The Optimal Management of Orchards*

Daniel Tregeagle† Leo Simon‡

July 23, 2025

Abstract

Perennial crops, particularly fruit and tree nut orchards, represent a high-value sector in global agriculture, yet the theory of optimal management with production smoothing preferences in multi-age-class orchards remains unresolved. While previous analyses examine either single-tree models or multi-age orchards without production smoothing, we analyze orchards where growers prefer stable production over time and must manage trees of multiple ages simultaneously. This paper characterizes optimal tree replacement strategies and their long-run dynamics: when to replace trees and whether to replace them simultaneously or partially, and how orchards evolve toward steady-state management patterns. We first extend existing two-age-class models by characterizing complete transition dynamics, correcting an omission in previous work, and developing comparative statics for optimal cycle amplitude. We then introduce a three-age-class model that provides the first complete convergence characterization for orchards with non-monotonic yield curves and production smoothing preferences. For both models, we prove that optimally managed orchards exhibit cyclical production patterns with partial rather than simultaneous replacement, and demonstrate finite-time convergence: arbitrary initial age distributions converge to optimal cycles within two periods (two-age-class) or a bounded number of periods (three-age-class). We identify a 'cycle region' where initial allocations immediately generate optimal cycles. Our results are applicable to other point-input, flow-output capital assets with non-monotonic productivity.

Keywords: Age-Structured Dynamics, Dynamic Programming, Optimal Replacement, Cyclical Equilibrium, Point-Input Flow-Output Capital, Uniform Convergence, Perennial Crops, Production Smoothing

JEL Classification: C61, C62, D25, Q12, Q23

---

*Tregeagle dedicates this paper to the memory of Leo Simon; he is greatly missed. Material from chapter 3 of Tregeagle's dissertation, *The Dynamics of Perennial Crop Production and Processing* was used to develop this paper. We thank Mark Agerton, Julian Alston, Rachael Goodhue, Larry Karp, Jeffrey LaFrance, Zoë Plakias, Gordon Rausser, Dan Sumner, Clevo Wilson, Don Webber, David Zilberman, and the participants of the UC Berkeley Ag Policy Seminar, UC Davis Ag Policy Econ Workshop, Western Economic Association International Conference 2018, the Western Economic Association International Early Career Workshop 2018, the University of Sydney Agricultural Economics Reading Group, and the Monash Early Career Economists Conference 2018 for useful comments and discussions

†Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, NC, 27695. Corresponding Author: tregeagle@ncsu.edu. Nelson Hall, Campus Box 8109, 28o1 Founders Drive, Raleigh, NC, 27695. Ph: +1 (919) 515-6091.

‡Deceased. Department of Agricultural and Resource Economics, University of California Berkeley, Berkeley, CA, 94720.

1 Introduction

Perennial crops, particularly fruit and tree nut orchards, represent a critically important sector in global agriculture due to their substantial economic and environmental contributions, with a farm-gate value accounting for at least 0.5 trillion USD in 2022 (FAO, 2024). Despite their importance, the optimal economic management of orchards has remained understudied, especially regarding the combined complexities of multiple age-class structures and non-monotonic yield patterns. This paper provides the first comprehensive analysis of optimal orchard management that fully characterizes the convergence dynamics under both these characteristics, identifying cyclical long-run optimal production and demonstrating that optimally managed orchards converge to this cycle in finite time—within two periods for the two-age-class model and within a bounded number of periods for the three-age-class model. We also derive comparative statics describing how the amplitude of the production cycle responds to economic parameters.

The central research questions addressed in this paper are explicitly as follows: (1) At what age should a grower replace orchard trees to maximize the enterprise value? (2) Should the grower replace some or all of the trees simultaneously? (3) What is the steady-state of an optimally managed orchard, and how rapidly does the orchard converge to this state from an arbitrary initial condition?

These questions are complicated by three key features of perennial crop production that distinguish orchards from conventional capital replacement problems. First, perennial crops exhibit non-monotonic 'hump-shaped' yield profiles over their lifespans, with productivity initially rising, then declining in later years. Second, commercial orchards typically contain trees of multiple ages rather than homogeneous stands, creating complex interdependencies between replacement decisions across different age cohorts. Third, the combination of non-monotonic yields and heterogeneous age structures means that replacement timing affects not only individual tree productivity but also the temporal pattern of total orchard output. As a result, optimal replacement strategies must account for how timing decisions influence the smoothness of production streams over time.

Production-smoothing preferences, reflecting a grower's inclination to avoid substantial year-to-year variations in output, play an important role in orchard management. These preferences arise due to various economic and operational factors, such as credit constraints, processing capacity limitations, and supply contracts, which incentivize growers to maintain stable production levels. In this study, such preferences are modeled using a strictly concave utility function to reflect diminishing marginal returns and the grower's aversion to output fluctuations (Mitra et al., 1991; Feinerman and Tsur, 2014).

A distinguishing biological feature of perennial crops, which significantly complicates orchard management, is their non-monotonic ("hump-shaped") yield trajectory. Typically, orchard trees experience an initial period of yield growth, followed by a plateau at peak production, and eventually yield decline as trees age (Siegle et al., 2024; Mitra et al., 1991). This characteristic necessitates careful timing in replacement decisions, as optimal replacement must balance the decreasing productivity of older trees against the costs of establishing new ones. The presence of multiple age-classes within a single orchard further complicates this decision-making process. For instance, the almond orchards of California exhibit a wide distribution of tree ages, with trees uniformly distributed from one to 20 years of age, after which density declines (Luckstead and Devadoss, 2024). Such multiple-age structures reflect real-world complexities that cannot be captured by a simplified, identically-aged orchard model where all trees are assumed to have the same age and are replaced as a single unit.

This analysis relates closely to the literature on capital-theoretic models of forestry rotation (Mitra and Wan, 1985, 1986; Salo and Tahvonen, 2003, 2004) but differs fundamentally in the timing of payoffs. Forestry is typically modeled as a point-input, point-output system, where trees are planted at one time and harvested at a later discrete point. In contrast, perennial orchards represent a point-input, flow-output system, where trees are planted once but provide a continuous stream of outputs over their productive life, necessitating orchard management decisions that emphasize ongoing yield optimization rather than a sequence of harvest events (Mitra et al., 1991).

Methodologically, this paper significantly extends the existing theoretical framework for orchards. We first develop a two-age-class model that mirrors the models of Mitra et al. (1991) and Wan

(1993). While the qualitative structure of this model remains similar to the earlier literature, we contribute three major refinements: we derive comparative statics for the cycle amplitude, provide a complete characterization of the transition dynamics to the steady-state, and clarify an omission in the solution proposed by Wan (1993). Second, we introduce a new three-age-class model, which accommodates non-monotonic yields. We analytically characterize the steady-state of this model, showing that while even-aged age structures represent an optimal solution, they are not uniquely optimal, with a continuum of initial allocations generating optimal three-period cycles, and prove that all initial conditions converge uniformly to a steady-state cycle within a finite, bounded number of periods. Our results show that there exists a 'cycle region' of initial allocations that immediately generate optimal three-period cycles, while all other allocations transition to this region. This finite-time convergence result represents a significant strengthening of previous neighborhood convergence results in the literature.

2 Replacing a single tree

Before analyzing heterogeneous orchard management, we first examine optimal replacement decisions for individual trees, establishing the foundation for our broader analysis. This approach follows the methodology used by Mitra et al. (1991) and Wan (1993), where single-tree models illuminate key economic trade-offs in perennial crop management. Understanding these basic principles informs our subsequent analysis of two-age-class orchards (Section 3) and three-age-class systems with non-monotonic yields (Section 4), enabling us to characterize optimal management strategies with heterogeneous age structures.

2.1 The yield curve

The yield curve describes how the yield of the plant changes over its life cycle. We use a deterministic yield curve, abstracting from the fact that observed yield curves in empirical applications will be stochastic functions of multiple variables, including rainfall, temperature profile throughout the growing season, soil type, inputs applied (e.g. fertilizer, pesticide), hours of labor, etc.

Adapting the single tree replacement framework of Mitra et al. (1991), we specify a generic yield curve with four integers demarking the non-bearing period, the period of increasing yield, the period of constant yield, and the period of decreasing yields. That is, integers P, Q, R, and T mark the end of each period, with $1 \leq P < Q \leq R \leq T$, such that

$$f_P \leq f_Q = \ldots = f_R \geq \ldots \geq f_T$$

and at least one strict inequality between f_1 and f_Q, where f_i is the yield of an i-year old tree. The integer P corresponds to the end of the non-bearing period, Q corresponds with the beginning of the yield plateau, R with the end of the yield plateau, and T with the end of the tree's lifespan. A stylized yield curve is shown in figure 1.

In their solution to the single-tree replacement problem, Mitra et al. (1991) require only mono-

tonicity for the increasing and decreasing phases of the age-yield relationship.[1] Their model allows an initial non-bearing period.

In this paper we focus on two special cases of their general relationship: a two-age class model (young trees and old trees), where old trees are less productive than young trees; and a three-age-class model (young, mature, and old trees) where the old age-class may be either more or less productive than mature trees. For ease of exposition we will refer to these age classes as y, o in the two-age-class model (rather than $1, 2$), and y, m, o in the three-age-class model (rather than $1, 2, 3$). Moreover, to avoid carrying around an explict planting cost term, assume without loss of generality that the yield of young trees, f_y, is net of replanting costs.

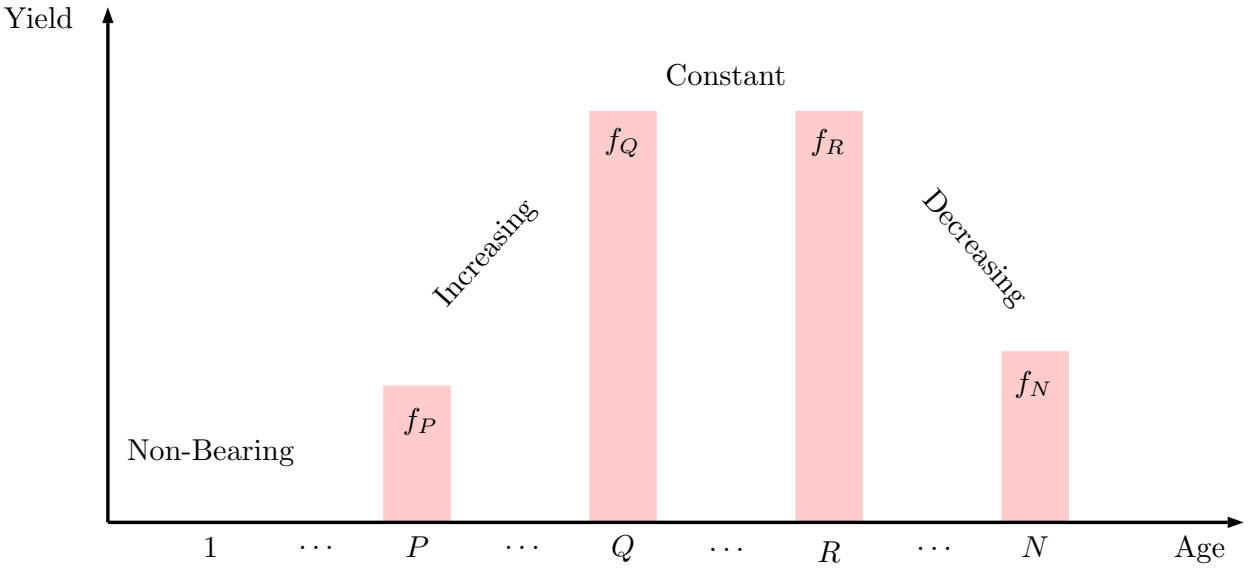


Figure 1: A generic non-monotonic (hump-shaped) yield curve. It is not necessary for a perennial yield curve to have a non-bearing year.

2.2 The optimal replacement age for a two-age-class tree

Before studying the orchard problem, we begin by considering the optimal replacement age for a single tree. Consider a single plot of land that is continuously replanted. fter harvest, a new tree is immediately planted. The tree lives at most two periods, and the grower chooses either a 1-period

[1] Mitra et al. (1991) find the optimal replacement for a single tree with the generic yield curve above. However their non-convergence result (Proposition 5.2) assumes $R = T$ (monotonically increasing yields), while their neighborhood turnpike theorem (Proposition 5.3) applies more generally

or 2-period rotation over an infinite horizon. Normalize the price of output to 1. Let β (< 1) be the discount factor. The grower's objective is to maximize the discounted benefits from the stream of harvests over an infinite time horizon. The net present value of the tree grown on a 1-period rotation is

$$\text{NPV}_1 = f_y + \beta f_y + \beta^2 f_y + \beta^3 f_y + \ldots = \frac{f_y}{1-\beta}$$

The net present value of the tree grown on a 2-period rotation is

$$\text{NPV}_2 = f_y + \beta f_o + \beta^2 f_y + \beta^3 f_o + \ldots = \frac{f_y + \beta f_o}{1-\beta^2}$$

The difference between the two values is

$$\text{NPV}_1 - \text{NPV}_2 = \frac{\beta(f_y - f_o)}{1-\beta^2}$$

Hence if $f_y > f_o$ the optimal replacement age is y (let i^* be the optimal replacement age, so ($i^* = y$). If $f_y < f_o$ the optimal replacement age is o ($i^* = o$). Finally, if $f_y = f_o$ ($i_1^* = y$; $i_2^* = o$). [2]

The optimal rotation age does not depend on discounting in this model. Discounting will affect the size of the net present value, but not the relative size of NPV_1 and NPV_2. This is a special result, due to the assumption of a two-age-class model. In a more general model, the optimal single tree replacement age is a function of the discount factor, as shown by Mitra et al. (1991) in proposition 3.1.

2.3 The optimal replacement age for a three-age-class tree

While the two-age-class model provides important insights, it cannot capture the crucial biological reality of non-monotonic yields where productivity first rises and then falls. However, a three-age-class tree is the simplest case that addresses this limitation, allowing us to model the important

[2] Theorem 3.1 in Mitra et al. (1991), adapted to our notation, states that there will either be one or two optimal replacement ages, i^* or i_1^* and i_2^*. The age(s) will occur during the declining section of the yield curve. If there are two optimal replacement ages, then they will differ by one period. That is $Q \leq i_1^* \leq i_2^* \leq i_1^* + 1 \leq T$.

hump-shaped yield patterns common in perennial crops.

Figure 2 shows the optimal replacement for a single three-age-class tree, normalizing the yield of the second age-class to one ($f_m = 1$). The horizontal axis is the yield of the first age-class relative to the second, and the vertical axis is the yield of the third age-class relative to the second. The optimal replacement age also depends on the discount factor. The dotted line represents parameter combinations where a grower with near-zero discount rate would be indifferent between replacement strategies, while the dashed line shows the same boundary for a patient grower (β approaching 1).

Regardless of the discount factor, there are parameter sets giving an optimal replacement age of three where the yield of the third age-class is less than the second age-class, that is, with non-monotonic yield. Devadoss and Luckstead (2010) identified a declining final period as a key feature of perennials. While Mitra et al. (1991) could theoretically analyze the entire $i^* = 3$ region in figure 2, their specific convergence results were limited. They proved non-convergence for the monotonically increasing case $f_o \geq 1$ and only neighborhood convergence generally. Our analysis provides definitive finite-time convergence results for the entire $i^* = 3$ region, including the biologically important hump-shaped yields where $f_3 < 1$.

These single-tree models show how yield patterns and economic parameters determine optimal replacement timing. The two-age-class model demonstrates that when old trees outproduce young ones, keeping trees for multiple periods is optimal. The three-age-class model extends this insight to account for the biologically realistic hump-shaped yield curves, showing that trees may be optimally replaced even after their productivity begins declining. These single-tree optimal replacement ages form the foundation for our orchard-level analysis in subsequent sections.

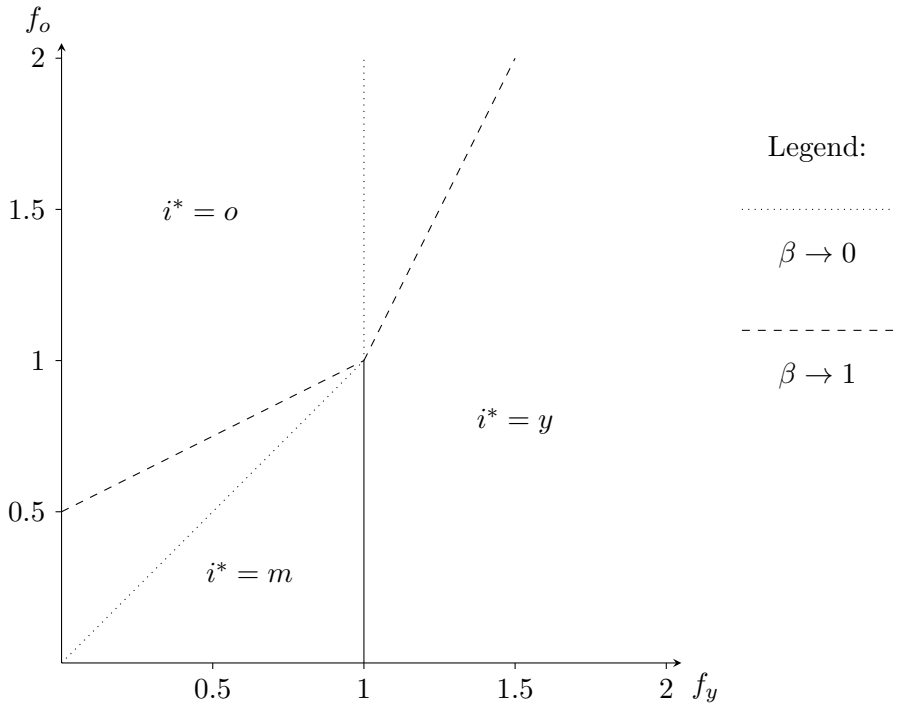


Figure 2: The optimal replacement age for a three-age-class tree. The yield of a mature tree is normalized to 1 ($f_m = 1$). The boundaries between the two lower optimal replacement ages $i^* = y, m$ and the oldest replacement age $i^* = o$ depend on the discount factor β. The boundary between $i^* = 1$ and $i^* = 2$ is fixed with respect to β.

3 A two-age-class model

In this section, we present a two-age-class, infinite horizon model of a perennial orchard. An age class refers to all trees planted in the same season. Every period, each cohort advances to the next age class. The model allows us to analyze optimal allocation sequences under different yield relationships between young and old trees, with particular focus on when old trees have higher yield than young trees. We show that optimal sequences are typically cyclical when old trees outperform young trees, and examine how cycle characteristics respond to changes in economic and biological parameters.

