must exist at least one positive root to the equation $f_t(\rho) = 0$. Suppose there are two different roots, denoted by ρ_1 and ρ_2. Without loss of generality, let $\rho_1 > \rho_2$. Then we have

$$\begin{aligned} b\theta^t\rho_1^{t+1} - (a_C - a_I)\zeta\rho_1 - (a_C - a_I) &= 0 \\ b\theta^t\rho_2^{t+1} - (a_C - a_I)\zeta\rho_2 - (a_C - a_I) &= 0, \end{aligned}$$

which implies

$$b\theta^t(\rho_1^{t+1} - \rho_2^{t+1}) = (a_C - a_I)\zeta(\rho_1 - \rho_2) \Leftrightarrow (a_C - a_I)\zeta = \frac{b\theta^t(\rho_1^{t+1} - \rho_2^{t+1})}{\rho_1 - \rho_2} > b\theta^t\rho_2^t,$$

where the last equality follows from $\rho_1 > \rho_2$. Since $(a_C - a_I)\zeta > b\theta^t\rho_2^t$, $b\theta^t\rho_2^{t+1} - (a_C - a_I)\zeta\rho_2 - (a_C - a_I) < 0$, leading to the contradiction. Hence, $\tilde{\rho}_t$ is the unique positive root, being well-defined. Further, since $f_t(1/\theta) = -b(a_C - a_I)/(a_I\theta) < 0$ and we know $f_t(\rho)$ is positive for ρ sufficiently large, $\tilde{\rho}_t > 1/\theta$.

By definition, we have

$$\begin{aligned} f_{t+1}(\tilde{\rho}_{t+1}) = 0 &\Leftrightarrow b\theta^{t+1}\tilde{\rho}_{t+1}^{t+2} - (a_C - a_I)\zeta\tilde{\rho}_{t+1} - (a_C - a_I) = 0 \\ f_t(\tilde{\rho}_t) = 0 &\Leftrightarrow b\theta^t\tilde{\rho}_t^{t+1} - (a_C - a_I)\zeta\tilde{\rho}_t - (a_C - a_I) = 0. \end{aligned}$$

Since $\tilde{\rho}_{t+1} > 1/\theta$, or equivalently, $\tilde{\rho}_{t+1}\theta > 1$,

$$f_t(\tilde{\rho}_{t+1}) = b\theta^t\tilde{\rho}_{t+1}^{t+1} - (a_C - a_I)\zeta\tilde{\rho}_{t+1} - (a_C - a_I) < f_{t+1}(\tilde{\rho}_{t+1}) = 0.$$

Further, we know $f_t(\rho) > 0$ for ρ sufficiently large, so $\tilde{\rho}_{t+1} < \tilde{\rho}_t$.

Now we turn to characterizing the optimal policy correspondence for $x \in (a_C(1 + (1-d)/\zeta) - a_I/\zeta, a_I/(1-d))$.

Pick the smallest integer t_0 such that $\theta^{t_0}a_C(1-d) > a_I$. By construction, $t_0 \geq 1$ and $\theta^{t_0-1}a_C(1-d) \leq a_I$, so $\theta^{t_0}a_C(1-d) \leq \theta a_I \leq a_C(1+(1-d)/\zeta) - a_I/\zeta$, where the last inequality follows from $\zeta \leq 1$ and $a_I\theta \leq a_C$ (also see the proof for Proposition 5).

Pick $x \in (a_C(1+(1-d)/\zeta)-a_I/\zeta, a_I/(1-d))$. According to case (iii) in Proposition 5, the stock for the next period, x', has to be in $[\max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x\}, a_I]$, so $x' \geq a_C(1-d)$. Pick the smallest integer t_1 such that $\theta^{t_1}x' > a_I$. Since $x' \geq a_C(1-d)$, by construction, $1 \leq t_1 \leq t_0$ and $\theta^{t_1-1}a_C(1-d) \leq a_I$, so $\theta^{t_1}a_C(1-d) \leq \theta a_I \leq a_C(1+(1-d)/\zeta)-a_I/\zeta$. For any stock below a_I, notice that the optimality mandates the stock in the following period to grow up by θ times. Following x', the stock for the next t_1 periods are given by $\{\theta^t x'\}_{t=1}^{t_1}$. Since $\theta^{t_1}x' \subset (a_I, a_C(1+(1-d)/\zeta)-a_I/\zeta]$, after t_1+1 periods, the total value loss of the remaining periods will be zero, so we focus on the total value loss for the first t_1+1 periods.

Consider the (t_1+1)-period value loss associated with (x, x') and $\{(\theta^t x', \theta^{t+1}x')\}_{t=0}^{t=t_1-1}$.

$$
\begin{aligned}
\ell_{t_1}(x') &\equiv \delta^\rho(x,x') + \sum_{t=0}^{t_1-1}\rho^{t+1}\delta^\rho(\theta^t x', \theta^{t+1}x') \\
&= \frac{1-\rho^{t_1}}{1-\rho}u(\hat{x},\hat{x}) + (\rho^{t_1}-1)\hat{p}\hat{x} - (1-(1/b)(x'-(1-d)x)) - \hat{p}(\rho^{t_1+1}\theta^{t_1}x'-x)
\end{aligned}
$$

Then we have

$$
\frac{\partial \ell_{t_1}(x')}{\partial x'} = \frac{1}{b} - \hat{p}\rho^{t_1+1}\theta^{t_1} = \frac{1}{b} - \frac{\rho^{t_1+1}\theta^{t_1}}{(a_C-a_I)(1+\rho\zeta)} = \frac{-f_{t_1}(\rho)}{b(a_C-a_I)(1+\rho\zeta)}.
$$

By construction of $\tilde{\rho}_{t_1}$, we know $\partial\ell_{t_1}(x')/\partial x' > 0$ if $\rho < \tilde{\rho}_{t_1}$; $\partial\ell_{t_1}(x')/\partial x' = 0$ if $\rho = \tilde{\rho}_{t_1}$; $\partial\ell_{t_1}(x')/\partial x' < 0$ if $\rho > \tilde{\rho}_{t_1}$.