3.1 A two-age-class, infinite horizon orchard model

We adapt the forestry model of Salo and Tahvonen (2002) to perennials by changing the production function from a point payoff to a flow payoff, and adopt the aging constraint structure from Salo and Tahvonen (2004). Instead of including explicit choice variables for replanting, we formulate a reduced form dynamic optimization problem (Mitra, 2000). For clarity, throughout this section we refer to young trees ($i = 1$) with subscript y, and old trees ($i = 2$) with subscript o. The grower's objective is to maximize the discounted benefits from the stream of harvests from each type of tree over an infinite time horizon.

$$V(x_{y0}, x_{o0}) = \max_{\{x_{yt}, x_{ot}\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^t u(c_t) \tag{1a}$$

subject to

$$c_t \equiv f_y x_{yt} + f_o x_{ot} \tag{1b}$$

$$x_{o,t+1} \leq x_{yt} \tag{1c}$$

$$x_{yt} + x_{ot} \leq L \quad (= x_{y0} + x_{o0}) \tag{1d}$$

$$x_{it} \geq 0 \tag{1e}$$

where $\{x_{yt}\}_{t=1}^{\infty}$ is the sequence of land allocations to young trees, $\{x_{ot}\}_{t=1}^{\infty}$ is the sequence of land allocations to old trees, f_y is the yield of young trees, f_o is the yield of old trees,[3] $u(c_t)$ is the benefit to the grower from growing/consuming fruit in period t, and L is the area of land available for the orchard. We assume that the benefit function exhibits diminishing marginal returns: $u'(.) > 0$ and $u''(.) < 0$. Finally, c_t is the total quantity of fruit available for consumption in period t. This model differs from Salo and Tahvonen (2002) in two important ways. First, the definition of consumption, c_t, represents harvested fruit rather than timber. Second, we added an explicit total land constraint. In a forestry model, benefits are obtained only when trees are cut (replaced). In contrast, in this orchard model, fruit is harvested before any replanting decisions are made, and there is no direct benefit from replacing a tree.

The aging constraint (1c) ensures that the number of old trees in the next period cannot exceed the number of young trees in the current period.[4] Old trees cannot be bought and planted; they must grow from young trees.

This optimization problem is convex since the objective function is strictly concave and the constraints are linear. Therefore any solution to the Karush-Kuhn-Tucker conditions will also be a solution to the constrained optimization problem posed above.

The corresponding Lagrangian function is

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \left[u(c_t) + \lambda_t (x_{yt} - x_{o,t+1}) + \psi_t (L - x_{yt} - x_{ot}) \right] \tag{2a}$$

[3] These productivities are defined as net of planting costs.

[4] Wan (1993) calls this constraint the *cross-vintage bound*. Other authors such as Salo and Tahvonen (2002, 2003, 2004) follow this terminology. We feel that *aging constraint* is more intuitive.

The associated KKT conditions for $t \geq 1$ are

$$\beta^{-t} \frac{\partial \mathcal{L}}{\partial x_{yt}} = u'(c_t) f_y + \lambda_t - \psi_t \leq 0 \tag{2b}$$

$$\beta^{-t} \frac{\partial \mathcal{L}}{\partial x_{ot}} = u'(c_t) f_o - \frac{\lambda_{t-1}}{\beta} - \psi_t \leq 0 \tag{2c}$$

$$\lambda_t \geq 0;\ \lambda_t(x_{yt} - x_{o,t+1}) = 0 \tag{2d}$$

$$\psi_t \geq 0;\ \psi_t(L - x_{yt} - x_{ot}) = 0 \tag{2e}$$

$$x_{it} \geq 0;\ x_{it} \frac{\partial \mathcal{L}}{\partial x_{it}} = 0 \tag{2f}$$

where λ_t is the Lagrangian multiplier corresponding to the aging constraint and ψ_t is the Lagrangian multiplier corresponding to the total land constraint.

3.1.1 Interpreting ψ_t

The variable ψ_t represents the marginal increase in the orchard's value in period t from a permanent marginal addition to total land. If ψ_t is positive, the total land constraint is binding, meaning all available land is being used productively. This implies that old trees are immediately replanted at the end of their lifespan—land is never left fallow.

Throughout this paper, we proceed assuming ψ_t is always positive. This assumption holds when productivity coefficients are non-negative, since the marginal utility of harvest is always positive. If young tree productivity were negative (which might occur if planting costs are substantial), additional land might not be valuable. However, in such a case, the grower would have no incentive to maintain the orchard at all.

3.1.2 Interpreting λ_t

The variable λ_t represents the marginal increase in the orchard's value in period t from relaxing the aging constraint between periods t and $t+1$. This constraint is relaxed by planting additional young trees in period t.

Equation (2b) shows the marginal benefit in period t from increasing the land allocated to young

trees. It consists of three components: the immediate marginal benefit from additional young trees in period t, the future marginal benefit from additional young trees in period t, and the immediate cost of using scarce land.

$$\beta^{-t}\frac{\partial \mathcal{L}}{\partial x_{yt}} = \underbrace{u'(c_t)f_y}_{\substack{\text{Current}\\ \text{marginal}\\ \text{benefit}}} \quad \underbrace{+\lambda_t}_{\substack{\text{Future}\\ \text{marginal}\\ \text{benefit}}} \quad \underbrace{-\psi_t}_{\substack{\text{Marginal}\\ \text{cost of}\\ \text{land}}}$$

Growing an additional young tree in period t allows additional old trees to be grown in period $t+1$. Assuming an interior solution ($x_{yt}, x_{ot} > 0$ and rearranging equation (2c) we see that the net marginal benefit of an additional old tree in period $t+1$ (valued in period t) is the marginal utility from an old tree, less the cost of allocating the land to that tree

$$\lambda_t = \beta(u'(c_{t+1})f_o - \psi_{t+1})$$

When the total land constraint binds ($\psi_t > 0$), the aging multiplier becomes

$$\lambda_t = \beta(u'(c_{t+1})f_o - u'(c_{t+1})f_y - \lambda_{t+1}))$$

This expression shows that the marginal benefit of an additional old tree in period $t+1$ equals the marginal benefit of that old tree minus the forgone marginal benefit of a young tree in period $t+1$ (which includes both its harvest and the option for having an old tree in period $t+2$). This formulation is key for analyzing the stationary solutions to the two-age-class model.

3.2 Characterizing optimal allocation sequences

We now characterize optimal allocation sequences for the two-age-class model, first identifing a necessary and sufficient condition for a two-period cyclical solution, with the steady-state (one-period cycle) emerging as a special case. We focus on interior solutions with a binding total land constraint, where $x_{ot} > 0$ and $L - x_{yt} - x_{ot} = 0$, ensuring that equation (2c) is satisfied with equality and $\psi_t > 0$ for all t.

In a two-period cycle, the land allocation will repeat every two periods, so $x_{ot} = x_{o,t+2}$. If the land allocations repeat, the harvest values must also repeat, $c_t = c_{t+2}$. To show that such a cycle is optimal, we must find a set of non-negative λ_t to satisfy the KKT conditions.

Throughout this analysis we assume that old trees are more productive than young trees,[5] $f_o > f_y$, and that $0 < \beta < 1$. This assumption corresponds to the case where the single-tree Faustmann optimal replacement age is m. When the Faustmann age is y (i.e., when $f_y > f_o$), the orchard management problem becomes trivial: the optimal policy is to allocate all land to young trees, since they are more productive. Any initial allocation with old trees would transition to an all-young-tree allocation in the first period, making the dynamic optimization problem uninteresting

Proposition 3.1. *For all $f_y < f_o$ and $0 < \beta < 1$, a land allocation sequence $\{x_{yt}, x_{ot}\}_{t=1}^{\infty}$ that follows a two-period cycle is optimal if and only if $\beta \leq \frac{u'(c_{t+1})}{u'(c_t)} \leq \frac{1}{\beta}$ for all t.*

Proof. See appendix A.3.1. □

Corollary 1. *For all $f_y < f_o$ and $0 < \beta < 1$, the even-aged sequence with $x_{ot} = \frac{L}{2}$ for all t is an optimal solution to the two-age-class orchard management problem.*

Proof. See Appendix A.3.2. □

3.3 Intuition for the existence of cyclical sequences

Why would cyclical production be optimal when a constant, even-aged allocation is also available? To understand the optimality of cycles, consider the marginal value of deviating from an existing cycle.

Let a be the fraction of land allocated to young trees in period t, $a = x_{yt}$, and, without loss of generality, that $a < \frac{L}{2}$. Assume that the orchard at time t follows the cyclical land allocation, $\{\boldsymbol{x}_t, \boldsymbol{x}_{t+1}, \ldots\} = \{(a, L-a), (L-a, a), \ldots\}$, with the corresponding consumption sequence $\{c_t, c_{t+1}, \ldots\} = \{f_y a + f_o(L-a), f_y(L-a) + f_o\, a, \ldots\}$. From the period t, the value of this orchard

[5] If $f_y > f_o$ there cannot be any old trees in a stationary solution to the problem. This yield assumption implies that equation (2c) must be strictly negative, which implies that x_{ot} must be zero by the complementary slackness condition (2f).

is

$$\beta^{-t}V_t = u(c_t) + \beta u(c_{t+1}) + \beta^2 u(c_t) + \dots$$

This cyclical sequence will be optimal if the grower has no incentive to adjust it. There are two ways to adjust: increase young trees in period $t+1$ by replacing young trees at the end of period t, or increase young trees in period $t+2$ by replacing young trees at the end of period $t+1$.

If the marginal change in value from a marginal increase in young trees in both period $t+1$ and $t+2$ is non-positive, then the cycle will be optimal.

Focusing on the change in young trees in period $t+1$, the marginal change in value is

$$\begin{aligned}
\beta^{-t}\frac{\partial V_t}{\partial x_{y,t+1}} &= 0 + \beta u'(c_{t+1})(f_y - f_o) + \beta^2 u'(c_{t+2})(f_o - f_y) + \dots && (3)\\
&= \frac{\beta(f_o - f_y)}{1-\beta^2}(\beta u'(c_{t+2}) - u'(c_{t+1}))\\
&= \frac{\beta(f_o - f_y)}{1-\beta^2}(\beta u'(c_t) - u'(c_{t+1})) && \text{(because } c_t = c_{t+2}\text{)}
\end{aligned}$$

There is no incentive to adjust the area of young trees in period $t+1$ if this expression is non-positive

$$\beta^{-t}\frac{\partial V_t}{\partial x_{y,t+1}} \leq 0 \Leftrightarrow \frac{u'(c_{t+1})}{u'(c_t)} \geq \beta$$

Similarly for increasing young trees in period $t+2$

$$\beta^{-t}\frac{\partial V_t}{\partial x_{y,t+2}} \leq 0 \Leftrightarrow \frac{u'(c_{t+1})}{u'(c_t)} \leq \frac{1}{\beta}$$

Combining these inequalities gives the same restriction on the ratio of marginal utilities derived from the KKT conditions in proposition 3.1.

Figure 3 shows utility as a function of the area allocated to young trees, x_y, assuming all land is used. This figure also assumes that $a < \frac{L}{2}$. When there are few young trees, $x_y = a$, there are many old trees, $x_o = L - a$, harvest is $(c_t = f_y a + f_o(L-a))$ is high, total utility is high, and

marginal utility is low; *vice versa* when there are many young trees. Proposition 3.1 requires the ratio of the marginal utilities between each two periods to be sufficiently close to 1 and the land allocations sufficiently close to $\frac{L}{2}$.

The grower faces competing incentives from time and consumption smoothing preferences. Time preference (positive discount rate) makes the grower willing to trade more future consumption for less present consumption. Consumption smoothing preference gives greater utility from more consistent year-to-year consumption.

These preferences can either conflict or align. The grower can increase utility in period $t+2$ by replacing young trees at the end of period t, increasing old trees and thus consumption in period $t+2$. However, this reduces utility in period $t+1$ by decreasing old trees that period. Whether the utility gain in $t+2$ outweighs the loss in $t+1$ depends on the discounting effect. Additionally, this operation may either increase or decrease consumption variation depending on the orchard's age-structure in period t. If replacing young trees reduces variation, the smoothing benefit is weighed against discounting losses. If it increases variation, both effects work against early replacement. Section 3.5 explores how initial conditions affect these trade-offs.

Figure 4 illustrates an example where the grower has eight trees each year with $x_{y,0} = \frac{5}{8}$. This grower is facing discrete replacement decisions, but it serves to illustrate the intuition behind whether to replace trees early, which is the same as the marginal case above. Subfigure 4a shows the trajectory of this orchard if the grower follows Faustmann replacement only and does not engage in smoothing.[6] In contrast, subfigure 4b shows the same initial orchard where the grower engages in smoothing by replacing one of the young trees. By engaging in early replacement, the grower achieves a even-aged orchard in period 2, but has lower production in period 1. Whether this operation increases the grower's utility depends on the discount factor.

[6] Faustmann replacement is defined as a replacement strategy where trees are replanted if and only if their age is equal to i^*, i.e. the age that maximizes the discounted net present value of an infinite sequence of orchards, each replaced at age i.

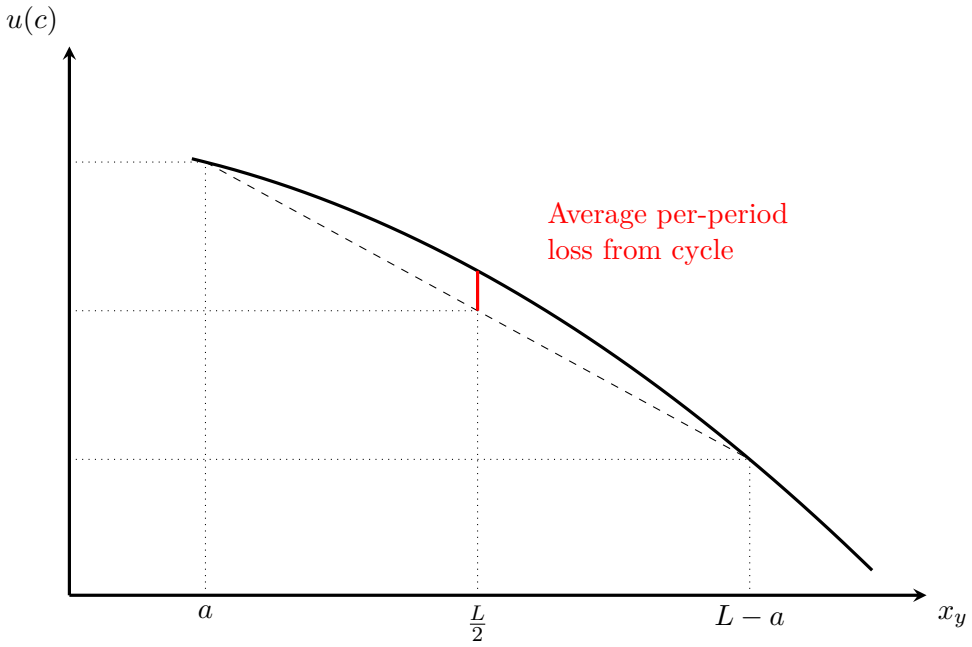


Figure 3: Cyclical production leads to a loss in average utility.

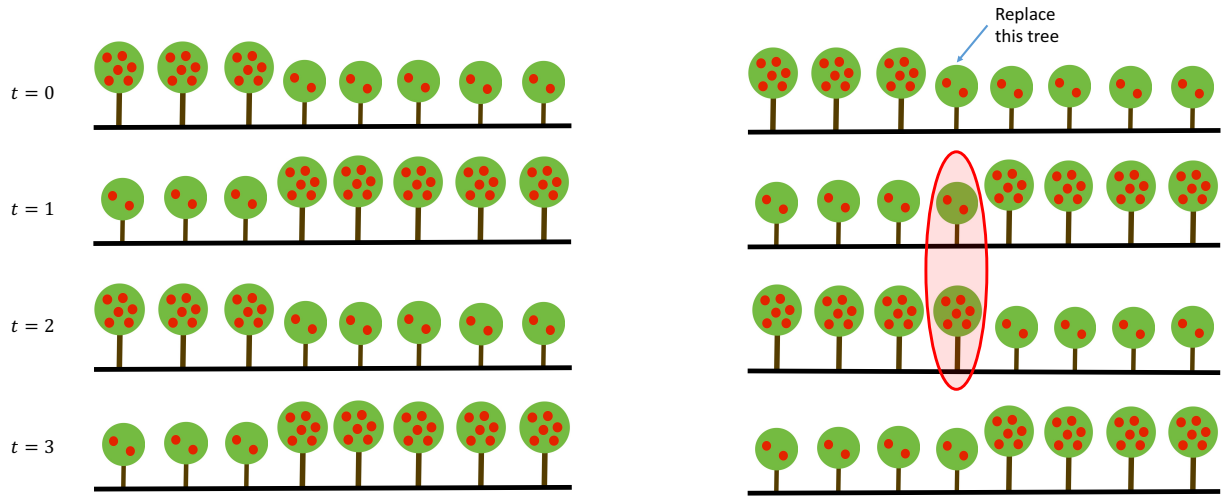
3.4 Comparative statics of cycle amplitude

The inequalities in proposition 3.1 define a set of land allocations from which the optimal sequence of allocations is cyclical. How does the size of this set change with the discount factor, the total area of land, and the relative productivity of young and old trees? First we need to identify this set, and then calculate comparative statics to examine how it changes with parameters.

To analyze how this set changes with parameters, we first revisit the relationship between land allocations and harvest levels. Let $c(x_{yt})$ be the harvest when x_{yt} units of land are planted with young trees. Using this definition, the inequalities from proposition 3.1 become

$$\beta \leq \frac{u'(c(x_{y,t+1}))}{u'(c(x_{yt}))} \leq \frac{1}{\beta}$$

In a two-period cycle, one period will have higher harvest than the other. Without loss of generality, we can label the periods so that the cycle peak (higher harvest) occurs in period t, which means $c_t > c_{t+1}$ and therefore $u'(c_t) < u'(c_{t+1})$ (since u is strictly concave). Given that $f_o > f_y$, period t has higher harvest because it has a larger allocation of old trees relative to young trees than



(a) Without smoothing – no short term loss, but long run loss from larger cycle.

(b) With smoothing – grower forgoes an old tree in period 2, but smoother in the long run.

Figure 4: The age-structure of an orchard without and with smoothing.

period $t+1$. Hence the right hand inequality is the relevant inequality. Requiring this inequality to be satisfied with equality gives the largest difference in marginal utilities such that a cycle will be optimal. Writing the allocation of young trees in terms of deviations from the even-aged orchard, $x_{yt} = \frac{L}{2} - \phi$ and $x_{y,t+1} = \frac{L}{2} + \phi$, implicitly defines the maximum cycle amplitude given the parameters of the model.

$$\frac{u'(c(\frac{L}{2} - \phi))}{u'(c(\frac{L}{2} + \phi))} = \beta \tag{4}$$

Using this implicit definition of the maximum cycle amplitude, we can use the implicit function theorem to find the comparative statics of the maximum cycle amplitude with respect to the parameters of the orchard management problem.