Consider two possible cases: (1) $t_0 = 1$; (2) $t_0 > 1$.

For (1), $t_0 = 1$, so $t_1 = 1$. Hence, we only need to consider the two-period value loss. If $\rho < \tilde{\rho}_1$, the total value loss attains its minimum when x' attains its minimum, suggesting that $h(x) = \max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x\}$. If $\rho = \tilde{\rho}_1$, then the total value loss is constant with respect to x', so $h(x) = [\max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x\}, a_I]$. If $\rho > \tilde{\rho}_1$, the total value loss attains its minimum when x' attains its maximum, which implies that $h(x) = \{a_I\}$.

For (2), $(1/\theta, \infty)$ is partitioned by $\{\tilde{\rho}_t\}_{t=1}^{t_0}$: $(1/\theta, \tilde{\rho}_{t_0})$, $\{\tilde{\rho}_{t_0}\}$, $(\tilde{\rho}_{t_0}, \tilde{\rho}_{t_0-1})$,..., $(\tilde{\rho}_2, \tilde{\rho}_1)$, $\{\tilde{\rho}_1\}$, and $(\tilde{\rho}_1, \infty)$.

Consider $\rho > \tilde{\rho}_1$. The two-period value loss is minimized and total value loss is equal to the two-period value loss when $x' = a_I$, so $h(x) = \{a_I\}$.

Consider $\rho = \tilde{\rho}_{t_1}$ for t_1 taking value from $\{1, 2, ..., t_0\}$. The (t_1+1)-period value loss and also the total value loss is constant with respect to x' for a fixed t_1. Since $\rho = \tilde{\rho}_{t_1}$, $\rho > \tilde{\rho}_{t_1'}$ for any $t_1' > t_1$ and $\rho < \tilde{\rho}_{t_1'}$ for any $t_1' < t_1$. Since $\rho > \tilde{\rho}_{t_1'}$ for any $t_1' > t_1$, the total value loss decreases with x' for $t_1' > t_1$, or equivalently, for $x'\theta^{t_1} \leq a_I$.[28] Since $\rho < \tilde{\rho}_{t_1'}$ for any $t_1' < t_1$, the total value loss increases with x' for

[28] Here we implicitly rely on the continuity of the value function.

$t_1' < t_1$, or equivalently, for $x'\theta^{t_1-1} > a_I$. If $\max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x\} < a_I/\theta^{t_1}$, then $h(x) = [a_I/\theta^{t_1}, a_I/\theta^{t_1-1}]$. If $\max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x\} \in [a_I/\theta^{t_1}, a_I/\theta^{t_1-1}]$, then $h(x) = [\max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x\}, a_I/\theta^{t_1-1}]$. If $\max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x\} > a_I/\theta^{t_1-1}$, then $h(x) = \max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x\}$. In sum, for $\rho = \tilde{\rho}_{t_1}$, $h(x) = [\max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x, a_I/\theta^{t_1}\}, \max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x, a_I/\theta^{t_1-1}\}]$.

Consider $\rho \in (\tilde{\rho}_{t_1+1}, \tilde{\rho}_{t_1})$ for t_1 taking value from $\{1,...,t_0-1\}$. The (t_1+1)-period value loss and also the total value loss is minimized when x' attains its minimum for a fixed t_1. Since $\rho > \tilde{\rho}_{t_1+1}$, $\rho > \tilde{\rho}_{t_1'}$ for any $t_1' > t_1$, which implies that the total value loss decreases with x' for $t_1' > t_1$, or equivalently, for $x'\theta^{t_1} \le a_I$. Since $\rho < \tilde{\rho}_{t_1}$, $\rho < \tilde{\rho}_{t_1'}$ for any $t_1' < t_1$, which implies that the total value loss increases with x' for $t_1' < t_1$, or equivalently, for $x'\theta^{t_1-1} > a_I$. Hence, we have $h(x) = \max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x, a_I/\theta^{t_1}\}$.

Last, consider $\rho < \tilde{\rho}_{t_0}$. Since we know $\tilde{\rho}_{t_0} \le \tilde{\rho}_t$ for any $t = 1,2,...,t_0$, $\rho < \tilde{\rho}_{t_1}$ for any t_1. This suggests that the (t_1+1)-period value loss and also the total value loss increases with x' for *any* given t_1. Then the total value loss is minimized when x' attains its minimum. Hence, $h(x) = \max\{\zeta(\hat{x}-x)+\hat{x},(1-d)x\}$.

We have now obtained the desired conclusion.

References

[1] Asimakopulos, A., 1977, Review of "The Theory of Equilibrium Growth" by A. K. Dixit. *Journal of Political Economy* **85**, 1311-1312.

[2] Barna, T., 1957, Review of "Essays in the Theory of Economic Growth" by Joan Robinson. *Economic Journal* **67**, 490-493.

[3] Benhabib, J. (eds.), 1992, *Cycles and Chaos in Economic Equilibrium*, Princeton University Press.

[4] Benhabib, J., and Nishimura, K., 1985, Competitive equilibrium cycles. *Journal of Economic Theory* **35**, 284-306.

[5] Benhabib, J., Nishimura, K. and Venditti, A., 2002, Indeterminacy and cycles in two-sector discrete-time model. *Economic Theory* **20**, 217-235

[6] Burmeister, E. and Dobell, A. R., 1970. *Mathematical Theories of Economic Growth.* New York: MacMillan.

[7] Chen, B., Nishmiura, K. and Shimomura, K., 2004, An Oniki-Uzawa dynamic two-country model of international trade: back to Heckscher and Ohlin. Kyoto University, mimeo,

[8] Dana, R., la Van, C., Mitra, T. and Nishimura, K. (eds.), 2006, *Handbook on Optimal Growth I* (Berlin: Springer-Verlag).

[9] Dixit, A. K., 1976, *The Theory of Equilibrium Growth.* Oxford: Oxford University Press.

[10] Fujio, M., 2005. The Leontief two-sector model and undiscounted optimal growth with irreversible investment: the case of labor-intensive consumption goods. *Journal of Economics* **86**, 145-159.

[11] Fujio, M., 2006, *Optimal Transition Dynamics in the Leontief Two-sector Growth Model*, Ph.D. Dissertation, The Johns Hopkins University.

[12] Fujio, M., 2006, Undiscounted optimal growth in a Leontief two-sector model with circulating capital: the case of capital-intensive consumption goods sector. *Journal of Economic Behavior and Organization*, **66**, 420-436.

[13] Fujio, M., 2009, Optimal transition dynamics in the Leontief two-sector growth model with durable capital: the case of capital-intensive consumption goods. *Japanese Economic Review*, **60**, 490-511.

[14] Fujio, M. and Khan, M. Ali, 2006, Ronald W. Jones and Two-Sector Growth: Ramsey Optimality in the RSS and Leontief Cases. *Asia-Pacific Journal of Accounting & Economics* **13**, 87-110

[15] Hamberg, D., 1963, Review of "Essays in the Theory of Economic Growth" by Joan Robinson. *American Economic Review* **53**, 1109-1114.

[16] Haque, W., 1970, Sceptical notes on Uzawa's "Optimal growth in a two-sector model of capital accumulation", and a precise characterization of the optimal path. *Review of Economic Studies* **37**, 377-394.

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

[18] Khan, M. Ali and Mitra, T., 2006, Discounted optimal growth in the two-sector RSS model: a geometric investigation. *Advances in Mathematical Economics* **8**, 349-381.

[19] Khan, M. Ali, Mitra, T., 2007a. Optimal growth in a two-sector model without discounting: a geometric investigation. *Japanese Economic Review* **58**, 191-225.

[20] Khan, M. Ali and Mitra, T., 2007b, Optimal growth under discounting in the two-sector Robinson-Solow-Srinivasan model: a dynamic programming approach, *Journal of Difference Equations and Applications* **13**, 151-168.

[21] Khan, M. Ali and Mitra, T., 2012, Impatience and dynamic optimal behavior: a bifurcation analysis of the Robinson-Solow-Srinivasan model, *Nonlinear Analysis* **75** 1400-1418.

[22] Khan, M. Ali and Mitra, T., 2013, Optimal growth in a two-sector RSS model with discounting: A further geometric investigation, *Advances in Mathematical Economics* **17**, 1-33.

[23] Khan, M. Ali and Piazza, A., 2011, Optimal cyclicity and chaos in the 2-sector RSS model: An anything-goes construction. *Journal of Economic Behavior & Organization* **80**, 397-417.

[24] Lancaster, L., 1960, Mrs. Robinson's dynamics, *Economica* **27**, 63-70

[25] Le Van, C. and Dana, R.A., 2003, *Dynamic Programming in Economics* Dordrecht: Kluwer Academic Publishers.

[26] Little, I. M. D., 1957, Classical growth. *Oxford Economic Papers* **9**, 152-177.

[27] Magee, S. P., 1973, Factor market distortions, production, and trade: a survey. *Oxford Economic Papers* **25**, 1-43.

[28] Magee, S. P., Brock, W. A. and Young, L., 1989, *Black Hole Tariffs and Endogenous Policy Theory* Cambridge: Cambridge University Press.

[29] Majumdar, M., Mitra, T., and Nishimura, K. (eds.) 2000, *Optimization and Chaos.* Berlin: Springer-Verlag.

[30] May, R., 1976, Simple mathematical models with very complicated dynamics, *Nature* **261**, 459-67.

[31] McKenzie, L., 1986, Optimal economic growth, turnpike theorems, and comparative dynamics, *Handbook of Mathematical Economics* (Arrow, K. and Intrilligator, M., eds.), **3** 1281-1355.

[32] Morishima, M., 1965, The multi-sectoral theory of economic growth. In Bruno de Finetti (ed.) *Mathematical Optimization in Economics,* pp. 81-163, Fondazione CIME, Berlin: Springer-Verlag.

[33] Morishima, M., 1969, *Theory of Economic Growth*, Oxford University Press.

[34] Nishimura, K. and Yano, M., 1995, Nonlinear dynamics and chaos in optimal growth: an example. *Econometrica*, **63**, 981-1001.

[35] Nishimura, K. and Yano, M., 1996, Chaotic solutions in dynamic linear programming. *Chaos, Solitons, and Fractals*, **7**, 1941-1953.

[36] Oswald, A. S., The economic theory of trade unions: a survey. *Scandinavian Journal of Economics*, **87**, 160-193.

[37] Robinson, J., 1956, *The Accumulation of Capital.* London: Macmillan. Reprinted by Palgrave Macmillan in 2013.

[38] Robinson, J., 1962, *Essays in the Theory of Economic Growth.* London: Macmillan.

[39] Robinson, J., 1970, Review of "The Neoclassical Theory of Production and Distribution" by C. E. Ferguson. *Economic Journal* **80**, 336-339.

[40] Robinson, J., 1971, *Economic Heresies: Some Old-Fashioned Questions in Economic Theory.* New York: Basic Books.

[41] Robinson, J., 1971, *Contributions to Modern Economics.* New York: Academic Press.

[42] Shinkai, Y., 1960, On equilibrium growth of capital and labor. *International Economic Review* **1**, 107-111.

[43] Solow, R., 1999, Neoclassical growth theory, pp. 637-667 in J. B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, Volume 1A. Amsterdam: Elsevier Science B. V.