Proposition 3.2. *The comparative statics of the maximum cycle amplitude with respect to the parameters β, f_y and f_o are presented in table 1, where $A(c) = \frac{-u''(c)}{u'(c)}$ measures the preference for consumption smoothing.*[7]

[7] This has the same mathematical form as the Arrow-Pratt measure of absolute risk aversion (Pratt, 1964), but in our deterministic context it captures the strength of preference for avoiding consumption variability over time.

Proof. See appendix A.3.3 □

The preference for consumption smoothing $A(c) = \frac{-u''(c)}{u'(c)}$ appearing in these comparative statics captures how strongly the grower values stable production levels. Larger values of $A(c)$ represent a stronger smoothing preference. This parameter reflects the curvature of the utility function and governs intertemporal substitution behavior in deterministic models. In our orchard context, higher values of $A(c)$ indicate stronger aversion to production variability, with the comparative statics in Table 1 reflecting how the intensity of smoothing preferences interacts with other parameters to determine optimal cycle amplitude.

α	$\frac{\partial \phi}{\partial \alpha}$
β	(< 0)
f_y	$(> 0) \Leftrightarrow \frac{A(c(\frac{L}{2}+\phi))}{A(c(\frac{L}{2}-\phi))} < \frac{(\frac{L}{2}+\phi)}{(\frac{L}{2}-\phi)}$
f_o	$(< 0) \Leftrightarrow \frac{A(c(\frac{L}{2}+\phi))}{A(c(\frac{L}{2}-\phi))} < \frac{(\frac{L}{2}-\phi)}{(\frac{L}{2}+\phi)}$
L	$(> 0) \Leftrightarrow \frac{A(c(\frac{L}{2}+\phi))}{A(c(\frac{L}{2}-\phi))} > 1$

Table 1: Signs of comparative statics of cycle amplitude, ϕ, with respect to β, f_y, and f_o. Here $A(c)$ measures the preference for consumption smoothing.

These results provide intuition for how economic parameters affect the willingness to accept production cycles. More patient growers (higher β) always choose smaller cycle amplitudes since they place greater weight on future periods, strengthening the smoothing motive relative to immediate benefits from optimal replacement timing. The effects of productivity parameters depend on how the strength of smoothing preferences varies with consumption levels. When the smoothing preference measure is higher at higher consumption levels relative to the consumption difference, increasing young-tree productivity leads to larger optimal cycles, while increasing old-tree productivity leads to smaller cycles. This reflects how the marginal value of smoothing changes along the cycle. Orchard managers with greater utility curvature face a greater trade-off between replacement timing and production stability, with the optimal response depending on the specific consumption

levels achieved at different points in the cycle.

3.5 Converging to the cycle

We now examine how an orchard transitions from an arbitrary initial state to the optimal cycle. Proposition 3.3 defines the optimal transition rule between periods, which identifies the trajectory to the steady-state from any initial orchard.

Proposition 3.3. *Assuming the land constraint binds every period and letting x_{ot} be the land allocated to old trees in period t, the optimal transition rule is given by*

$$x_{o,t+1} = P(x_{o,t}) = \begin{cases} \frac{L}{2} + \phi & \text{for } x_{o,t} \in [0, \frac{L}{2} - \phi) \\ L - x_{o,t} & \text{for } x_{o,t} \in [\frac{L}{2} - \phi, \frac{L}{2} + \phi] \\ L - x_{o,t} & \text{for } x_{o,t} \in (\frac{L}{2} + \phi, L] \end{cases}$$

where ϕ is the maximum cycle amplitude, as defined in equation (4).

Proof. See appendix A.3.4. □

Figure 5 shows the optimal transition map between old trees in period t and old trees in period $t+1$. The horizontal axis denotes the area allocated to old trees in period t and the vertical axis denotes old trees in period $t+1$. This diagram is drawn assuming that the total land constraint binds every period. The downward sloping dashed line represents the aging constraint, $x_{o,t+1} \leq L - x_{ot}$. The bold black line represents the optimal transition rule, showing the optimal area allocated to old trees in period $t+1$ given an allocation of old trees in period t.

For allocations of old trees larger than $L - \phi$, the aging constraint binds. However, when there are few old trees in period t (i.e. $x_{ot} < L - \phi$), some young trees are replaced at the end of period t so $x_{o,t+1} < L - x_{ot}$. When there are many young trees, the opportunity cost of replacing a young tree early is low.

An implication of this diagram is that any optimal sequence $\{\mathbf{x}_t\}_{t=1}^{\infty}$ will converge to an optimal cycle in at most two periods. If the initial orchard has $x_{ot} < L - \phi$, then an optimal cycle will be

reached in the next period. If the initial orchard has $x_{ot} > L + \phi$, then an optimal cycle will be reached in two periods.

3.6 Comparison to Wan (1993) and forestry models

Our characterization of the optimal transition rule for a unit-area, two-age-class orchard both clarifies and extends earlier vintage-capital results.

First, in a two-age orchard with higher productivity of old trees $f_o > f_y$, Wan (1993, p. 418) obtains the "boundary-cycle" policy $x_{o,t} = 1 - x_{o,t-1}$. Wan's derivation implicitly assumes linear utility so that the replacement shadow price is everywhere positive (the equivalent in our model is λ_t, the Lagrange multiplier on the aging constraint (1c)). Under strictly concave $u(\cdot)$ the shadow price is zero whenever the orchard is sufficiently uneven-aged, giving rise to an interior regime that Wan's rule omits, as seen in the three regime transition map in our proposition 3.3. Region I ($x_{o,t} < L/2 - \phi$) features early replacement of part of the young cohort; the first-order condition $u'(c_t) = \beta u'(c_{t+1})$ with an inactive aging constraint pins the next-period stock at the cycle boundary $x_{o,t+1} = L/2 + \phi$. Regions II–III, where $\lambda_t > 0$, coincide with Wan's downward-sloping aging constraint $x_{o,t+1} = L - x_{o,t}$.

Second, our results illuminate how the replacement motive differs between perennial crops and forests. In the two-age forestry model of Salo and Tahvonen (2002), utility arises only when timber is harvested; standing trees provide no contemporaneous benefit. When the stock of old stands is low, marginal harvest utility is high and the optimal map is upward-sloping—part of the young cohort is logged to raise current harvest and enlarge the stock of old trees next period. As the old tree stock grows, this incentive weakens and the map bends down in a piecewise fashion until it follows the aging constraint $x_{o,t+1} = L - x_{o,t}$.

In contrast, in orchards fruit is harvested before any replacement decision, so replacing a young tree yields no immediate return; early replacement is valued only for smoothing future harvests. The resulting map is flat at $x_{o,t+1} = L/2 + \phi$ in Region I and follows the aging constraint thereafter. If the initial old-tree stock lies below $L/2 - \phi$, this flat segment sets $x_{o,1} = L/2 + \phi$. If instead

$x_{o,0} > L/2 + \phi$, the aging-constraint sends the system to $x_{o,1} = L - x_{o,0} < L/2 - \phi$, and the following period applies the flat segment, yielding $x_{o,2} = L/2 + \phi$. Thus every orchard sequence with $i^* = 2$ enters the two-point cycle $\{L/2 - \phi,\ L/2 + \phi\}$ in at most two periods, independent of parameters. To our knowledge, no other vintage-capital model of perennial crop dynamics has documented an explicit bound on convergence to equilibrium. On the other hand, Proposition 3 of (Salo and Tahvonen, 2002) shows only that forest sequences converge in finite time, with the number of periods increasing in both the distance of the initial stock from the cycle and reductions in their harvest return parameter a. Orchard adjustment is therefore generally faster and less sensitive to initial conditions.

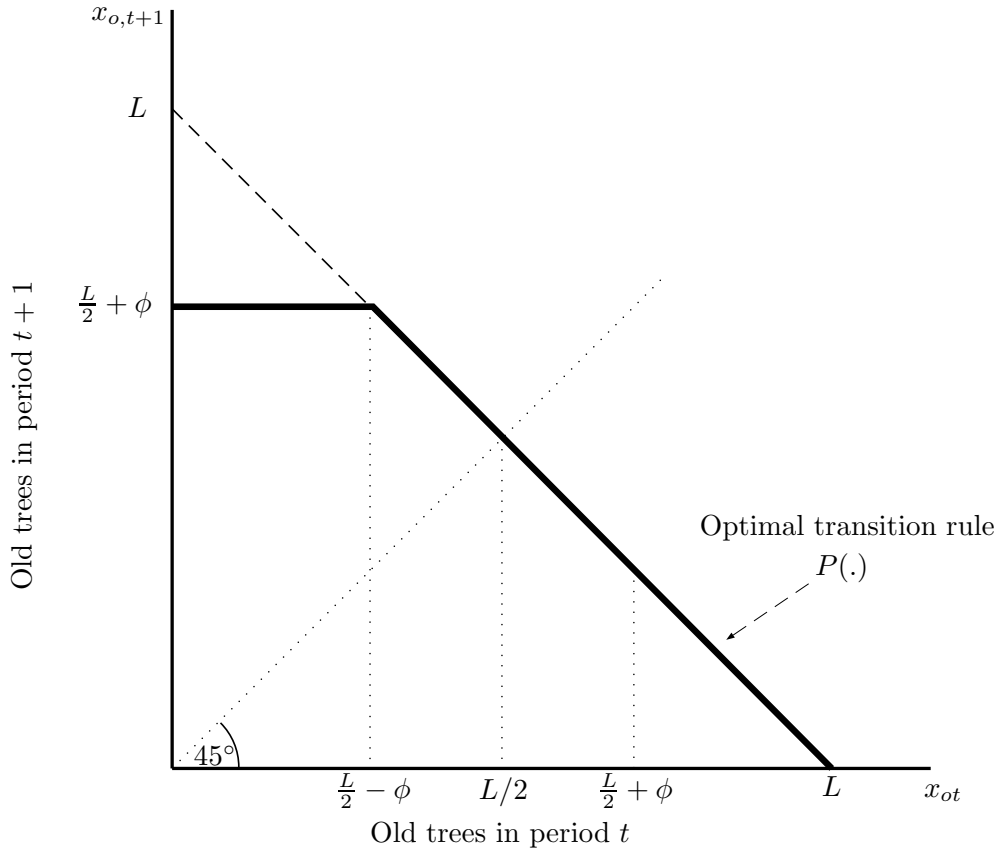


Figure 5: Transition map showing the optimal transition rule (solid line) for old trees in period t to old trees in period $t+1$ for a two-age-class orchard model. The aging constraint is shown by the dashed line. The cycle region is given by the set $[\frac{L}{2} - \phi, \frac{L}{2} + \phi]$.

4 A three-age-class orchard model with non-monotonic yields

We now extend the model to include a third age-class, which allows us to capture the biological reality of non-monotonic yields in perennial crops. In this model, trees live for three periods before dying. The three age-classes, indexed by i, are labeled: young (y), mature (m), and old (o) trees. As before, the grower's objective is to maximize the discounted benefits from the stream of harvests from each of type tree over an infinite time horizon. Assume the per-period payoff, $u(c_t(\mathbf{x}_t))$ is C^2, with $u'(c_t) > 0$, $u''(c_t) < 0$. Without loss of generality, normalize the total area of available land, L, to 1. See section A.4 for a glossary of notation used in this section.

$$V(x_{y0}, x_{m0}, x_{o0}) = \max_{\mathbf{x}_t} \sum_{t=1}^{\infty} \beta^t u(c_t(\mathbf{x}_t)) \tag{5a}$$

subject to

$$c_t \equiv f_y x_{yt} + f_m x_{mt} + f_o x_{ot}$$

$$x_{m,t+1} \leq x_{yt} \tag{5b}$$

$$x_{o,t+1} \leq x_{mt} \tag{5c}$$

$$x_{yt} + x_{mt} + x_{ot} \leq 1 \tag{5d}$$

$$x_{it} \geq 0 \quad \text{for } i \in \{y, m, o\} \tag{5e}$$

The Lagrangian of the reduced problem is

$$\mathcal{L} = \sum_{t=1}^{\infty} \beta^t \left[u(c_t) + \lambda_{1t}(x_{yt} - x_{m,t+1}) + \lambda_{2t}(x_{mt} - x_{o,t+1}) + \psi_t(1 - x_{yt} - x_{mt} - x_{ot}) + \sum_{i \in \{y,m,o\}} s_{it}(x_{it}) \right]$$

This leads to the KKT conditions for $t \geq 1$

$$\beta^{-t}\frac{\partial \mathcal{L}}{\partial x_{yt}} = u'(c_t)f_y + \lambda_{1t} - \psi_t - s_{yt} = 0 \tag{6a}$$

$$\beta^{-t}\frac{\partial \mathcal{L}}{\partial x_{mt}} = u'(c_t)f_m - \frac{\lambda_{1,t-1}}{\beta} + \lambda_{2t} - \psi_t - s_{mt} = 0 \tag{6b}$$

$$\beta^{-t}\frac{\partial \mathcal{L}}{\partial x_{ot}} = u'(c_t)f_o - \frac{\lambda_{2,t-1}}{\beta} - \psi_t - s_{ot} = 0 \tag{6c}$$

$$\lambda_{1t} \geq 0;\ \lambda_{1t}(x_{yt} - x_{m,t+1}) = 0 \tag{6d}$$

$$\lambda_{2t} \geq 0;\ \lambda_{2t}(x_{mt} - x_{o,t+1}) = 0 \tag{6e}$$

$$\psi_t \geq 0;\ \psi_t(1 - x_{yt} - x_{mt} - x_{ot}) = 0 \tag{6f}$$

$$x_{it} \geq 0;\ s_{it}(x_{it}) = 0 \quad \text{for } i \in \{y, m, o\} \tag{6g}$$

The key distinction of this model is that it allows for hump-shaped yield patterns, where ($f_m > f_y$) and ($f_m > f_o$). This is a common feature of many perennial crops, where productivity first increases as trees mature, then decreases in old age. Such non-monotonic yield patterns create more complex economic trade-offs than in the two-age-class model.

We focus on the case where the single-tree Faustmann optimal replacement age is o. When the Faustmann age is y, the problem reduces to maintaining only young trees (since they are most productive), making the multi-age-class dynamics trivial. When the Faustmann age is m, the long-run dynamics reduce to cycling between only young and mature trees - essentially replicating the two-age-class model analyzed in the previous section. Only when the Faustmann age is o do we obtain the three-period cyclical dynamics that justify the full three-age-class analysis.

Definition 1 (Faustmann transitions)**.** A Faustmann transition is where the grower replaces all trees whose age is greater than or equal to the Faustmann age. For $i^* = o$, a Faustmann transition simply applies the aging map, $\mathcal{F}(x_y, x_m, x_o) = (x_o, x_y, x_m)$. △

4.1 Properties of the value function

Lemma 2 (Basic Properties of the Value Function)**.** *Under the assumptions specified above, the value function $V : S \to \mathbb{R}$ has the following properties:*

1. Existence and uniqueness *There exists a unique bounded function V satisfying the Bellman equation*

$$V(\mathbf{x}_t) = \max_{\mathbf{x}_{t+1} \in \Gamma(\mathbf{x}_t)} \{u(c(\mathbf{x}_t)) + \beta V(\mathbf{x}_{t+1})\}$$

2. Continuity*: V is continuous on S*

3. Concavity*: V is weakly concave on S*

Proof. Existence, uniqueness, boundedness, and continuity: Define the Bellman operator T on the space of bounded continuous functions $\mathcal{B}(S)$ by

$$(Tf)(\mathbf{x}_t) = \max_{\mathbf{x}_{t+1} \in \Gamma(\mathbf{x}_t)} \{u(c(\mathbf{x}_t)) + \beta f(\mathbf{x}_{t+1})\}$$

Since u is bounded on compact sets, S is compact, and $\Gamma(\mathbf{x}_t)$ is compact-valued and continuous, the operator T is well-defined and maps $\mathcal{B}(S)$ into itself. Moreover, T satisfies Blackwell's sufficient conditions for a contraction: it is monotone and satisfies the discounting property with modulus $\beta \in (0, 1)$. By Theorem 3.2 in (Stokey et al., 1989), T has a unique fixed point $V \in \mathcal{B}(S)$, establishing existence, uniqueness, boundedness, and continuity.

Concavity: We show that T preserves concavity. Let $f : S \to \mathbb{R}$ be bounded and concave. For arbitrary $\mathbf{x}^0, \mathbf{x}^1 \in S$ and $\lambda \in [0, 1]$, define $\mathbf{x}^\lambda = \lambda \mathbf{x}^0 + (1 - \lambda)\mathbf{x}^1$. Let $\mathbf{x}^0_{t+1}$ and $\mathbf{x}^1_{t+1}$ achieve the maxima defining $(Tf)(\mathbf{x}^0)$ and $(Tf)(\mathbf{x}^1)$, respectively.

Since the constraint correspondEnce Γ has a convex graph (the constraints are linear), we have $\mathbf{x}^\lambda_{t+1} = \lambda \mathbf{x}^0_{t+1} + (1 - \lambda)\mathbf{x}^1_{t+1} \in \Gamma(\mathbf{x}^\lambda)$. Therefore:

$$(Tf)(\mathbf{x}^\lambda) \geq u(c(\mathbf{x}^\lambda)) + \beta f(\mathbf{x}^\lambda_{t+1})$$

Since u is concave and c is linear in $\mathbf{x}$:

$$u(c(\mathbf{x}^\lambda)) \geq \lambda u(c(\mathbf{x}^0)) + (1-\lambda)u(c(\mathbf{x}^1))$$

By concavity of f:

$$f(\mathbf{x}^\lambda_{t+1}) \geq \lambda f(\mathbf{x}^0_{t+1}) + (1-\lambda)f(\mathbf{x}^1_{t+1})$$

Combining these inequalities:

$$(Tf)(\mathbf{x}^\lambda) \geq \lambda (Tf)(\mathbf{x}^0) + (1-\lambda)(Tf)(\mathbf{x}^1)$$

Thus T preserves concavity. Since the zero function is concave and $V = \lim_{n\to\infty} T^n 0$, and since concavity is preserved under uniform limits on compact sets (Stokey et al., 1989, , Theorem 4.8 under weak inequality), V is (weakly) concave. □

Lemma 3 (Properties of the Optimal Policy Correspondence)**.** *Let the feasible set correspondence $\Gamma : S \rightrightarrows S$ be defined by the constraints (5b)–(5e), and by assumption the utility function u is continuous and strictly increasing. Then the value function,*

$$V(\mathbf{x}_t) := \max_{\mathbf{x}_{t+1}\in\Gamma(\mathbf{x}_t)} f(\mathbf{x}_t, \mathbf{x}_{t+1})$$

is continuous and the optimal policy correspondence

$$\boldsymbol{\xi}^S(\mathbf{x}_t) := \arg\max_{\mathbf{x}_{t+1}\in\Gamma(\mathbf{x}_t)} \{u(c(\mathbf{x}_t)) + \beta V(\mathbf{x}_{t+1})\}$$

is nonempty, compact-valued, and upper hemicontinuous on the state space.[8]

[8]The superscript S indicates that the policy correspondence is defined across the full state space, S. The remainder of the proof will focus on a subset of this correspondence.

Proof. Define the objective function

$$f(\mathbf{x}_t, \mathbf{x}_{t+1}) := u(c(\mathbf{x}_t)) + \beta V(\mathbf{x}_{t+1})$$

where $c(\mathbf{x}_t) = f_y x_{yt} + f_m x_{mt} + f_o x_{ot}$ is linear and $u(\cdot)$ is continuous and strictly increasing. By Lemma 2, the value function $V : S \to \mathbb{R}$ is continuous and bounded. Hence, $f(\mathbf{x}_t, \mathbf{x}_{t+1})$ is continuous in both arguments.

The correspondence $\Gamma : S \rightrightarrows S$ is defined by the system of linear constraints (5b)–(5e), so each constraint is affine and continuous in $\mathbf{x}_t$ and Γ is non-empty and compact-valued. Therefore, Γ is continuous (de la Fuente, 2000, , Theorem 2.2).

By Berge's Maximum Theorem (de la Fuente, 2000, , Theorem 2.1), since $f(\mathbf{x}_t, \mathbf{x}_{t+1})$ is continuous in both $\mathbf{x}_t$ and $\mathbf{x}_{t+1}$, and $\Gamma(\mathbf{x}_t)$ is nonempty, compact-valued,

$$V(\mathbf{x}_t) := \max_{\mathbf{x}_{t+1} \in \Gamma(\mathbf{x}_t)} f(\mathbf{x}_t, \mathbf{x}_{t+1})$$

is continuous, and the optimal policy correspondence

$$\boldsymbol{\xi}^S(\mathbf{x}_t) := \arg \max_{\mathbf{x}_{t+1} \in \Gamma(\mathbf{x}_t)} f(\mathbf{x}_t, \mathbf{x}_{t+1})$$

is nonempty-valued, compact-valued, and upper hemicontinuous. □

4.2 Stationary solutions to the three-age-class model

Our first major result concerns the existence and characterization optimal stationary three-period cycles in the three-age-class orchard model, assuming the land constraint is always binding. We begin by listing the dual variables generated by a candidate cycle $\{\mathbf{x}_t, \mathbf{x}_{t+1}, \mathbf{x}_{t+2}, \ldots\}$.

$$\begin{bmatrix} \lambda_{1t} \\ \lambda_{1,t+1} \\ \lambda_{1,t+2} \end{bmatrix} = \frac{\beta}{1-\beta^3} \begin{bmatrix} \beta^2(f_y - f_o) & (f_m - f_y) & \beta(f_o - f_m) \\ \beta(f_o - f_m) & \beta^2(f_y - f_o) & (f_m - f_y) \\ (f_m - f_y) & \beta(f_o - f_m) & \beta^2(f_y - f_o) \end{bmatrix} \begin{bmatrix} u'(c_t) \\ u'(c_{t+1}) \\ u'(c_{t+2}) \end{bmatrix} \tag{7a}$$

$$\begin{bmatrix} \lambda_{2t} \\ \lambda_{2,t+1} \\ \lambda_{2,t+2} \end{bmatrix} = \frac{\beta}{1-\beta^3} \begin{bmatrix} \beta^2(f_m - f_o) & (f_o - f_y) & \beta(f_y - f_m) \\ \beta(f_y - f_m) & \beta^2(f_m - f_o) & (f_o - f_y) \\ (f_o - f_y) & \beta(f_y - f_m) & \beta^2(f_m - f_o) \end{bmatrix} \begin{bmatrix} u'(c_t) \\ u'(c_{t+1}) \\ u'(c_{t+2}) \end{bmatrix}. \tag{7b}$$

Complementarity requires $s_{it}\, x_{it} = 0$ and $s_{it} \geq 0$. Because the matrices in (7) depend continuously on $\mathbf{x}_t$, all dual variables are continuous functions of the state.

Proposition 4.1 (Steady states of the three-age-class model)**.** *Assume $\beta \in (0,1)$ and a three-age-class yield vector $\mathbf{f} = (f_y, f_m, f_o)$ such that the single-tree Faustmann optimal replacement age is unique and equal to $i^* = o$, i.e.*

$$\frac{f_y}{1-\beta} < \frac{f_y + \beta f_m + \beta^2 f_o}{1-\beta^3}, \qquad \frac{f_y + \beta f_m}{1-\beta^2} < \frac{f_y + \beta f_m + \beta^2 f_o}{1-\beta^3}. \tag{8}$$

Denote the simplex by $\Delta^2 := \{\mathbf{x} \in \mathbb{R}^3 : x_y + x_m + x_o = 1,\ x_i \geq 0 \ \text{ for } i \in \{y, m, o\}\}$ and the aging map by $\mathcal{F}(x_y, x_m, x_o) = (x_o, x_y, x_m)$.

1. Even-Aged optimum*: The allocation $\bar{\mathbf{x}} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ is optimal.*
2. Local 3-cycles*: There exists $\varepsilon > 0$ and an open ball $\mathcal{U} := \{\mathbf{x} \in \Delta^2 : \|\mathbf{x} - \bar{\mathbf{x}}\| < \varepsilon\}$ such that every $\mathbf{x} \in \mathcal{U}$ generates the optimal 3-period cycle $\{\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}, \ldots\}$.*
3. Cycle region*: Define the cycle region*

   $K_0 := \left\{\mathbf{x} \in \Delta^2 : \text{the 3-cycle } \{\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}\} \text{ satisfies all KKT conditions}\right\}$

Equivalently,

$$K_0 := \left\{ \mathbf{x} \in \Delta^2 : \begin{array}{c} \boldsymbol{\lambda}_1(\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}) \geq \mathbf{0} \\ \boldsymbol{\lambda}_2(\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}) \geq \mathbf{0} \\ \mathbf{s}(\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}) \geq \mathbf{0} \end{array} \right\}$$

where $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2$ *are the* 3×1 *vectors from* (7a) *and* (7b)*, and* $\mathbf{s}$ *represents all slack variables across the cycle.*

Then K_0 *is closed and contains both* $\bar{\mathbf{x}}$ *and* $\mathcal{U}$*. Moreover, for any initial allocation* $\mathbf{x} \in \Delta^2$*, the 3-cycle* $\{\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}, \ldots\}$ *is optimal if and only if* $\mathbf{x} \in K_0$*.*

Proof. By Lemma 2, the value function V is weakly concave. The constraints defining the dynamic optimization problem—aging constraints (5b)–(5c), land constraint (5d), and non-negativity constraints (5e)—are all linear. Since the Linear Independence Constraint Qualification (LICQ) is satisfied on S, and the objective is concave, the KKT conditions (6) are both necessary and sufficient for optimality in this problem.

Even-Aged optimum: Consider the even-aged allocation $\bar{\mathbf{x}} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. The corresponding 3-cycle is $\{\bar{\mathbf{x}}, \bar{\mathbf{x}}, \bar{\mathbf{x}}\}$ since $\mathcal{F}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$. All three periods have equal consumption:

$$c_t = c_{t+1} = c_{t+2} = \frac{1}{3}(f_y + f_m + f_o) =: \bar{c}$$

Therefore, $u'(c_t) = u'(c_{t+1}) = u'(c_{t+2}) = u'(\bar{c})$.

From (7a):

$$\begin{bmatrix} \lambda_{1t} \\ \lambda_{1,t+1} \\ \lambda_{1,t+2} \end{bmatrix} = \frac{\beta u'(\bar{c})}{1-\beta^3} \begin{bmatrix} \beta^2(f_y - f_o) & (f_m - f_y) & \beta(f_o - f_m) \\ \beta(f_o - f_m) & \beta^2(f_y - f_o) & (f_m - f_y) \\ (f_m - f_y) & \beta(f_o - f_m) & \beta^2(f_y - f_o) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{9}$$

$$= \frac{\beta u'(\bar{c})}{1-\beta^3} \begin{bmatrix} \Delta_1 \\ \Delta_1 \\ \Delta_1 \end{bmatrix} \tag{10}$$

where $\Delta_1 = (f_m - f_y) + \beta(f_o - f_m) + \beta^2(f_y - f_o)$. Under the Faustmann condition (8), $\Delta_1 > 0$, so all three components $\lambda_{1t}, \lambda_{1,t+1}, \lambda_{1,t+2} > 0$.

Similarly for $\boldsymbol{\lambda}_2$ from (7b):

$$\begin{bmatrix} \lambda_{2t} \\ \lambda_{2,t+1} \\ \lambda_{2,t+2} \end{bmatrix} = \frac{\beta u'(\bar{c})}{1-\beta^3} \begin{bmatrix} \Delta_2 \\ \Delta_2 \\ \Delta_2 \end{bmatrix}$$

where $\Delta_2 = (f_o - f_y) + \beta(f_y - f_m) + \beta^2(f_m - f_o) > 0$ under the $i^* = 3$ assumption.

From (**??**) evaluated at $\bar{\mathbf{x}}$, all slack variables equal zero since $x_i = \frac{1}{3} > 0$: $s_y = s_m = s_o = 0$ for all three periods. All aging and land constraints bind with equality at $\bar{\mathbf{x}}$, and all non-negativity constraints are inactive ($x_i > 0$), so complementary slackness conditions are satisfied. Since all KKT conditions (6) are satisfied and the problem is concave, the even-aged allocation $\bar{\mathbf{x}}$ is optimal.

Local 3-cycles: The maps $\mathbf{x} \mapsto \boldsymbol{\lambda}_1(\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x})$, $\mathbf{x} \mapsto \boldsymbol{\lambda}_2(\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x})$, and $\mathbf{x} \mapsto \mathbf{s}(\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x})$ are continuous because the consumption functions $c_t(\mathbf{x}) = f_y x_y + f_m x_m + f_o x_o$, $c_{t+1}(\mathcal{F}\mathbf{x})$, $c_{t+2}(\mathcal{F}^2\mathbf{x})$ are linear, hence continuous, the utility function $u'(\cdot)$ is continuous by assumption, and the matrix operations and slack variable computations are continuous.

At $\bar{\mathbf{x}}$, all dual variables are strictly positive ($\lambda_{1i} > 0$, $\lambda_{2i} > 0$ for all i) and all primal variables are strictly feasible ($x_j > 0$ for all j). By continuity, there exists $\varepsilon > 0$ such that for the open ball

$\mathcal{U} := \{\mathbf{x} \in \Delta^2 : \|\mathbf{x} - \bar{\mathbf{x}}\| < \varepsilon\}$, all constraint inequalities defining K_0 remain satisfied. For any $\mathbf{x} \in \mathcal{U}$, we must verify that the 3-cycle $\{\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}\}$ satisfies the complete KKT system (6).

For $\mathbf{x} \in \mathcal{U}$, since $\lambda_{1i} > 0$ and $\lambda_{2i} > 0$ for all i by continuity from $\bar{\mathbf{x}}$, complementary slackness requires the aging constraints to bind:

$$\lambda_{1t}(x_{yt} - x_{m,t+1}) = 0 \Rightarrow x_{m,t+1} = x_{yt} \tag{11}$$

$$\lambda_{2t}(x_{mt} - x_{o,t+1}) = 0 \Rightarrow x_{o,t+1} = x_{mt} \tag{12}$$

This forces optimal transitions to follow the aging map $\mathcal{F}$, so the optimal trajectory is $\{\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}\}$. Since $x_j > 0$ for all j in $\mathcal{U}$, the non-negativity constraints are inactive ($s_j = 0$ for all j).

With aging constraints binding ($x_{m,t+1} = x_{yt}$, $x_{o,t+1} = x_{mt}$) and slack variables zero ($s_j = 0$), the first-order conditions (6a)–(6c) reduce to the matrix equations (7a) and (7b), which are satisfied for the 3-cycle by definition of $\mathcal{U}$. Since the complete KKT system is satisfied and the optimization problem is concave with linear constraints, the 3-cycle is optimal.

Cycle region: K_0 is defined as the intersection of sets of the form $\{\mathbf{x} : g_i(\mathbf{x}) \geq 0\}$ where each g_i represents a constraint function (components of $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$, or $\mathbf{s}$). Since each g_i is continuous, each set $\{\mathbf{x} : g_i(\mathbf{x}) \geq 0\}$ is closed as the preimage of the closed set $[0, \infty)$. The intersection of finitely many closed sets is closed, so K_0 is closed. By construction from the previous parts, $\bar{\mathbf{x}} \in K_0$ and $\mathcal{U} \subset K_0$.

If $\mathbf{x} \in K_0$, then by definition the 3-cycle $\{\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}\}$ satisfies all KKT conditions across all three periods. Since the dynamic optimization problem has a concave objective (Lemma 2) and linear constraints, the KKT conditions are sufficient for optimality. Therefore, the 3-cycle is optimal.

If the 3-cycle $\{\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}\}$ is optimal for the dynamic optimization problem, then by the necessary conditions for optimality, it must satisfy the KKT conditions (6) at each period of the cycle. This means all constraint inequalities defining K_0 must hold, hence $\mathbf{x} \in K_0$. □

This result establishes that the even-aged orchard (with equal land allocated to each age class) is optimal, but it is not the only optimal solution. In fact, there exists a continuum of initial

allocations (the cycle region, K_0) that generate optimal three-period cycles. This is analogous to the cycle region we found in the two-age-class model, but with a three-period cycle structure.

4.3 Convergence in the three-age-class model

Our second major result concerns how orchards with arbitrary initial age distributions converge to optimal cycles. We combine results from all four ridge segments to prove that every allocation reaches K_0 within $N = 6+N_1+N_2$ periods. The proof accounts for all starting locations: allocations not on the ridge reach it within three periods, then follow optimal transitions from the ridge.

Proposition 4.2. *For the problem defined by conditions (5a)-(5e), there exists $N = 6 + N_1 + N_2$ such that the optimal trajectory starting from any $\mathbf{x}_0 \in \Delta^2$ converges to K_0 in at most N periods (N_1 is defined in Corollary 14 and N_2 is defined in Corollary 26).*

The convergence proof is based on understanding the geometry of optimal transitions in the state space, which is represented by the simplex of possible land allocations.

The core of the proof is to manually construct the optimal policy function using Faustmann transitions ($\mathbf{x}_{t+1} = \mathcal{F}\mathbf{x}_t$) and 'ridge' transitions, where the 'ridge' is defined as the set of value maximizing allocations of old trees for each level of x_m. The optimal path from any initial allocation is characterized as follows. First, for any initial allocation outside the cycle region and not on the ridge, the optimal policy brings the orchard to the ridge within at most 3 periods. Second, once on the ridge, the optimal policy follows a structured transition path that we fully characterize, eventually entering the cycle region (K_0) in at most $N_1 + N_2 + 3$ periods.

4.3.1 Restricting focus to only replacing mature trees

This subsection establishes that under hump-shaped yields, the aging constraint $x_{m,t+1} \leq x_{yt}$ binds with equality whenever $x_{mt} > 0$. The proof uses the first-order conditions and the assumption that the Faustmann age is o (implying $f_m > f_y$ and $f_m > f_o$) to show it is never optimal to replace young trees early. This result reduces the optimization problem to choosing $x_{o,t+1}$ given $x_{m,t+1} = x_{yt}$, allowing us to focus exclusively on decisions about mature tree replacement.

Lemma 4. *Consider the dynamic forestry problem with hump-shaped yields satisfying $f_m > f_y$,*

$f_m > f_o$. If $x_{mt} > 0$, then the aging constraint $x_{m,t+1} \leq x_{yt}$ binds with equality.

Proof. From the first-order conditions (6b) and (6c), we have:

$$u'(c_t)f_m - \frac{\lambda_{1,t-1}}{\beta} + \lambda_{2t} - \psi_t - s_{mt} = 0 \tag{13}$$

$$u'(c_t)f_o - \frac{\lambda_{2,t-1}}{\beta} - \psi_t - s_{ot} = 0 \tag{14}$$

Subtracting equation (14) from equation (13) yields:

$$\frac{\lambda_{1,t-1}}{\beta} = u'(c_t)(f_m - f_o) + \lambda_{2t} + \frac{\lambda_{2,t-1}}{\beta} - s_{mt} + s_{ot}$$

Since $x_{mt} > 0$ by assumption, the boundary constraint for mature trees is non-binding, implying $s_{mt} = 0$ by complementary slackness. Under the hump-shaped yield assumption, $f_m > f_o$, and since $u'(c_t) > 0$, we have $u'(c_t)(f_m - f_o) > 0$.

Given that $\lambda_{2t} \geq 0$, $\lambda_{2,t-1} \geq 0$, and $s_{ot} \geq 0$, it follows that:

$$\frac{\lambda_{1,t-1}}{\beta} = u'(c_t)(f_m - f_o) + \lambda_{2t} + \frac{\lambda_{2,t-1}}{\beta} + s_{ot} > 0$$

Therefore, $\lambda_{1,t-1} > 0$. By the complementary slackness condition $\lambda_{1,t-1}(x_{yt} - x_{m,t+1}) = 0$, this implies $x_{yt} - x_{m,t+1} = 0$, establishing that the aging constraint binds with equality. □

Since Lemma 4 applies almost everywhere, to streamline notation we will often write the triple (x_y, x_m, x_o) as (x_m, x_o) with the understanding that x_y is implicitly defined by the condition $x_y = 1 - x_m - x_o$.

4.3.2 Defining transitions

Having established that only mature trees are replaced, we formalize the restricted transition set $\Gamma^r(\mathbf{x}_t)$ and introduce the ridge correspondence $R(x_m)$, which maps each mature tree allocation to the value-maximizing old tree allocations on the iso-mature set $\mathcal{M}(x_m)$. The ridge $\mathcal{R}$ constitutes

the optimal policy surface that trajectories seek to reach. We define the optimal transition function that chooses ridge transitions when $R(x_{yt}) \cap [0, x_{mt}] \neq \emptyset$ and Faustmann transitions via the aging map $\mathcal{F}$ otherwise.

Definition 2. Restricted Transition Sets

If the young-to-mature aging constraint binds with equality, $x_{m,t+1} = x_{yt}$, the transition set available to the grower from an allocation $\mathbf{x}_t = (x_{yt}, x_{mt}, x_{ot})$ is

$$\begin{aligned} x_{y,t+1} &= 1 - x_{m,t+1} + x_{o,t+1} \\ x_{m,t+1} &= x_{yt} \\ 0 \leq x_{o,t+1} &\leq x_{mt} \end{aligned}$$

Let $\Gamma^r(\mathbf{x}_t)$ denote the set of allocations satisfying (15). That is, $\mathbf{x_{t+1}} \in \Gamma^r(\mathbf{x}_t) \subset \Gamma(\mathbf{x}_t)$. This set can be drawn on the simplex as an iso-mature (horizontal) line extending from the $x_o = 0$ boundary towards the $x_y = 0$ boundary of the simplex. Figure 7 shows five example intital allocations and their restricted transition sets. △

Definition 3. The "Ridge": Correspondence of optimal x_o given x_m

Fix $x_m \in [0, 1]$. The iso-mature set is $\mathcal{M}(x_m) = \{\mathbf{x} \in \Delta^2 : \; x_{mt} = x_m, \; 0 \leq x_{ot} \leq 1 - x_m\}$.