[44] Solow, R., 2000, *Growth Theory, An Exposition* Oxford: Oxford University Press.

[45] Stiglitz, J. E. and Uzawa, H., 1969, *Readings in the Modern Theory of Economic Growth* Cambridge: MIT Press.

[46] Worswick G. D. N., 1959, Mrs. Robinson on simple accumulation: a comment with algebra. *Oxford Economic Papers* **11**, 125-142.

[47] Worswick, G. D. N., 1963, Review of "Essays in the Theory of Economic Growth" by Joan Robinson. *Economic Journal* **73**, 295-297.

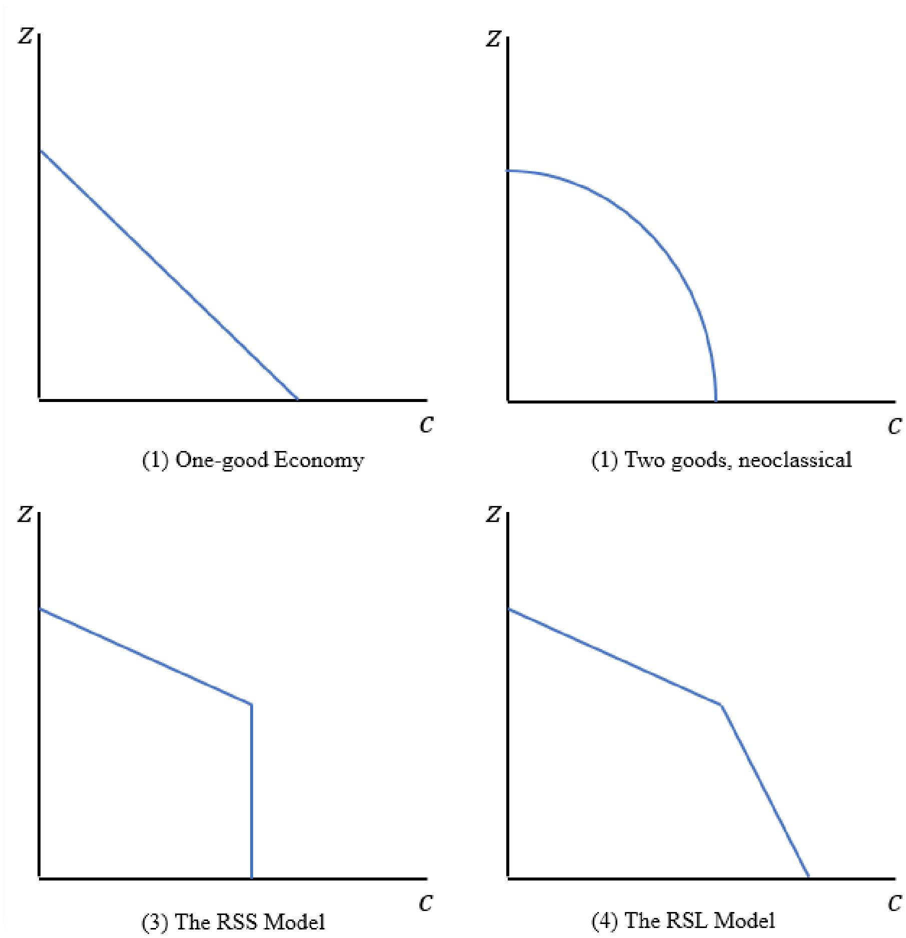

Figure 1: The Production Possibility Frontier

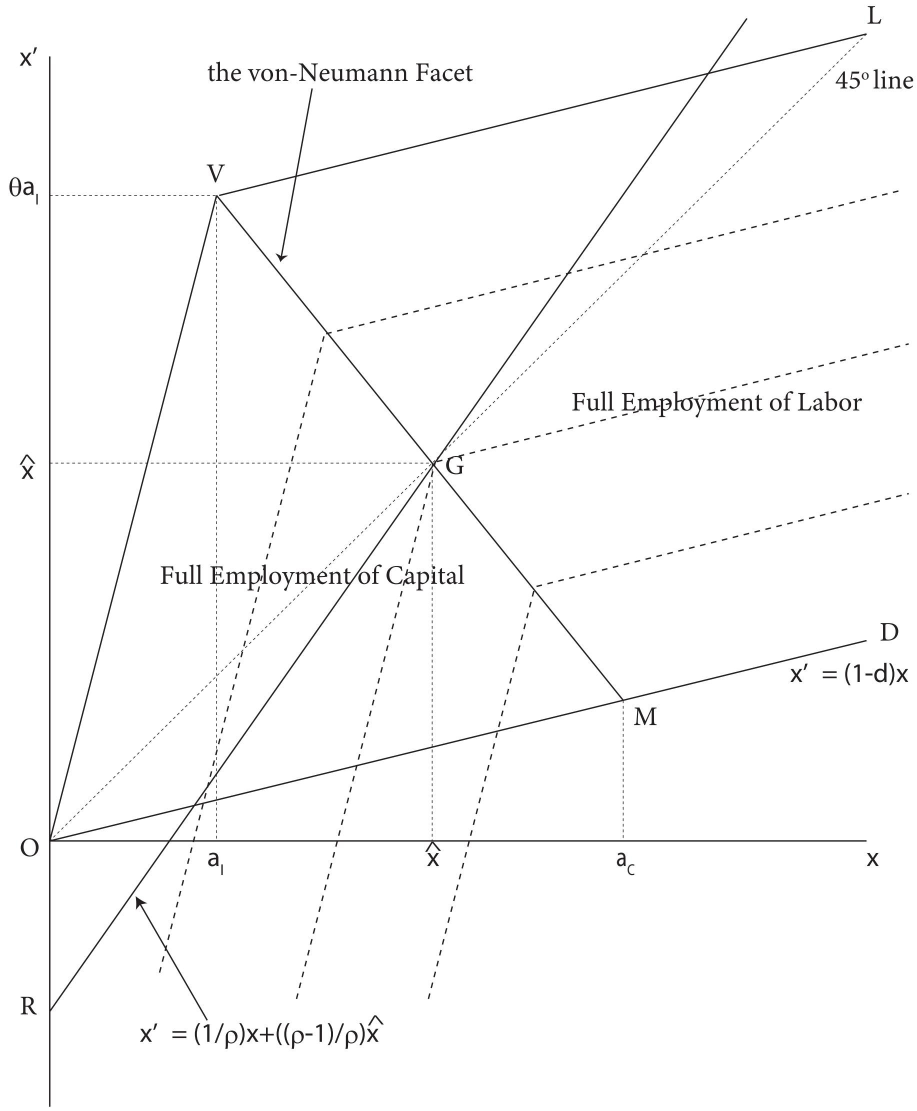

Figure 2: The Basic Geometry

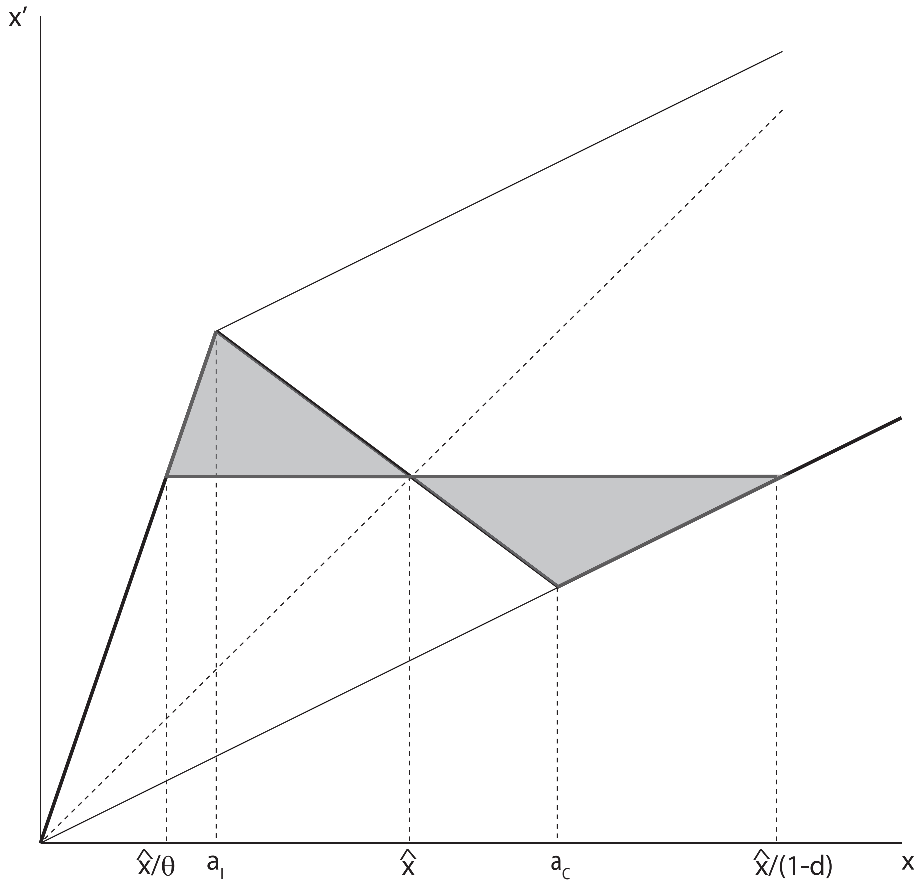