Since the value function is weakly concave by Lemma 2, the set of maximizers may be non-singleton. Define the correspondence of optimal old-tree areas as

$$R(x_m) = \{x_o : (x_m, x_o) \in \arg \max_{(x_m, x_o) \in \mathcal{M}(x_m)} V(x_m, x_o)\}$$

That is, $R(x_m)$ maps the mature-tree share x_m to the set of optimal old-tree shares x_o, constrained on the $\mathcal{M}(x_m)$ line. A selection $r(x_m) \in R(x_m)$, or, if unambiguous, just r, is an arbitrary optimal value of x_o for the given x_m.

From Lemma 3, the correspondence $R(x_m)$ is non-empty, compact-valued, and upper hemicontin-

uous. Since $R(x_m)$ is compact-valued, it attains its maximum and minimum values; denote these by $\overline{r}(x_m) = \max R(x_m)$ and $\underline{r}(x_m) = \min R(x_m)$.

The ridge, $\mathcal{R}$, is defined as the graph of the correspondence R:

$$\mathcal{R} = \{(x_m, x_o) : x_m \in [0,1], x_o \in R(x_m)\}$$

Since R is non-empty, compact-valued, and upper hemicontinuous, and the domain $[0,1]$ is compact, the ridge $\mathcal{R}$ is non-empty, compact, and closed.

For a given mature tree allocation x_m, the *graph slice* at x_m is the set of all allocations on the ridge with that mature component:

$$\mathbf{R}(x_m) := \{(1 - x_m - x_o, x_m, x_o) : x_o \in R(x_m)\}$$

This represents the portion of the ridge $\mathcal{R}$ corresponding to a specific level of mature trees. Note the relationships: a selection $r \in R(x_m)$ corresponds to the allocation $(1 - x_m - r, x_m, r) \in \mathbf{R}(x_m)$, and the complete ridge is the union $\mathcal{R} = \bigcup_{x_m \in [0,1]} \mathbf{R}(x_m)$.

We call this correspondence the 'ridge' because when plotted as a surface over the simplex, the value function $V(x_m, x_o)$ exhibits a ridge-like peak along the graph of $R(x_m)$, where the value is maximized for each level of mature trees (see Figure 8). △

Figure 6 shows the transition set (where both young and mature trees can be replaced) and the restricted transition set (where only mature trees are replaced) from an example allocation. The example allocation is $a = (0.2, 0.2, 0.6)$. Assuming that either or both of young and mature trees can be replaced, if no young or mature trees are replaced, a will make a Faustmann transition to the allocation $(0.6, 0.2, 0.2)$; if all young and mature trees are replaced, a will transition to $(1, 0, 0)$; if only young trees are replaced, a will transition to $(0.8, 0, 0.2)$; and if only mature trees are replaced, a will transition to $(0.8, 0.2, 0)$. Any allocation in the set $\mathcal{A}$ can be reached as a transition from a by a convex combination of replacing young and mature trees. Assuming that only mature trees

can be replaced, allocation a can transition to any allocation on the line segment A, a subset of the region $\mathcal{A}$.

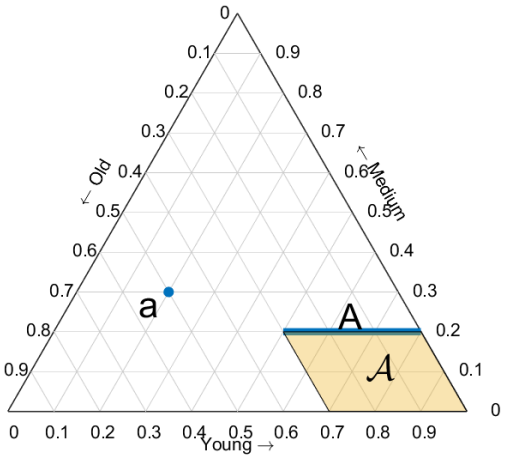


Figure 6: From allocation a, the unrestricted transition set is the shaded parallellogram $\mathcal{A}$, i.e. $\Gamma(a) = \mathcal{A}$, and the restricted transition set is the line segment A, i.e. $\Gamma^r(a) = A$.

Definition 4. Optimal Restricted Transition Correspondence: Ridge and Faustmann Transitions The optimal restricted transition from state $\mathbf{x}_t$ depends on whether the ridge correspondence intersects with the feasible transition set.

Recall that the restricted transition set from $\mathbf{x}_t = (x_{yt}, x_{mt}, x_{ot})$ is

$$\Gamma^r(\mathbf{x}_t) = \{(x_{y,t+1}, x_{m,t+1}, x_{o,t+1}) : x_{m,t+1} = x_{yt},\ x_{y,t+1} = 1 - x_{yt} - x_{o,t+1},\ 0 \leq x_{o,t+1} \leq x_{mt}\}$$

Since the aging constraint binds with $x_{m,t+1} = x_{yt}$, the grower chooses $x_{o,t+1}$ to maximize value given that period $t+1$ will have x_{yt} mature trees. The ridge correspondence $R(x_{yt})$ provides exactly these value-maximizing old tree allocations for the future mature tree level. The feasible range for

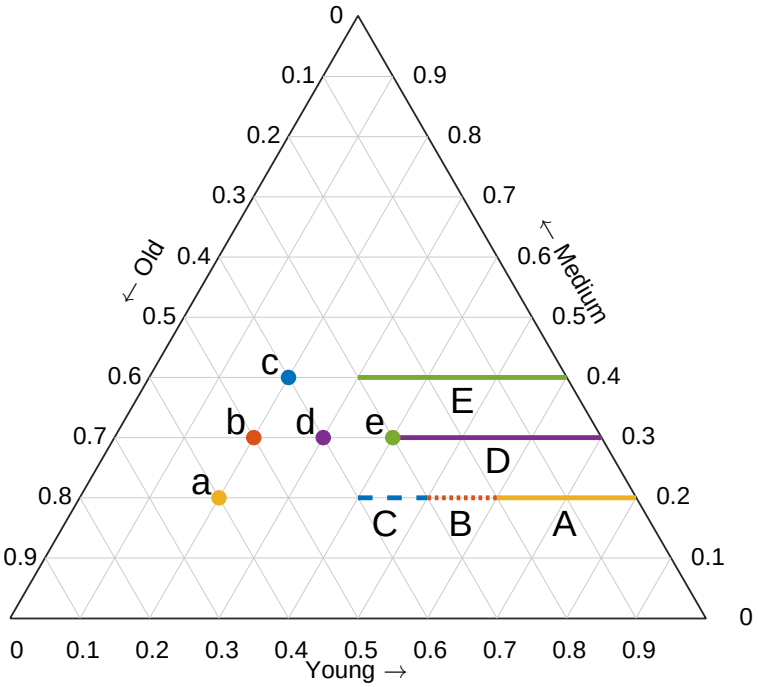


Figure 7: Example restricted transition sets. The restricted transition set of allocation a is line segment A, $\Gamma^r(a) = A$. Similarly, $\Gamma^r(b) = A+B$, $\Gamma^r(c) = A+B+C$, $\Gamma^r(d) = D$, and $\Gamma^r(e) = E$. As an allocation decreases its old allocation along an iso-young line (moves up), it's restricted transition set remains on the same iso-mature line, but becomes longer (a to b to c). As an allocation decreases its old allocation along an iso-mature line (moves right), its restricted transition set moves to a higher iso-mature line, but remains the same length (b to d to e).

old trees in the next period is $[0, x_{mt}]$. The optimal transition function is:

$$\mathbf{x}^*_{t+1} = \boldsymbol{\xi}(\mathbf{x}_t) \tag{16}$$

$$= \begin{cases} \text{Any } \mathbf{x}_{t+1} \in \mathbf{R}(x_{yt}) \cap \Gamma^r(\mathbf{x}_t) & \text{If } R(x_{yt}) \cap [0, x_{mt}] \neq \emptyset \quad \text{(Ridge transition)} \\ \mathcal{F}\mathbf{x}_t & \text{If } R(x_{yt}) \cap [0, x_{mt}] = \emptyset \quad \text{(Faustmann transition)} \end{cases}$$

where $\mathcal{F}(x_y, x_m, x_o) = (x_o, x_y, x_m)$ is the aging map and $\mathbf{R}(x_m)$ is the correspondence of value-maximizing allocations on the iso-mature set $\mathcal{M}(x_m)$.

When a ridge transition is feasible, the grower can reach the ridge by choosing any allocation in $\mathbf{R}(x_{yt}) \cap \Gamma^r(\mathbf{x}_t)$, which involves replacing some mature trees to achieve the value-maximizing old tree allocation for the next period's mature tree level. Equivalently, since $R(x_{yt}) \subseteq [0, 1 - x_{yt}]$, a ridge transition is feasible if and only if $\underline{R}(x_{yt}) \leq x_{mt}$.

When a Faustmann transition is optimal, the ridge is not reachable within the feasible transition set, so the grower follows the aging map $\mathcal{F}$. $\triangle$

4.3.3 Characterizing K_0

This subsection characterizes the cycle region K_0 as the set of allocations generating optimal three-period cycles satisfying all KKT conditions. We prove that $\mathrm{bd}(K_0)$ consists of three curves that are images of each other under $\mathcal{F}$, with one curve, named $\mathcal{R}_3$, coinciding with the ridge where $\lambda_{2,t+2} = 0$. We establish that $R(x_m)$ is singleton-valued within K_0 due to strict concavity within this region, and that no ridge points lie outside K_0 for the relevant domain D.

[Domain of Ridge Intersection with the $\lambda_{2,t+2} = 0$ Boundary of K_0] Let K_0 denote the cycle region from Proposition 4.1. The subset

$$D = \{x_m \in [0,1] \mid \mathcal{M}(x_m) \cap \{\mathbf{x} \in K_0 : \lambda_{2,t+2}(\mathbf{x}) = 0\} \neq \emptyset\}$$

collects all mature shares whose iso-mature line intersects the $\lambda_{2,t+2} = 0$ boundary of K_0.

Lemma 5 (Strict Concavity on Interior Ridge Slice)**.** *If $x_m \in D$, then $V(x_m, x_o)$ is strictly concave on the set $K_0 \cap \mathcal{M}(x_m)$, and the ridge correspondence $R(x_m)$ is a singleton.*

Proof. From Proposition 4.1, the value function on K_0 satisfies:

$$V(\mathbf{x}_t)|_{\mathbf{x}_t \in K_0} = \frac{1}{1-\beta^3}\Big[u(c_t) + \beta u(c_{t+1}) + \beta^2 u(c_{t+2})\Big] \tag{17}$$

where $c_t = c(\mathbf{x}_t)$, $c_{t+1} = c(\mathcal{F}\mathbf{x}_t)$, and $c_{t+2} = c(\mathcal{F}^2\mathbf{x}_t)$.

For $\mathbf{x} = (x_m, x_o) \in K_0 \cap \mathcal{M}(x_m)$, the first derivative of V with respect to x_o is

$$\frac{\partial V}{\partial x_o} = \frac{1}{1-\beta^3}\Big[u'(c_t)(f_o - f_y) + \beta u'(c_{t+1})(f_y - f_m) + \beta^2 u'(c_{t+2})(f_m - f_o)\Big].$$

From equation (7b) and the KKT conditions, we have:

$$\beta \frac{\partial V}{\partial x_o} = \lambda_{2,t+2}(x_o) \quad \Rightarrow \quad \frac{\partial V}{\partial x_o} = \frac{\lambda_{2,t+2}(x_o)}{\beta} \geq 0.$$

Equality holds if and only if $\mathbf{x}$ lies on the $\lambda_{2,t+2} = 0$ boundary of K_0.

The second derivative is:

$$\frac{\partial^2 V}{\partial x_o^2} = \frac{1}{1-\beta^3}\left[u''(c_t)(f_o - f_y)^2 + \beta u''(c_{t+1})(f_y - f_m)^2 + \beta^2 u''(c_{t+2})(f_m - f_o)^2\right] < 0,$$

using $u'' < 0$ and the strict inequalities $f_m > f_y, f_m > f_o$ from the Faustmann $i^* = 3$ assumption.

Therefore, $V(x_m, x_o)$ is strictly concave on the compact domain $K_0 \cap \mathcal{M}(x_m)$ and attains a unique maximum. Hence $R(x_m)$ is a singleton in K_0. □

Lemma 6 (No Ridge Points Outside K_0)**.** *Let $x_m \in D$. Then $\mathbf{R}(x_m) \subseteq K_0$. That is, no point in $\mathbf{R}(x_m)$ lies outside K_0.*

Proof. By definition of D, there exists an element of $R(x_m)$ that lies in K_0. By Lemma 5, $R(x_m)$ is singleton within K_0, and since $\frac{\partial V}{\partial x_o} \geq 0$ on $K_0 \cap \mathcal{M}(x_m)$ with equality at the $\lambda_{2,t+2} = 0$ boundary, this K_0 element must be the minimum selection of $R(x_m)$. Therefore, we can write it as $\underline{r}(x_m)$, where $(x_m, \underline{r}(x_m)) \in K_0$.

Observe that $\underline{r} \neq 0$, since K_0 also contains interior points of $\mathcal{M}(x_m)$ (Proposition 4.1) and $V(x_m, x_o)$ is strictly concave on $K_0 \cap \mathcal{M}(x_m)$ (Lemma 5), the boundary value at $x_o = 0$ is strictly dominated. Moreover, if $\underline{r} = 1 - x_m$, R must be singleton since $\underline{r}$ is the minimum of R. Hence non-singleton R are only possible for $0 < \underline{r} < 1 - x_m$. We now assume $0 < \underline{r}(x_m) < 1 - x_m$, and show that no $\delta > 0$ can exist such that $\underline{r}(x_m) + \delta \in R(x_m) \setminus K_0$.

Fix an interior $\underline{r}$ and assume towards a contradiction that there exists some $r(x_m) \in R(x_m)$ such that $(x_m, r(x_m)) \notin K_0$. Since $R(x_m)$ is compact-valued (Lemma 3), we can write $r(x_m) = \underline{r}(x_m) + \delta$

for some $\delta > 0$, so that

$$\mathbf{x}_t = (x_{mt}, \underline{r}(x_{mt}) + \delta) \in \mathbf{R}(x_{mt}) \setminus K_0.$$

Consider the following four exhaustive cases:

Case 1: First transition is to the ridge. Let $r(x_{m,t+1})$ be a selection from $R(x_{m,t+1})$. Then:

$$\mathbf{x}_{t+1} = (x_{m,t+1}, r(x_{m,t+1})), \quad x_{m,t+1} = 1 - x_{mt} - r(x_{mt}) - \delta.$$

Since R is upper hemicontinuous and K_0 is interior, $r(x_{m,t+1}) > 0$. Hence, the transition is interior.

By Benveniste-Scheinkman:

$$\left.\frac{\partial V}{\partial x_o}\right|_{\mathbf{x}_t} = u'(c_t)(f_o - f_y) = 0 \;\Rightarrow\; f_o = f_y,$$

contradicting the Faustmann $i^* = o$ conditions.

Case 2: First is Faustmann, second is ridge. Let $r(x_{m,t+2})$ be a selection from $R(x_{m,t+2})$. Then:

$$\mathbf{x}_{t+1} = \mathcal{F}(\mathbf{x}_t), \quad \mathbf{x}_{t+2} = (x_{m,t+2}, r(x_{m,t+2})), \quad x_{m,t+2} = r(x_{mt}) + \delta.$$

The transition is interior. Let:

$$h(x_o) := (f_o - f_y)u'(c_t(x_o)) + \beta(f_y - f_m)u'(c_{t+1}(x_o)) + \beta^2(f_m - f_o)u'(c_{t+2}(x_o))$$

h corresponds to $\frac{1-\beta^3}{\beta}\lambda_{2,t+2}(x_o)$. Then:

$$0 = u'(c_t)(f_o - f_y) + \beta u'(c_{t+1})(f_y - f_m) + \beta^2 \left.\frac{\partial V}{\partial x_m}\right|_{(x_{m,t+2}, r(x_{m,t+2}))}.$$

Rearranging:

$$\frac{\partial V}{\partial x_m}\Big|_{(x_{m,t+2},r(x_{m,t+2}))} = \frac{1}{\beta^2(f_m - f_o)}\left[h(r(x_{mt}) + \delta) - (u'(c_t)(f_o - f_y) + \beta u'(c_{t+1})(f_y - f_m))\right] > 0,$$

a contradiction.

Case 3: Two Faustmann transitions, then ridge.

$$\mathbf{x}_{t+2} = \mathcal{F}^2(\mathbf{x}_t) = (x_{mt}, r(x_{mt}) + \delta, 1 - x_{mt} - r(x_{mt}) - \delta).$$

Then:

$$0 = u'(c_t)(f_o - f_y) + \beta u'(c_{t+1})(f_y - f_m) + \beta^2 u'(c_{t+2})(f_m - f_o) = h(r(x_{mt}) + \delta).$$

But h is strictly decreasing and $h(r(x_{mt})) = 0$, so this contradicts strict monotonicity.

Case 4: Three Faustmann transitions. $\mathbf{x}_{t+3} = \mathcal{F}^3(\mathbf{x}_{mt}) = \mathbf{x}_t$ implies $\mathbf{x}_t \in K_0$, contradicting the assumption.

Hence, no such $\delta > 0$ exists. □

[Uniqueness of Ridge Correspondence on the $\lambda_{2,t+2}$ Boundary of K_0] Let $x_m \in D$. Then the ridge correspondence $R(x_m)$ is a singleton, and the unique ridge allocation lies in K_0.

Proof. By definition of D, we have $\mathcal{M}(x_m) \cap K_0 \neq \emptyset$. Lemma 5 establishes that V is strictly concave on the compact set $K_0 \cap \mathcal{M}(x_m)$, ensuring a unique maximum exists. By Lemma 6, no point in $\mathbf{R}(x_m)$ lies outside K_0, so the unique ridge allocation must lie within K_0. Therefore, $R(x_m)$ is a singleton for all $x_m \in D$. □

Definition 5. Largest allocation of mature-aged trees intersecting the $\lambda_{2,t+2}$ boundary

Let $\bar{x}_m := \sup D$, where D is the set of mature shares whose iso-mature lines intersect the $\lambda_{2,t+2} = 0$ boundary of K_0 (as defined above).