Figure 3: Illustration of Proposition 1

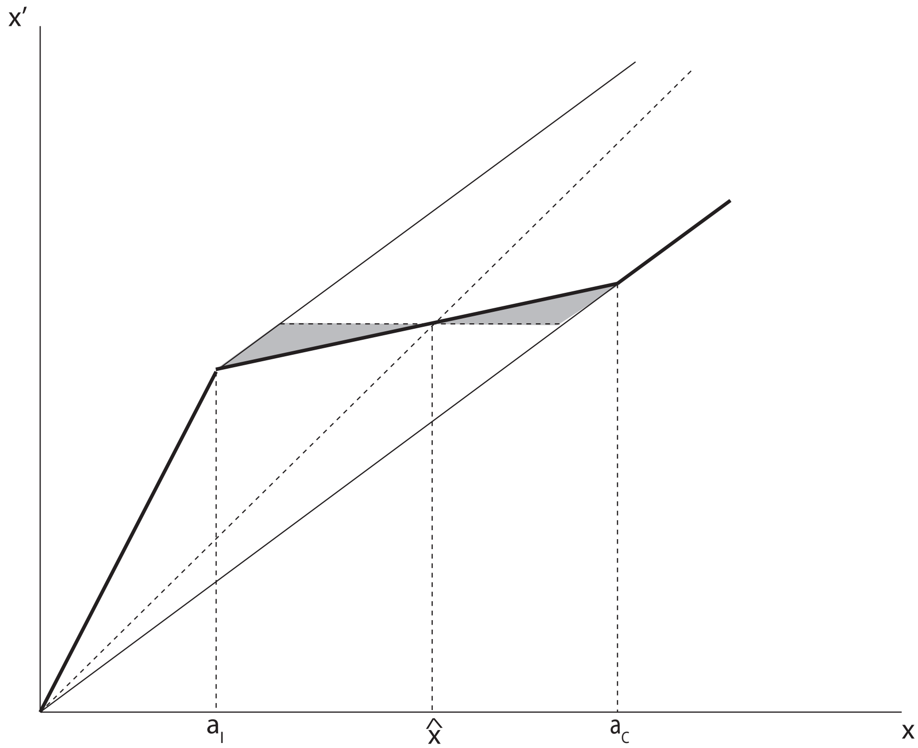

Figure 4: Illustration of Proposition 3

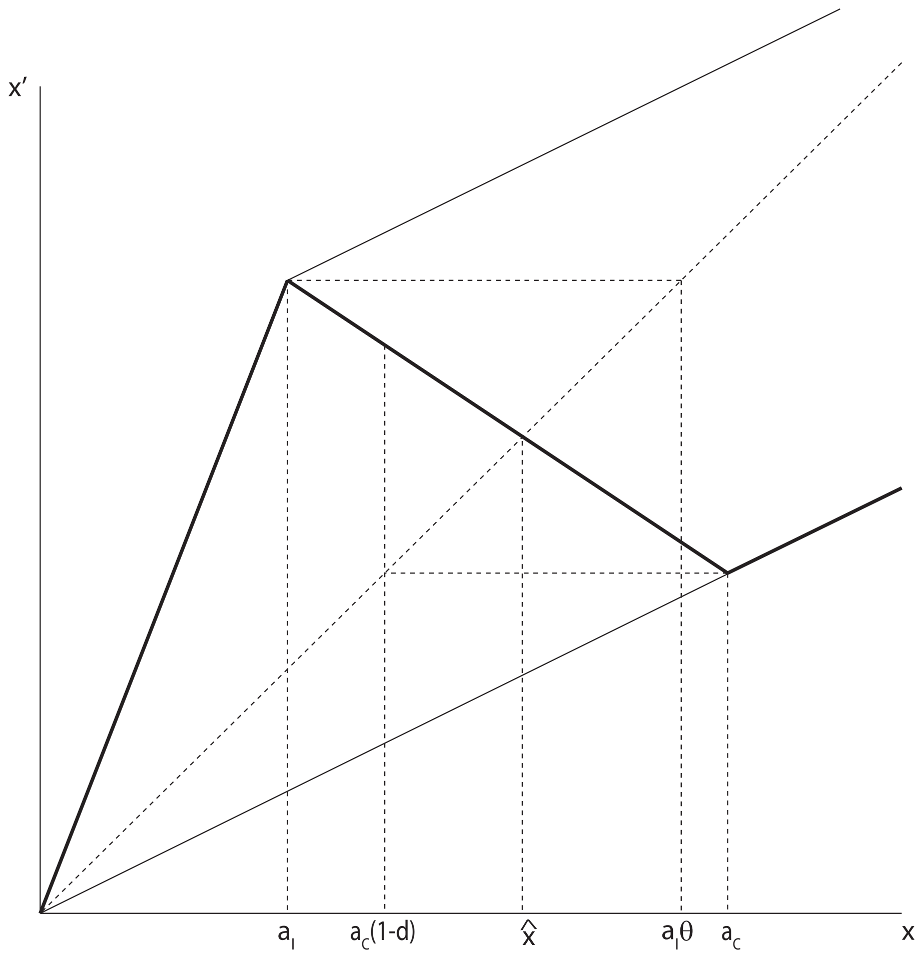

Figure 5: Illustration of Proposition 5: Case (i)

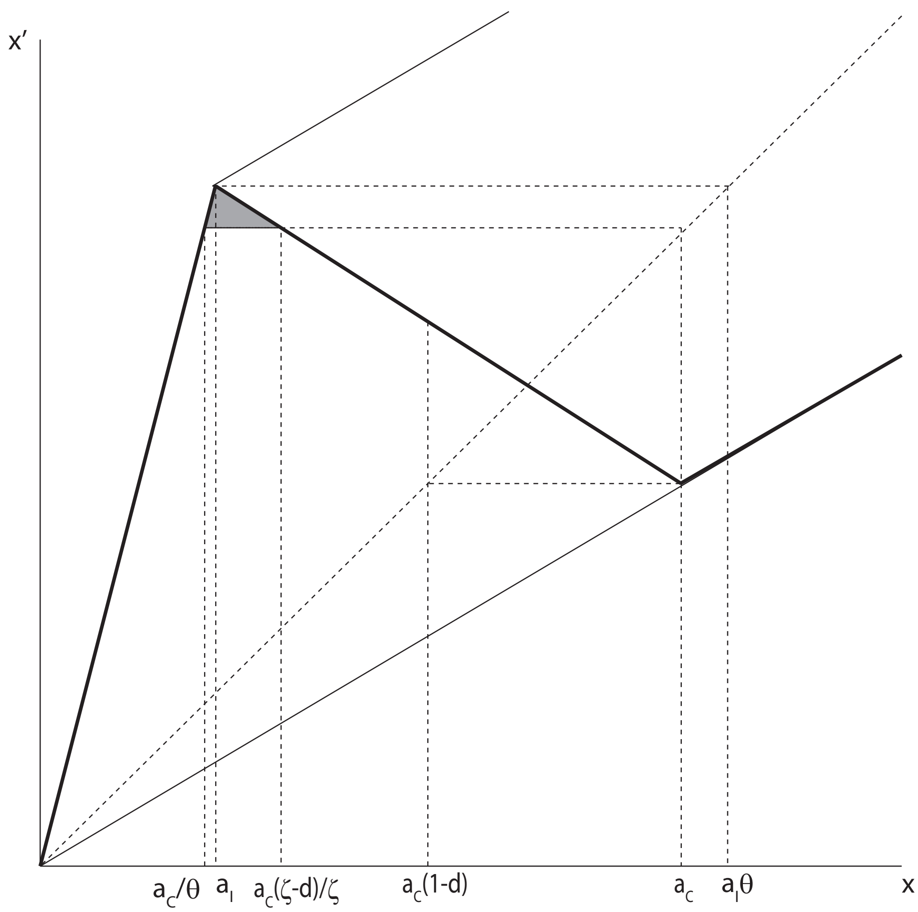

Figure 6: Illustration of Proposition 5: Case (ii)

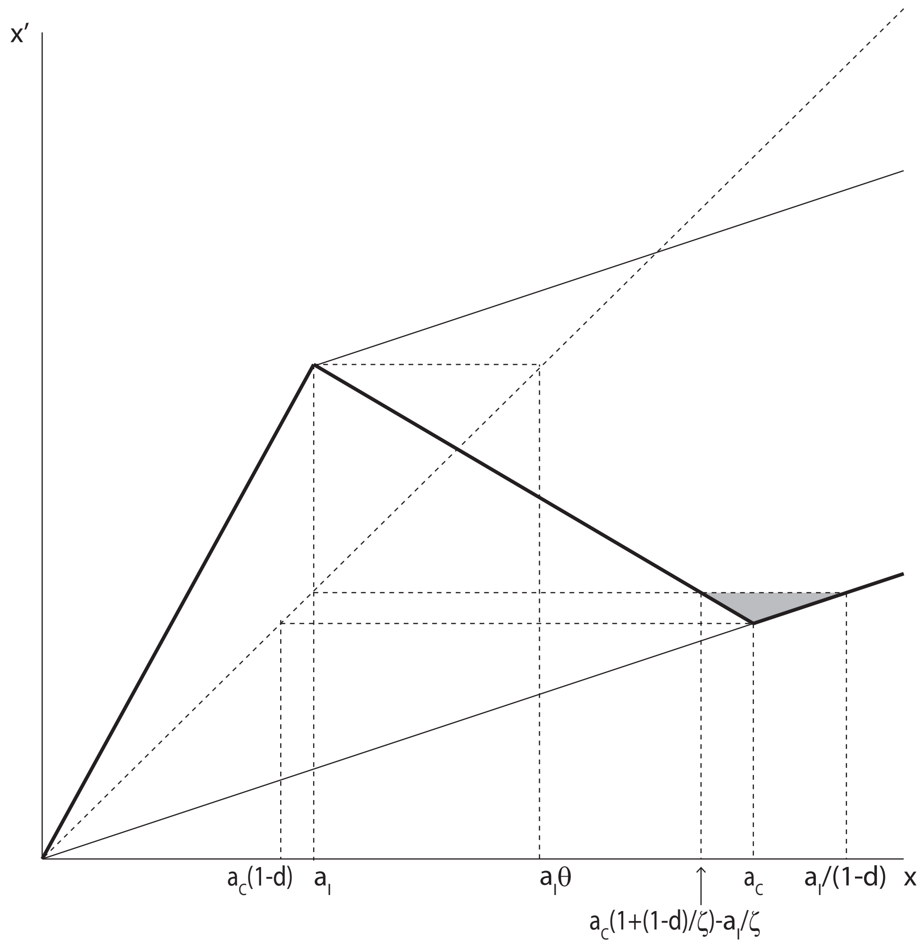

Figure 7: Illustration of Proposition 5: Case (iii)