Since K_0 is closed and bounded (hence compact) and $\lambda_{2,t+2}$ is continuous, the set $\{\mathbf{x} \in K_0 : \lambda_{2,t+2}(\mathbf{x}) = 0\}$ is compact. The projection of a compact set is compact, so D is a compact subset of $[0, 1]$. Therefore, the supremum is attained: $\bar{x}_m = \max D$. △

Lemma 7 (Location of $\mathcal{F}(\bar{x}_m, \bar{r})$ on the $\lambda_{2,t+2}$ Boundary)**.** *Let $\bar{x}_m$ denote the maximum value of x_m such that $\mathcal{M}(x_m) \cap K_0 \neq \emptyset$, and let $\bar{r}$ be the corresponding ridge value. Then:*

1. *The allocation $(1 - \bar{x}_m - \bar{r}, \bar{x}_m)$ lies on the ridge and on the $\lambda_{2,t+2} = 0$ boundary of K_0.*
2. *The value $1 - \bar{x}_m - \bar{r}$ is the minimum element of D.*

Proof. Part 1: By definition of $\bar{x}_m$, we have $(\bar{x}_m, \bar{r}) \in K_0$. Since K_0 is the set of allocations generating optimal 3-cycles, the aging map transition from $(\bar{x}_m, \bar{r})$ must also lie in K_0. The aging map $\mathcal{F}$ applied to $(\bar{x}_m, \bar{r})$ gives $\mathcal{F}(\bar{x}_m, \bar{r}) = (1 - \bar{x}_m - \bar{r}, \bar{x}_m)$. Therefore, $(1 - \bar{x}_m - \bar{r}, \bar{x}_m) \in K_0$.

We show that $\bar{x}_m$ is the largest x_o value on $\mathcal{M}(1 - \bar{x}_m - \bar{r})$ within K_0. Suppose there exists $\delta > 0$ such that $(1-\bar{x}_m-\bar{r}, \bar{x}_m+\delta) \in K_0$. Then $\mathcal{F}^2$ applied to this allocation yields $\mathcal{F}^2(1-\bar{x}_m-\bar{r}, \bar{x}_m+\delta) = (\bar{x}_m + \delta, 1 - \bar{x}_m - \bar{r})$. This implies $(\bar{x}_m + \delta, 1 - \bar{x}_m - \bar{r}) \in K_0$, so $\mathcal{M}(\bar{x}_m + \delta) \cap K_0 \neq \emptyset$, contradicting the maximality of $\bar{x}_m$.

By Corollary 4.3.3, since $(1-\bar{x}_m-\bar{r}, \bar{x}_m)$ is the value-maximizing allocation on $\mathcal{M}(1-\bar{x}_m-\bar{r}) \cap K_0$, we have $(1 - \bar{x}_m - \bar{r}, \bar{x}_m) \in \mathbf{R}(1 - \bar{x}_m - \bar{r})$, which gives $R(1 - \bar{x}_m - \bar{r}) = \bar{x}_m$. Since $\bar{x}_m$ represents the maximum x_o value along its iso-mature line within K_0 and this maximum is achieved at a ridge point, the allocation $(1 - \bar{x}_m - \bar{r}, \bar{x}_m)$ must lie on the boundary of K_0 where $\lambda_{2,t+2} = 0$.

Part 2: We prove $1 - \bar{x}_m - \bar{r} = \min D$ by contradiction. Suppose there exists $x'_m < 1 - \bar{x}_m - \bar{r}$ such that $x'_m \in D$. Then there exists (x'_m, x'_o) on the $\lambda_{2,t+2} = 0$ boundary of K_0. Since $(x'_m, x'_o) \in K_0$, the complete 3-cycle must lie in K_0:

$$(x'_m, x'_o) \in K_0$$

$$\mathcal{F}(x'_m, x'_o) = (1 - x'_m - x'_o, x'_m) \in K_0$$

$$\mathcal{F}^2(x'_m, x'_o) = (x'_o, 1 - x'_m - x'_o) \in K_0$$

From the third element of the 3-cycle, $\mathcal{F}^2$, we have $\mathcal{M}(x'_o) \cap K_0 \neq \emptyset$, which by definition of $\bar{x}_m$ requires $x'_o \leq \bar{x}_m$.

Since $x'_m < 1-\bar{x}_m-\bar{r}$, we have $1-x'_m > \bar{x}_m+\bar{r}$. Therefore, $1-x'_m-x'_o > \bar{x}_m+\bar{r}-x'_o \geq \bar{x}_m+\bar{r}-\bar{x}_m = \bar{r}$

From the second element of the 3-cycle, we have $(1-x'_m-x'_o, x'_m) \in K_0$ with $1-x'_m-x'_o > \bar{r} \geq \bar{x}_m$, where the last inequality follows from the fact that $(\bar{x}_m, \bar{r}) \in K_0$ lies on the ridge. This means $\mathcal{M}(1-x'_m-x'_o) \cap K_0 \neq \emptyset$ with $1-x'_m-x'_o > \bar{x}_m$, contradicting the definition of $\bar{x}_m$ as the maximum value of x_m such that $\mathcal{M}(x_m) \cap K_0 \neq \emptyset$.

Therefore, no such $x'_m < 1 - \bar{x}_m - \bar{r}$ can exist in D, establishing that $1 - \bar{x}_m - \bar{r} = \min D$. □

Proposition 4.3 (Complete Characterization of the K_0 Boundary)**.** *Let* $\mathcal{R}_3 = \{(x_m, R(x_m)) : x_m \in [1 - \bar{x}_m - \bar{r}, \bar{x}_m]\}$ *denote the subset of the ridge graph that lies in* K_0*. The boundary of the cycle region* K_0 *consists of exactly three curves:*

$$\mathrm{bd}(K_0) = \mathcal{R}_3 \cup \mathcal{F}(\mathcal{R}_3) \cup \mathcal{F}^2(\mathcal{R}_3)$$

where:

$$\mathcal{F}(\mathcal{R}_3) = \{(1 - x_m - R(x_m), x_m) : x_m \in [1 - \bar{x}_m - \bar{r}, \bar{x}_m]\}$$
$$\mathcal{F}^2(\mathcal{R}_3) = \{(R(x_m), 1 - x_m - R(x_m)) : x_m \in [1 - \bar{x}_m - \bar{r}, \bar{x}_m]\}$$

Proof. From Corollary 4.3.3 and Lemma 7, the $\lambda_{2,t+2} = 0$ boundary of K_0 is precisely $\mathcal{R}_3$.

Since K_0 consists of allocations generating optimal 3-cycles, it is invariant under the aging map $\mathcal{F}$: if $\mathbf{x} \in K_0$, then $\{\mathbf{x}, \mathcal{F}\mathbf{x}, \mathcal{F}^2\mathbf{x}\} \subset K_0$. This 3-fold rotational symmetry implies that the complete boundary is:

$$\mathrm{bd}(K_0) = \mathcal{R}_3 \cup \mathcal{F}(\mathcal{R}_3) \cup \mathcal{F}^2(\mathcal{R}_3)$$

Since $\mathcal{F}(x_m, x_o) = (1 - x_m - x_o, x_m)$ and $\mathcal{F}^2(x_m, x_o) = (x_o, 1 - x_m - x_o)$, we obtain the explicit

characterization above. The three curves together form the complete boundary $\text{bd}(K_0)$. □

4.3.4 The feasible transition sets

We partition the state space into sets $K_0, K_1, K_2, \ldots$ where K_i contains allocations such that $\Gamma^r(\mathbf{x}) \cap K_{i-1} \neq \emptyset$ but $\mathbf{x} \notin \bigcup_{j=0}^{i-1} K_j$. This hierarchical partition transforms the convergence problem into showing that optimal trajectories move systematically through these sets toward K_0.

Definition 6. Feasible Transition Sets

Call the cycle region, defined by (7a) ≥ 0 and (7b) ≥ 0, K_0. The set K_i is the set of allocations that can feasibly transition to K_{i-1} but are not already elements of any K_j, $(j < i)$. That is $K_i = \{\mathbf{x} \in \Delta^2 \setminus \bigcup_{j=0}^{i-1} K_j : \Gamma^r(\mathbf{x}) \cap K_{i-1} \neq \emptyset\}$. Figure 8 shows the feasible transition sets for four examples with qualitatively different intersections between the ridge and the sets. △

Although definition 6 defines the K_i sets, we will need an explicit enumeration of these sets in the subsequent analysis. The following sections define the allocations that satisfy the definition above.

Definition 7. Explicit definition of K_1 The set K_1 consists of allocations that can feasibly transition to K_0 under the restricted transition rule but are not already elements of K_0.

From the restricted transition rule, allocation $\mathbf{x} = (x_y, x_m, x_o)$ can transition to any point in

$$\Gamma^r(\mathbf{x}) = \{(1 - x_y - x'_o, x_y, x'_o) : 0 \leq x'_o \leq x_m\}$$

For $\Gamma^r(\mathbf{x}) \cap K_0 \neq \emptyset$, the iso-mature line $\mathcal{M}(x_y)$ must intersect K_0, requiring $x_y \in [1 - \bar{x}_m - \bar{r}, \bar{x}_m]$. Since $K_0 \cap \mathcal{M}(x_y)$ is maximized at $x'_o = R(x_y)$ by Lemma **??**, the transition set reaches K_0 when $R(x_y) \leq x_m$. The strict inequality $R(x_y) < x_m$ ensures $\mathbf{x} \notin K_0$.

For the given x_y values, no other x_m values are admissible: if $x_m \leq R(x_y)$, then $\Gamma^r(\mathbf{x})$ cannot reach K_0, so $\mathbf{x} \notin K_1$. The upper constraint comes from requiring $\mathbf{x} \notin K_0$, which excludes certain x_m values where $(x_y, x_m, 1 - x_y - x_m) \in K_0$.

Therefore:

$$K_1 = \Big\{(x_y, x_m, x_o) \in \Delta^2 \;\Big|\; x_y \in [1 - \bar{x}_m - \bar{r}, \bar{x}_m],$$
$$R(x_y) < x_m,$$
$$x_o = 1 - x_y - x_m\Big\}.$$

△

Definition 8. Explicit definition of K_2

K_2 consists of two sets: K_2', the set of points that transition to K_1 via Faustmann transition ($x_{o,t+1} = x_{mt}$), and K_2'', the set of points that transition to K_1 via Non-Faustmann transition ($x_{o,t+1} < x_{mt}$). These sets are defined as:

$$K_2' = \Big\{\mathbf{x} \in \Delta^2 \;:\; \mathcal{F}\mathbf{x} \in K_1\Big\} = \Big\{(x_y, x_m, x_o) \in \Delta^2 \;: x_o \in [1 - \bar{x}_m - \bar{r}, \bar{x}_m],$$
$$x_m < 1 - x_o - R(x_o),$$
$$x_y = 1 - x_m - x_o\Big\}$$

$$K_2'' = \Big\{(x_y, x_m, x_o) \in \Delta^2 \;\Big|\; x_y \in [\bar{x}_m, \bar{x}_m + \bar{r}],$$
$$x_o \leq 1 - \bar{x}_m - \bar{r},$$
$$x_m = 1 - x_y - x_o\Big\}$$

Therefore, $K_2 = K_2' \cup K_2''$. △

Definition 9. Explicit definition of K_3

Like K_2, K_3 is composed of consists of two sets, K_3' the set of points that transition to K_2 via Faustmann transition, and K_3'', the set of points that transition to K_2 via Non-Faustmann transition.

These sets are defined as:

$$
\begin{aligned}
K_3' &= \Big\{ \mathbf{x} \in \Delta^2 \ : \ \mathcal{F}\mathbf{x} \in K_2 \Big\} \\
&= \Big\{ (x_y, x_m, x_o) \in \Delta^2 \ : \ x_m \in [1 - \bar{x}_m - \bar{r}, \bar{x}_m], \\
&\qquad\qquad x_y < 1 - x_m - R(x_m), \\
&\qquad\qquad x_o = 1 - x_y - x_m \Big\} \\
&\cup \Big\{ (x_y, x_m, x_o) \in \Delta^2 \ : \ x_o \in [\bar{x}_m, \bar{x}_m + \bar{r}], \\
&\qquad\qquad x_m \leq 1 - \bar{x}_m - \bar{r}, \\
&\qquad\qquad x_y = 1 - x_m - x_o \Big\}
\end{aligned}
$$

and

$$
\begin{aligned}
K_3'' = \Big\{ (x_y, x_m, x_o) \in \Delta^2 \ \Big| \ & x_y \in [0, 1 - \bar{x}_m - \bar{r}], \\
& x_m > \bar{x}_m, \\
& x_o = 1 - x_y - x_m \Big\}
\end{aligned}
$$

Therefore, $K_3 = K_3' \cup K_3''$. △

Definition 10. Explicit definition of K_4 and K_5

There are two remaning regions on the simplex: $\{\mathbf{x} \in \Delta^2 : x_y > \bar{x}_m + \bar{r}\}$ and $\{\mathbf{x} \in \Delta^2 : x_o > \bar{x}_m + \bar{r}\}$. For all $\mathbf{x}$ in the first set, the restricted transition sets will have $x_m > \bar{x}_m + \bar{r}$ meaning the restricted transistion sets will be wholly contained in K_3. Hence this set must be a subset of K_4. For all $\mathbf{x}$ in the second set, the restricted transition sets cannot intersect any allocations with $x_y \leq \bar{x}_m + \bar{r}$, meaning these restricted transtition sets are wholly contained within the first set. Hence the second set must be a subset of K_5. By definition, there can be no common elements of K_i and K_j, so we have $K_4 = \{\mathbf{x} \in \Delta^2 : x_y > \bar{x}_m + \bar{r}\}$ and $K_5 = \{\mathbf{x} \in \Delta^2 : x_o > \bar{x}_m + \bar{r}\}$. △

4.3.5 Restricting focus to the ridge

We prove that any allocation $\mathbf{x}_t \notin K_0$ must optimally transition to the ridge $\mathcal{R}$ within at most three periods. This result reduces the convergence analysis to characterizing how allocations on each of the four ridge segments $\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3, \mathcal{R}_4$ transition toward K_0. The proof shows that avoiding the ridge for three consecutive periods would require Faustmann transitions satisfying $\mathcal{F}^3\mathbf{x}_t = \mathbf{x}_t$, contradicting $\mathbf{x}_t \notin K_0$.

Lemma 8. *Any $\mathbf{x}_t \notin K_0$ must optimally transition to $\mathcal{R}$ in at most three periods, i.e., if $\mathbf{x}^*_{t+1} \notin \mathcal{R}$ and $\mathbf{x}^*_{t+2} \notin \mathcal{R}$ then $\mathbf{x}^*_{t+3} \in \mathcal{R}$.*

Proof. Assume there exists $\mathbf{x}_t \notin K_0$, such that for $i{\in}\{1,2,3\}$, any selection $\xi(\mathbf{x}_{t+i}) \in \boldsymbol{\xi}(\mathbf{x}_{t+i}), \xi(\mathbf{x}_{t+i}) \notin \mathbf{R}(x_{m,t+i-1})$. This implies that for $i{\in}\{1,2,3\}$, $\mathbf{R}(x_{m,t+i}) \notin \Gamma^r(\mathbf{x}_{t+i})$ and hence $\mathbf{x}_{t+i} = \mathcal{F}^i\mathbf{x}_t$. But in this case $\mathcal{F}^3\mathbf{x}_t{=}\mathbf{x}_t$, i.e., $\mathbf{x}_t$ is part of a three-cycle, contradicting the assumption that $\mathbf{x}_0 \notin K_0$. □

Definition 11. Let $\tilde{\mathbf{x}} = (\tilde{x}, R(\tilde{x}))$ where $\tilde{x}_m = \min\{x_m : R(x_m) = \{0\}\}$, i.e. the lowest x_m for which the largest value of the ridge correspondence equals the $x_o = 0$ boundary of the simplex. △

From Lemma 8, the analysis reduces to focusing on the behavior of allocations contained within $\mathcal{R}$. To facilitate the following analysis, the ridge is partitioned into four subsets.

Definition 12. Ridge subsets

Let the set of allocations on the ridge, formally, $\{\mathbf{x} \in \Delta^2 : \mathbf{x} = (x_m, R(x_m)),\ x_m{\in}[0,1]\}$ be partitioned into the following, mutually exclusive subsets:

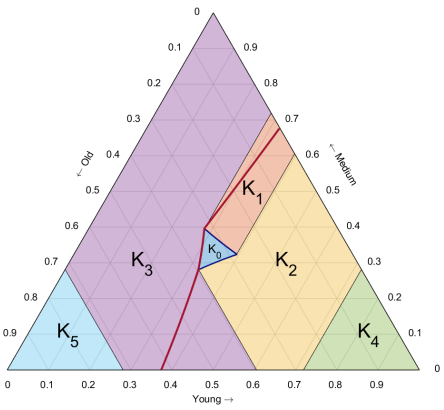
$$\mathcal{R}_1 := \{\mathbf{x} \in \mathcal{R} : x_m{\in}(\tilde{x}_m, 1]\} \tag{18a}$$

$$\mathcal{R}_2 := \{\mathbf{x} \in \mathcal{R} : x_m{\in}(\bar{x}_m, \tilde{x}_m]\} \tag{18b}$$

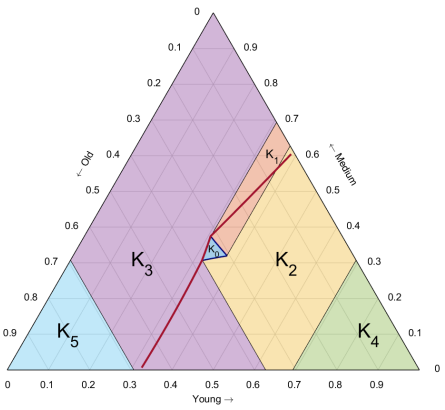
$$\mathcal{R}_3 := \{\mathbf{x} \in \mathcal{R} : x_m{\in}[R(\bar{x}_m), \bar{x}_m]\} \tag{18c}$$

$$\mathcal{R}_4 := \{\mathbf{x} \in \mathcal{R} : x_m{\in}[0, R(\bar{x}_m))\} \tag{18d}$$

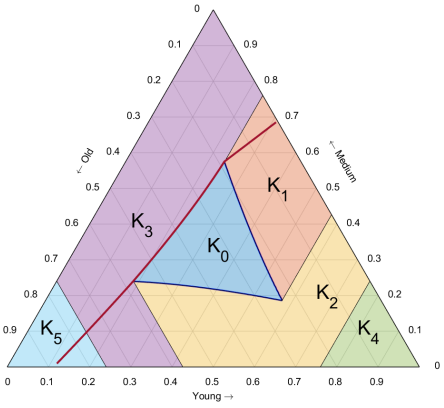
△



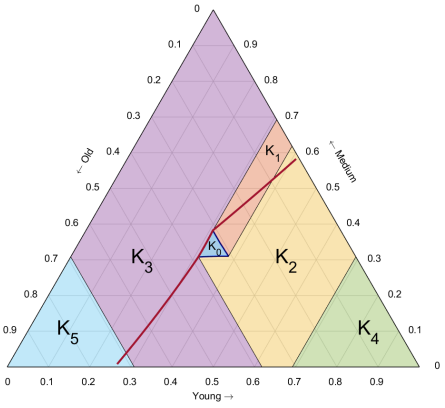
(a) Parameters generating $N_1 = 0, N_2 = 0$ ($u(\mathbf{c}) = \ln(\mathbf{c}), \mathbf{f} = (0.23, 1, 0.63), \beta = 0.75$).



(b) Parameters generating $N_1 > 0, N_2 = 0$ ($u(\mathbf{c}) = \ln(\mathbf{c}), \mathbf{f} = (0.38, 1, 0.60), \beta = 0.53$).



(c) Parameters generating $N_1 = 0, N_2 > 0$ ($u(\mathbf{c}) = \ln(\mathbf{c}), \mathbf{f} = (0.30, 1, 0.52), \beta = 0.21$).



(d) Parameters generating $N_1 > 0, N_2 > 0$ ($u(\mathbf{c}) = \ln(\mathbf{c}), \mathbf{f} = (0.27, 1, 0.48), \beta = 0.37$).

Figure 8: Example feasible transition-set plots for the four combinations of N_1 and N_2. (a) $N_1 = 0, N_2 = 0$; (b) $N_1 > 0, N_2 = 0$; (c) $N_1 = 0, N_2 > 0$; (d) $N_1 > 0, N_2 > 0$.

4.3.6 Transitions from $\mathcal{R}_2$

We analyze $\mathcal{R}_2 = \{\mathbf{x} \in \mathcal{R} : x_m \in (\bar{x}_m, \tilde{x}_m]\}$ and establish that optimal transitions are declining in mature tree allocations, with each step reducing the mature component by at least $\delta_1 > 0$. Using properties of $\partial V/\partial x_o$ along the boundary $\text{bd}(K_{12})$, which is established to be differentiable, we prove that allocations within δ^c of $\bar{x}_m$ transition directly to K_0. This monotonic decrease ensures convergence to K_0 within $N_1 = \lceil \frac{1-\bar{x}_m}{\delta_1} \rceil$ periods.

Throughout this section, we sometimes write $u(c(\mathbf{x}_t))$ as $u(\mathbf{x}_t)$ for conciseness.

Lemma 9. *If* $\mathbf{x}_t \in K_1$*, then* $\mathbf{x}^*_{t+1} \in K_0$.

Proof. As established in the proof of Lemma **??**, $\beta \left.\frac{dV(\mathbf{x}_t)}{dx_o}\right|_{\mathbf{x}_t \in K_0} = \lambda_{2,t+2}$, and the ridge corresponds to the $\lambda_{2,t+2} = 0$ boundary of K_0. By Definition 12, this boundary portion is $\mathcal{R}_3$.

For any $\mathbf{x}_t \in K_1$, Definition 7 ensures that $\Gamma^r(\mathbf{x}_t) \cap K_0 \neq \emptyset$, specifically that the iso-mature line $\mathcal{M}(x_{yt})$ intersects K_0 with $R(x_{yt}) \leq x_{mt}$. Since the ridge point $(x_{yt}, R(x_{yt}))$ maximizes value on $\mathcal{M}(x_{yt})$ by definition of the ridge, and this point lies within the feasible transition set $\Gamma^r(\mathbf{x}_t)$, the optimal transition function (Definition 4) yields $\boldsymbol{\xi}(\mathbf{x}_t) \in \mathcal{R}_3 \subset K_0$. □

Using the definitions of the feasible transition sets (definitions 7 and 9), denote the boundary between K_1 and K_3 as $\text{bd}(K_{13}) = \{(x_m, (\bar{x}_m + \bar{r} - x_m)) : x_m \in (\bar{x}_m, \bar{x}_m + \bar{r}]\}$ (see for example, figure 8). Note that this boundary coincides with the iso-young line defined by $x_y = 1 - \bar{x}_m - \bar{r}$.

Lemma 10. *For allocations in* $\text{bd}(K_{13})$*,* $\frac{\partial V}{\partial x_o} < 0$

Proof. Consider two allocations in $\text{bd}(K_{13})$, $\check{\mathbf{x}} = (\check{x}_m, \bar{x}_m + \bar{r} - \check{x}_m)$ and $\hat{\mathbf{x}} = (\hat{x}_m, \bar{x}_m + \bar{r} - \hat{x}_m)$ where $\hat{x}_m > \check{x}_m$. Therefore $u'(\hat{\mathbf{x}}) < u'(\check{\mathbf{x}})$ since $f_m > f_o$. Because these allocations are in K_1, lemma 9 applies and the optimal transition is to $\mathcal{R}_3$, which is in the interior of $\Gamma^r(\mathbf{x})$. Hence, the Benveniste-Scheinkman theorem applies and $V(\mathbf{x})$ is C^1 along $\text{bd}(K_{13})$. The derivatives of the

value function along the corresponding iso-mature lines at these two allocations are

$$\begin{aligned}\left.\frac{dV(\mathbf{x}_t)}{dx_o}\right|_{\check{\mathbf{x}}} &= u'(\check{\mathbf{x}})(f_o - f_y) + \beta\frac{dV(\boldsymbol{\xi}(\check{\mathbf{x}}))}{dx_o} \\ \left.\frac{dV(\mathbf{x}_t)}{dx_o}\right|_{\hat{\mathbf{x}}} &= u'(\hat{\mathbf{x}})(f_o - f_y) + \beta\frac{dV(\boldsymbol{\xi}(\hat{\mathbf{x}}))}{dx_o}\end{aligned} \tag{19}$$

Because $\check{\mathbf{x}}$ and $\hat{\mathbf{x}}$ are on the same iso-young line we have $\boldsymbol{\xi}(\check{\mathbf{x}}) = \boldsymbol{\xi}(\hat{\mathbf{x}})$. Therefore

$$\left.\frac{dV(\mathbf{x}_t)}{dx_o}\right|_{\hat{\mathbf{x}}} - \left.\frac{dV(\mathbf{x}_t)}{dx_o}\right|_{\check{\mathbf{x}}} = \underbrace{(u'(\hat{\mathbf{x}}) - u'(\check{\mathbf{x}}))}_{<0}\underbrace{(f_o - f_y)}_{>0}$$

Since $\hat{\mathbf{x}}$ and $\check{\mathbf{x}}$ are identical except for x_m, we can conclude that the derivative term in (19) decreases with x_m. Now take a $x_m \in (\bar{x}_m, \bar{x}_m + \bar{r}]$ and let it approach $\bar{x}_m$ from above. Since $(\bar{x}_m, \bar{r})$ is on the $\lambda_{2,t+2} = 0$ boundary of K_0 it follows from Lemma 9 that

$$\lim_{x_m \searrow \bar{x}_m} \frac{dV(x_m, \bar{x}_m - \bar{r} - x_m))}{dx_o} = \frac{dV(\bar{x}_m, \bar{r})}{dx_o} = 0$$

Therefore, for all $\mathbf{x} \in \mathrm{bd}(K_{13})$, $\frac{dV(\mathbf{x})}{dx_o} < 0$. □

Corollary 11. $\mathcal{R}_2 \subset K_1 \cup K_2$

Proof. From the explicit definitions of the K_i sets (definitions 7–10), $\mathcal{R}_2 \subset K_1 \cup K_2 \cup K_3$. Lemma 10 implies that for all $\mathbf{x} \in \mathcal{R}_2$, $1 - x_m - x_o > 1 - \bar{x}_m - \bar{r}$, so $\mathcal{R}_2 \cap K_3 = \emptyset$. □

Lemma 12. *There exists a constant $\delta^c > 0$, such that if $\mathbf{x} \in \mathcal{R}_2$ and $(x_m - \bar{x}_m) < \delta^c$, then $\mathbf{x} \in K_1$.*

Proof. Let $\mathrm{bd}(K_{12}) = \{\mathbf{x} \in \Delta^2 : x_y = \bar{x}_m,\ x_m \in (\bar{r}, 1 - \bar{x}_m]\}$ denote the common boundary of K_1 and K_2.

Case 1: $\mathcal{R}_2 \subset K_1$. Because $\bar{x}_m > \bar{r}$ (Proposition 4.1), every point of $\mathcal{R}_2$ lies at most $1 - 2\bar{x}_m$ units to the right of $\bar{x}_m$ in the mature coordinate, so the lemma holds with $\delta^c = 1 - 2\bar{x}_m$.

Case 2: $\mathcal{R}_2 \cap K_2 \neq \emptyset$. Define

$$A \;=\; \{x_m - \bar{x}_m > 0 : \exists\, x_o \text{ with } (x_m, x_o) \in \mathcal{R}_2 \cap \mathrm{bd}(K_{12})\}.$$

The set A is non-empty (because $\mathcal{R}_2 \cap K_2 \neq \emptyset$) and is bounded below by 0. Since the ridge graph is closed (definition 3), A is closed, so $\delta^c := \inf A > 0$ and

$$(x_m - \bar{x}_m) < \delta^c \implies \mathbf{x} \in \mathcal{R}_2 \subset K_1.$$

□

Lemma 13. *There exists $\delta_1 \in (0, \delta^c)$ (where δ^c was identified in Lemma 12), such that for every $\mathbf{x}_t \in \mathcal{R}_2$, the optimal transition satisfies $\xi_m(\mathbf{x}_t) < x_{m,t} - \delta_1$.*

Proof. First consider $\bar{\mathbf{x}}^c \in K_0$. Since K_0 is not a singleton set (by proposition 4.1), and $\bar{\mathbf{x}}^c$ is the allocation in K_0 with the largest mature element, there exists $\delta_0 > 0$ such that $\xi_m(\bar{\mathbf{x}}^c) < \bar{x}_m - 2\delta_0$. By continuity, there exists $\epsilon > 0$ such that for $\mathbf{x}_t \in \mathcal{R}_2$, if $x_{mt} - \bar{x}_m \leq \epsilon$, then $\xi_m(\mathbf{x}_t) < x_{mt} - \delta_0$. Now let $\hat{\mathcal{R}}_2 = \{\mathbf{x} \in \mathcal{R}_2 : x_t - \bar{x}_m \geq \epsilon\}$. If $\hat{\mathcal{R}}_2$ is empty, then the result is proved, by setting $\delta_1 = \delta_0$. Assume therefore that $\hat{\mathcal{R}}_2 \neq \emptyset$; we will show that

$$\text{for all } \mathbf{x}_t \quad \in \quad \mathcal{R}_2, \quad \xi_m(\mathbf{x}_t) \quad < \quad x_{mt}. \tag{20}$$

If, to the contrary, there exists $\mathbf{x}_t \in \mathcal{R}_2$ such that $\xi_m(\mathbf{x}_t) \geq x_{mt}$, then by continuity and the intermediate value theorem, there must exist $\mathbf{x}'_t \in \mathcal{R}_2$ such that $\xi_m(\mathbf{x}'_t) = x'_{mt}$. That is, by the definition of $\xi(\cdot)$, and noting that $\mathbf{x}_t$ is transitioning from ridge point to ridge point,

$$\xi(\mathbf{x}'_t) \quad = \quad \left(x'_{tm}, R\left(x'_{mt}\right)\right) \quad = \quad \left(\xi_m(\mathbf{x}'_t), R\left(\xi_m(\mathbf{x}'_t)\right)\right) \quad = \quad \xi(\xi(\mathbf{x}'_t)) \tag{21}$$

for some $x_o \in R(x'_{mt})$.

However, (21) would imply that $\mathbf{x}'_t$ is an optimal steady state of the three-age-class tree replacement

problem. An optimal steady state is also, degenerately, a three-period cycle. This is a contradiction since $\mathbf{x}'_t$ is not an element of K_0 (since by definition 18b of $\mathcal{R}_2$, $x'_{mt} > \bar{x}_m$) and K_0 is defined as the set of all optimal three period cycles. Having proved (20), it follows from the compactness of $\hat{\mathcal{R}}_2$ that there exists $\delta > 0$ such that for all $\mathbf{x}\in\hat{\mathcal{R}}_2$, $\xi_m(\mathbf{x}_t)\leq x_{mt}-\delta$. The Lemma now holds for $\delta_1 = \min(\delta, \delta_0)$. □

Corollary 14. *Beginning from an allocation $\mathbf{x}\in\mathcal{R}_2$, the optimal trajectory will enter K_0 in at most $N_1 = \left\lceil \frac{(1-\bar{x}_m)}{\delta_1} \right\rceil$ periods ($\delta_1>0$ was identified in Lemma 13).*

Proof. For each $\mathbf{x}\in\mathcal{R}_2$, let $q(\mathbf{x}) = \left\lceil \frac{x_m-\bar{x}_m}{\delta_1} \right\rceil$ (that is, x_m is no greater than $q(\mathbf{x})$ units of length δ_1 larger than $\bar{x}_m$). Since $\bar{\mathbf{x}}^c \notin \mathcal{R}_2$ and $x_m > \bar{x}_m$ (see Definition 18b) $\left\lceil \frac{x_m-\bar{x}_m}{\delta_1} \right\rceil \leq \left\lceil \frac{1-\bar{x}_m}{\delta_1} \right\rceil = N_1$, it follows that $q(\mathbf{x})\leq N_1$. We now prove the lemma by induction. First, fix $\mathbf{x}\in\mathcal{R}_2$ such that $q(\mathbf{x})=1$. Since $\delta_1<\delta^c$, Lemma 12 implies that $\mathbf{x}\in K_1$, so that from Lemma 9, $\mathbf{x}$ enters K_0 in $1\leq q(\mathbf{x})$ periods. Now fix $1<n\leq N_1$, and assume that for all $\mathbf{x}\in\mathcal{R}_2$ such that $q(\mathbf{x})<n$, $\mathbf{x}$ enters K_0 in at most $q(\mathbf{x})$ periods. (This assumption has been established to be true for $n = 2$.) Suppose there exists $\mathbf{x}_t\in\mathcal{R}_2$ such that $q(\mathbf{x}_t) = n$. From Lemma 13, $\xi_m(\mathbf{x}_t)<x_{m,t}-\delta_1$, implying that $q(\xi t + 1(\mathbf{x}_t)) \leq n-1$ and hence $\xi t+1(\mathbf{x}_t)$ enters K_0 in at most $q(\mathbf{x}_t)-1$ periods. Therefore, $\mathbf{x}$ enters K_0 in at most $q(\mathbf{x}_t)\leq N_1$ periods, completing the proof. □

4.3.7 Transitions from $\mathcal{R}_1$

We examine $\mathcal{R}_1 = \{\mathbf{x} \in \mathcal{R} : x_m \in (\tilde{x}_m, 1]\}$ and establish that $R(x_m) = 0$ for $x_m > \bar{x}_m + \bar{r}$ using monotonicity of the subdifferential $\partial_{x_o} V(x_m, x_o)$. From any point in $\mathcal{R}_1$, optimal transitions are necessarily to $\mathcal{R}_4$.

Lemma 15 (Monotonicity of the Subdifferential $\partial_{x_o} V(x_m, x_o)$ in x_m)**.** *Consider the restricted value function $V : \Delta^2 \to \mathbb{R}$. Fix $x_o \in [0, 1]$, and let $x_m, x'_m \in [0, 1 - x_o]$ with $x_m < x'_m$. Then for any subgradients*

$$g_1 \in \partial_{x_o} V(x_m, x_o), \quad g_2 \in \partial_{x_o} V(x'_m, x_o),$$

we have

$$g_1 \geq g_2.$$

That is, the subdifferential of V with respect to x_o is monotonically decreasing in x_m, even when x_o lies on the boundary of its domain.

Proof. By Lemma 2, the restricted value function $V : \Delta^2 \to \mathbb{R}$ exists, is unique, continuous, and weakly concave on the state space Δ^2.

Step 1: Measurable selection and envelope theorem setup By the definitions of the Ridge (Definition 3) and the restricted optimal transition function (Definition 4), the optimal policy correspondence

$$\boldsymbol{\xi}(\mathbf{x}_t) := \arg \max_{\mathbf{x}_{t+1} \in \Gamma^r(\mathbf{x}_t)} \{u(c(\mathbf{x}_t)) + \beta V(\mathbf{x}_{t+1})\}$$

is nonempty, compact-valued, and upper hemicontinuous on the state space Δ^2.

Define the stage return function:

$$\phi(\mathbf{x}_t, \mathbf{x}_{t+1}) := u(c(\mathbf{x}_t)) + \beta V(\mathbf{x}_{t+1})$$

where $c(\mathbf{x}_t) = f_y x_{yt} + f_m x_{mt} + f_o x_{ot}$.

Since u is strictly concave and increasing, and the argument $f_y x_{yt} + f_m x_{mt} + f_o x_{ot}$ is linear in $\mathbf{x}_t$ and bounded on the compact domain Δ^2, the utility term is continuous and bounded. By corollary 14.6 of Rockafellar and Wets (1998), there exists a measurable selection $\xi(\mathbf{x}_t) \in \boldsymbol{\xi}(\mathbf{x}_t)$.

Step 2: Concavity of the stage return function

The stage return function $\phi(\mathbf{x}_t, \mathbf{x}_{t+1})$ is concave in $\mathbf{x}_t$ because:

- The utility term $u(c(\mathbf{x}_t))$ is the composition of a strictly concave function u with the linear function $c(\mathbf{x}_t)$, hence concave in $\mathbf{x}_t$.
- The continuation value $\beta V(\mathbf{x}_{t+1})$ is constant in $\mathbf{x}_t$ for fixed $\mathbf{x}_{t+1}$, hence concave.

Since the constraint correspondence Γ^r has a convex graph (the constraints are linear), the value function is concave as the supremum of concave functions over a convex constraint set.

Step 3: Envelope theorem application

Define

$$\tilde{V}(\mathbf{x}_t) := \phi(\mathbf{x}_t, \xi(\mathbf{x}_t)) = V(\mathbf{x}_t).$$

Since ϕ is concave in $\mathbf{x}_t$ and $\xi(\mathbf{x}_t)$ is a maximizer, the subdifferentiation in parametric minimization theorem of Rockafellar and Wets (1998, Theorem 10.13) implies:

$$\partial_{x_o} V(\mathbf{x}_t) \supseteq \partial_{x_o} \phi(\mathbf{x}_t, \xi(\mathbf{x}_t)).$$

Step 4: Subdifferential monotonicity

Consider the mapping $x_m \mapsto V(x_m, x_o)$ for fixed x_o. This is the composition of V with the affine transformation $x_m \mapsto (x_m, x_o)$, where the image always lies in Δ^2. Since V is weakly concave on Δ^2 and the transformation is affine, the composite function is weakly concave in x_m.