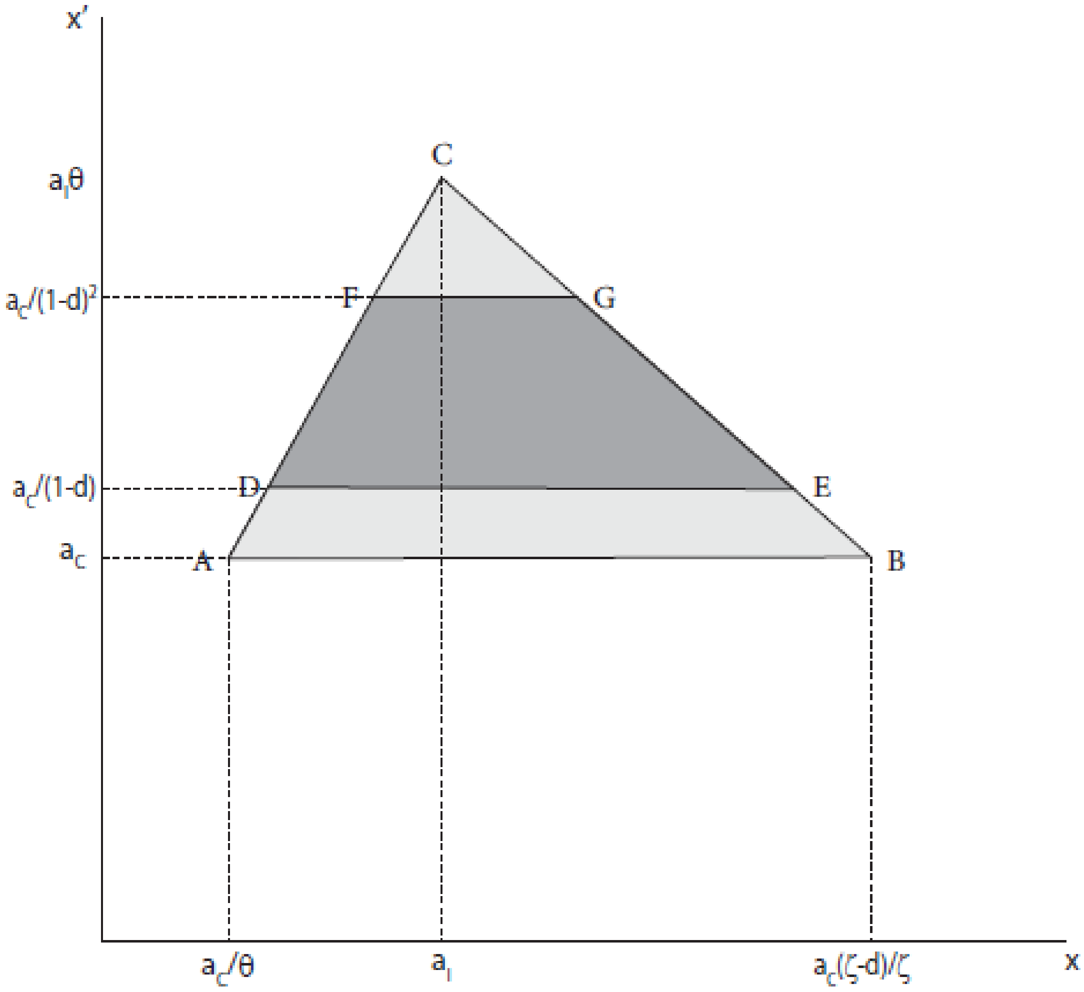

Figure 8: Illustration of Proposition 2

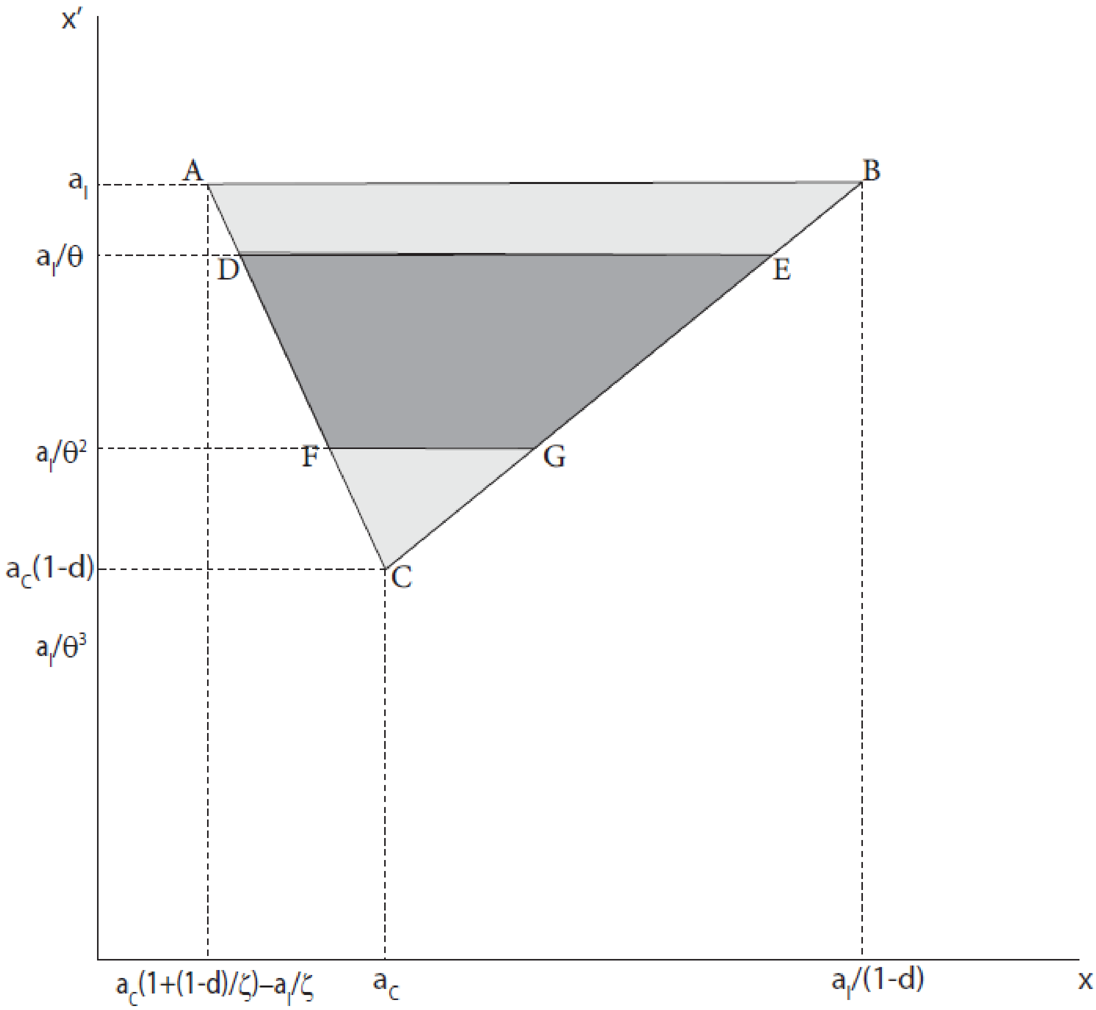

Figure 9: Illustration of Proposition 3

Table 1: Parameter Comparison with the Antecedent Literature

Model	Consumption Goods Labor	Consumption Goods Capital	Investment Goods Labor	Investment Goods Capital	Depreciation Rate	Marginal Transformation of Capital
RSL	1	a_C	$1/b$	a_I/b	$d \in (0,1)$	$\zeta = b/(a_C - a_I) - (1-d)$
RSS	1	1	a	0	$d \in (0,1)$	$\xi = 1/a - (1-d)$
Nishimura-Yano	α	1	β/μ	$1/\mu$	$d = 1$	$\alpha\mu/(\beta - \alpha)$
Benhabib-Nishimura	a_{00}	a_{10}	a_{01}	a_{11}	$g \in (0,1)$	$-[a_{01}(a_{11}/a_{01} - a_{10}/a_{00})]^{-1} - (1-g)$
Lancaster	a_{22}	a_{21}	a_{12}	a_{11}		
Worswick I	$1/c$	$1/(cn)$	m	0	$d \in (0,1)$	
Worswick II	$1/c$	$1/(cn)$	$1/h$	$1/(\mathrm{hl})$	$d \in (0,1)$	

Notes: (1) Benhabib-Nishimura works with a general production function and all the coefficients are those at the steady state; (2) Both Lancaster and Worswick analyze descriptive growth models with a given saving function.