By the monotonicity of one-sided derivatives for convex functions[9] (Rockafellar, 1970, Theorem 24.1) and the resulting monotonicity of subgradients for the concave function V, and because subdifferentials are non-empty on the relative interior of the domain (Rockafellar, 1970, Theorem 23.4), with one-sided limits supplying the endpoints, we have that for any $x_m < x'_m \in [0, 1 - x_o]$ and subgradients

$$g_1 \in \partial_{x_o} V(x_m, x_o), \quad g_2 \in \partial_{x_o} V(x'_m, x_o),$$

it holds that $g_1 \geq g_2$.

This conclusion holds whether the subdifferential is a singleton or a closed interval. □

[9] Apply Theorem 24.1 to the convex function $-g$. The resulting non-decreasing property of the subgradients of $-g$ translates into a non-increasing property for the subgradients of g.

Recall from Definition 11 that $\tilde{\mathbf{x}} = (\tilde{x}, R(\tilde{x}))$ where $\tilde{x} = \min\{x_m : R(x_m) = 0\}$, i.e. the lowest x_m for which the largest value of the ridge correspondence equals the $x_o = 0$ boundary of the simplex, i.e. $\min\{x_o : \bar{R} = \{0\}\}$.

Lemma 16 ($R(x_m) = 0$ for $x_m > \bar{x}_m + \bar{r}$)**.** *For all $x_m > \bar{x}_m + \bar{r}$, we have $R(x_m) = 0$.*

Proof. By Lemma 10, for $(\bar{x}_m + \bar{r}, 0)$, $\frac{\partial V}{\partial x_o}(\bar{x}_m + \bar{r}, 0)$ exists and is strictly negative.

For any $x_m > \bar{x}_m + \bar{r}$, by Lemma 15, the subdifferential $\partial_{x_o} V(x_m, 0)$ consists entirely of strictly negative values:

$$\forall g \in \partial_{x_o} V(x_m, 0), \quad g < 0.$$

Since V is weakly concave in x_o by Lemma 2, and the entire subdifferential at $x_o = 0$ lies strictly below zero for $x_m \geq \bar{x}_m + \bar{r}$, Proposition 27.4 of Rockafellar (1970) implies that $V(x_m, x_o)$ is strictly decreasing at $x_o = 0$. Therefore, the static problem

$$\max_{x_o \in [0, 1-x_m]} V(1 - x_m - x_o, x_m, x_o)$$

has the unique maximizer $x_o^*(x_m) = 0$, which means $R(x_m) = 0$. □

Corollary 17. *For all $\mathbf{x} \in \mathcal{R}_1$, $\boldsymbol{\xi}(\mathbf{x}) \in \mathcal{R}$.*

Proof. From Lemma 16, for all $\mathbf{x} \in \mathcal{R}_1$, $x_o = 0$. Since $x_o = 0$, the restricted transition set from $\mathbf{x}$ is the entire iso-mature line, $\mathcal{M}(1 - x_{mt})$. The allocation $(1 - x_{mt}, r(1 - x_{mt})) \in \mathcal{M}(1 - x_{mt})$, hence $\boldsymbol{\xi}(\mathbf{x}) = (1 - x_{mt}, r(1 - x_{mt})) \in \mathcal{R}$. □

4.3.8 Transitions from $\mathcal{R}_4$

We analyze $\mathcal{R}_4 = \{\mathbf{x} \in \mathcal{R} : x_m \in [0, \bar{r})\}$ and partition it based on intersections with feasible transition sets. For allocations in $\mathcal{R}_4 \cap K_5$, optimal three-period transitions increase the mature component by at least $\delta_2 > 0$, ensuring optimal trajectories enter $K_2 \cup K_3$ within $N_2 = 3\lceil \frac{x_m^\dagger}{\delta_2} \rceil$ periods.

Using the definitions of the feasible transition sets (definitions 8 and 10), denote the boundary between K_2 and K_4 as $\text{bd}(K_{24}) = \{(x_m, (1-\bar{x}_m-\bar{r}-x_m)) : x_m \leq 1-\bar{x}_m-\bar{r}\}$ (see for example, figure 8). Note that this boundary coincides with the iso-young line defined by $x_y = \bar{x}_m + \bar{r}$.

Lemma 18. *For allocations in* $\text{bd}(K_{24})$, $\frac{\partial V}{\partial x_o} > 0$.

Proof. For all $\mathbf{x}_t \in \text{bd}(K_{24})$, $x_{m,t+1} = \bar{x}_m + \bar{r}$. Since $R(\bar{x}_m + \bar{r})$ is a singleton $\{0\}$ (Lemma 16), $\text{Tr}(\mathbf{x}_t) \cap R(x_{m,t+1})$ contains the single point $(\bar{x}_m + \bar{r}, 0)$; hence every selection of the optimal policy correspondence chooses that point.

Since $(\bar{x}_m + \bar{r}, 0) \in K_1$, by Lemma 9 it optimally transitions to $(1 - \bar{x}_m - \bar{r}, R(1 - \bar{x}_m - \bar{r})) \in K_0$. Note that $(1 - \bar{x}_m - \bar{r}, R(1 - \bar{x}_m - \bar{r})) = (1 - \bar{x}_m - \bar{r}, \bar{x}_m)$ since $(1 - \bar{x}_m - \bar{r}, R(1 - \bar{x}_m - \bar{r})) \in K_0$, $(\bar{x}_m, \bar{r}) \in K_0$ by definition, and $\mathcal{F}(\bar{x}_m, \bar{r}) = (1 - \bar{x}_m - \bar{r}, \bar{x}_m)$. Therefore the value function for any point $\mathbf{x} \in \text{bd}(K_{24})$ can be written in Bellman equation form:

$$V(\mathbf{x})|_{\mathbf{x}\in\text{bd}(K_{24})} = u(\mathbf{x}) + \beta u(\bar{x}_m + \bar{r}, 0) + \beta^2 V(1 - \bar{x}_m - \bar{r}, \bar{x}_m) \tag{22}$$

This expression is differentiable because it is the affine composition of two C^2 functions and the value function evaluated at a point in K_0.

Consider two allocations in $\text{bd}(K_{24})$, $\check{\mathbf{x}}(\check{x}_m, (1-\bar{x}_m-\bar{r}-\check{x}_m))$ and $\hat{\mathbf{x}} = (\hat{x}_m, (1-\bar{x}_m-\bar{r}-\hat{x}_m))$, where $(1-\bar{x}_m-\bar{r}) \geq \hat{x}_m > \check{x}_m$. Therefore $u'(\hat{\mathbf{x}}) < u'(\check{\mathbf{x}})$ since $f_m > f_o$. The derivatives of the value function with respect to x_o along the corresponding iso-mature lines at these two allocations are

$$\frac{dV(\mathbf{x}_t)}{dx_o}\bigg|_{\hat{\mathbf{x}}\in\text{bd}(K_{24})} = u'(\hat{\mathbf{x}})(f_o - f_y) + \beta u'(\bar{x}_m + \bar{r}, 0)(f_y - f_m) + \beta^2 \frac{dV(1 - \bar{x}_m - \bar{r}, \bar{x}_m)}{dx_m}$$

$$\frac{dV(\mathbf{x}_t)}{dx_o}\bigg|_{\check{\mathbf{x}}\in\text{bd}(K_{24})} = u'(\check{\mathbf{x}})(f_o - f_y) + \beta u'(\bar{x}_m + \bar{r}, 0)(f_y - f_m) + \beta^2 \frac{dV(1 - \bar{x}_m - \bar{r}, \bar{x}_m)}{dx_m}$$

Therefore

$$\frac{dV(\hat{\mathbf{x}})}{dx_o} - \frac{dV(\check{\mathbf{x}})}{dx_o} = \underbrace{(u'(\hat{\mathbf{x}}) - u'(\check{\mathbf{x}}))}_{<0}\underbrace{(f_o - f_y)}_{>0} \tag{23}$$

Hence in $\mathrm{bd}(K_{24})$, we can conclude that the derivative of (22) with respect to x_o decreases with x_m. Observe that $(1-\bar{x}_m-\bar{r},\bar{x}_m)\in\mathcal{R}_3\subset K_0$, so $R(1-\bar{x}_m-\bar{r})>0$ and $\frac{dV}{dx_o}|_{(1-\bar{x}_m-\bar{r},0)}>0$.

Therefore $\frac{dV(\mathbf{x})}{dx_o}>0$ for all $\mathbf{x}\in\mathrm{bd}(K_{24})$. □

Corollary 19. *$R_4\cap K_4=\emptyset$.*

Proof. From the explicit definitions of the K_i sets (definitions 8–10), $\mathcal{R}_4\subset K_2\cup K_3\cup K_4\cup K_5$. Lemma 18 implies that for all $\mathbf{x}\in\mathcal{R}_4$, $1-x_m-x_o<\bar{x}_m+\bar{r}$, so $\mathcal{R}_2\notin K_4$. □

Lemma 20. *For all $x_m<1-\bar{x}_m-\bar{r}$, $1-x_m\notin R(x_m)$.*

Proof. Suppose to the contrary that there exists $x_m\in[0,1-\bar{x}_m-\bar{r})$ such that $1-x_m\in R(x_m)$.

Let $\mathbf{x}_0=(x_m,1-x_m,0)$. Since $x_m<1-\bar{x}_m-\bar{r}$, we have $1-x_m>\bar{x}_m+\bar{r}$. By the definition of $\bar{x}_m$, the largest mature share in K_0, this allocation $\mathbf{x}_0\notin K_0$.

The feasible transition set from $\mathbf{x}_0$ is the iso-mature line $\mathcal{M}(x_m)$. Under the assumption $1-x_m\in R(x_m)$, the ridge point $\mathbf{x}_1=(0,x_m,1-x_m)$ must be an optimal choice in $\mathcal{M}(x_m)$.

Since $1-x_m>\bar{x}_m+\bar{r}$, $\mathbf{x}_1\in K_5$. Since $\mathbf{x}_1\in K_5$, any feasible transition from $\mathbf{x}_1$ must lie in K_4. But by Corollary 19, the ridge does not intersect K_4, so no ridge point is feasible. Therefore, by Definition 4, the only optimal transition is the Faustmann point $\mathbf{x}_2=\mathcal{F}(\mathbf{x}_1)=(1-x_m,0,x_m)$.

From $\mathbf{x}_2$, the only feasible transition is the aging map $\mathcal{F}(\mathbf{x}_2)=(x_m,1-x_m,0)=\mathbf{x}_0$. Thus $\mathbf{x}_0\to\mathbf{x}_1\to\mathbf{x}_2\to\mathbf{x}_0$ forms a 3-cycle in which each transition is optimal. But then $\mathbf{x}_0$ would lie in K_0, contradicting the fact that its mature share $1-x_m>\bar{x}_m$. Hence the assumption must be false, and $1-x_m\notin R(x_m)$ for all $x_m<1-\bar{x}_m-\bar{r}$. □

Lemma 21. *For $\mathbf{x}_t\in K_4$, $\mathbf{x}^*_{t+2}\in\mathbf{R}(()x_{mt}+x_{ot})$.*

Proof. Let $\mathbf{x}_t=(x_{mt},x_{ot})$ be an allocation in K_4. By the definition of K_4 (definition 10), $x_{yt}>\bar{x}_m+\bar{r}$ $(>\tilde{x}_m)$, so $x_{m,t+1}>\bar{x}_m+\bar{r}$.

By Lemma 16, whenever $x_{m,t+1}>\bar{x}_m+\bar{r}$ the ridge on the iso-mature line $\mathcal{M}(x_{m,t+1})$ is the singleton $\{(x_{m,t+1},0)\}$. Because this point lies in the restricted transition set $\Gamma^r(\mathbf{x}_t)$, Definition 4 implies

that the optimal policy correspondence selects it. Hence

$$\mathbf{x}^*_{t+1} = (1 - x_{mt} - x_{ot}, 0).$$

Since $x_{o,t+1} = 0$, the restricted transition set from $\mathbf{x}_{t+1}$ is the entire iso-mature line, $\mathcal{M}(x_{mt} + x_{ot})$. Hence the optimal transition from $\mathbf{x}^*_{t+1}$ to $\mathbf{x}_{t+2}$ will be to $\mathcal{R}$, i.e., to some $\xi(\mathbf{x}^*_{t+1}) \in \mathbf{R}(()x_{mt} + x_{ot})$. □

Lemma 22. *If $\mathbf{x}_t \in \mathcal{R}_4 \cap K_3$, then $\boldsymbol{\xi}(\mathbf{x}_t) \subset K_2$*

Proof. Using definition 9, $\mathbf{x}_t \in \mathcal{R}_4 \cap K_3$ implies $x_m < 1 - \bar{x}_m - \bar{r}$ and $\bar{x}_m < x_o < \bar{x}_m + \bar{r}$, so for all $\mathbf{x}_t \in \mathcal{R}_4 \cap K_3$, $\Gamma^r(\mathbf{x}) \subset K_2 \cup K_4$. By corollary 19, $\mathcal{R}_4 \cap K_4 = \emptyset$, so every element of $\boldsymbol{\xi}(\mathbf{x}_t)$ is contained in K_2 . □

Lemma 23. *If $\mathcal{R}_4 \cap K_5 \neq \emptyset$, $\mathcal{R}_4 \cap \mathrm{bd}(K_3 \cap K_5)$ is a singleton.*

Proof. Along $\mathrm{bd}(K_3 \cap K_5)$ the optimal policy reaches the interior of K_0 after exactly three periods (first to K_4, then to K_1 as a corollary of Lemma 21, then into K_0). So, for allocations in $\mathrm{bd}(K_3 \cap K_5)$, the value function can be written explicitly as

$$V(\mathbf{x})|_{\mathbf{x} \in \mathrm{bd}(K_{35})} = u(\mathbf{x}) + \beta u(\mathcal{F}\mathbf{x}) + \beta^2 u(x_o, 0) + \beta^3 V(1 - \bar{x}_m - \bar{r}, \bar{x}_m)$$

Because the Bellman operator composes three C^2 stage-utility terms before evaluating V at a point in the C^2 region K_0, the composite expression for V on $\mathrm{bd}(K_3 \cap K_5)$ is C^2.

Taking the derivative of this expression with respect to x_o and noting that the continuation payoff is constant with respect to x_o, we get

$$\left.\frac{dV(\mathbf{x})}{dx_o}\right|_{\mathbf{x} \in \mathrm{bd}(K_{35})} = u'(\mathbf{x})(f_o - f_y) + \beta u'(\mathcal{F}\mathbf{x})(f_y - f_m) + \beta^2 u'(x_o, 0)(f_m - f_y)$$

Further taking the derivative with respect to x_m yields

$$\left.\frac{d^2V(\mathbf{x})}{dx_m dx_o}\right|_{\mathbf{x}\in \mathrm{bd}(K_{35})} = \underbrace{u''(\mathbf{x})}_{<0}\underbrace{(f_o - f_y)}_{>0}\underbrace{(f_m - f_y)}_{>0} + \beta\,\underbrace{u''(\mathcal{F}\mathbf{x})}_{<0}\underbrace{(f_y - f_m)}_{<0}\underbrace{(f_o - f_m)}_{<0} + 0 \quad (<0)$$

where the last term vanishes because $(x_o, 0)$ is constant with respect to x_m.

Regarding the statement of the lemma, suppose to the contrary there exist $(x_m, \bar{x}_m + \bar{r}), (x'_m, \bar{x}_m + \bar{r}) \in \mathrm{bd}(K_3 \cap K_5)$ such that $\frac{dV}{dx_o} = 0$. By Rolle's theorem, there exists $x''_m \in [x_m, x'_m]$ such that $\frac{d^2V}{d_m d_o} = 0$. This is a contradiction since for all $\mathbf{x} \in \mathrm{bd}(K_3 \cap K_5)$, $\frac{d^2V(\mathbf{x})}{dx_m dx_o} < 0$. Therefore, $|\mathrm{bd}(K_3 \cap K_5)| \leq 1$ and the result is proved. □

Lemma 24. *For $\mathbf{x}\in\mathcal{R}_4$, there exists $\delta^d>0$, such that $|1-\bar{x}_m-\bar{r}-x_m| > \delta^d$, if and only if $\mathbf{x} \in K_5$.*

Proof. From definition 10, K_5 is the set of allocations with $x_o > \bar{x}_m + \bar{r}$. Recall $\mathrm{bd}(\mathcal{R}_3) \cap \mathrm{bd}(\mathcal{R}_4)$ is the singleton $(1 - \bar{x}_m - \bar{r}, \bar{x}_m)$. By the interiority of K_0, $\bar{r} > 0$, so $(\mathrm{bd}(\mathcal{R}_3) \cap \mathrm{bd}(\mathcal{R}_4)) \cap K_5 = \emptyset$.

Because the correspondence $R(.)$ is upper hemicontinuous and compact valued, for any $\epsilon' > 0$ there exists a $\delta' > 0$ such that whenever $|1 - \bar{x}_m - \bar{r} - x_m| < \delta'$, every element $r \in R(x_m)$ satisfies $|\bar{x}_m - r| < \epsilon'$. This follows because $R(1 - \bar{x}_m - \bar{r})$ is the singleton $\{\bar{x}_m\}$.

If $\mathcal{R}_4 \subset K_2 \cup K_3$, $\mathcal{R}_4 \cap K_5 = \emptyset$. Let $\delta^d = 1 - \bar{x}_m - \bar{r}$ and the lemma holds vacuously.

On the other hand, if $\mathcal{R}_4 \cap K_5 \neq \emptyset$, by the ridge's single crossing of $\mathrm{bd}(K_{35})$ and that the ridge is a singleton at this crossing (Lemma 23), there exists a unique $x'_m \in [0, 1 - \bar{x}_m - \bar{r})$ such that $\mathcal{R}_4 \cap \mathrm{bd}(K_{35}) = (x'_m, r(x'_m)) = (x'_m, \bar{x}_m + \bar{r})$.

Let $\delta^d = 1 - \bar{x}_m - \bar{r} - x'_m$ and the result is proved. □

Corollary 25. *There exists $\delta_2 > 0$, such that for every state $\mathbf{x}_t\in\mathcal{R}_4 \cap K_5$ and every optimal three-period transition selection $\xi^3(\mathbf{x}_t) \in \boldsymbol{\xi}^3(\mathbf{x}_t)$, $\xi^3(\mathbf{x}_t) > x_{mt} + \delta_2$.*

Proof. For all $\mathbf{x}_t\in\mathcal{R}_4 \cap K_5$, $x_{ot} > \bar{x}_m + \bar{r}$ from the definition of K_5 (Definition 10), and $x_{y,t+1} > \bar{x}_m + \bar{r}$, so these allocations can only feasibly transition to K_4.

Using corollary 19, the opimal first step is the Faustmann transition, $\mathbf{x}^*_{t+1} = \mathcal{F}\mathbf{x}_t = (x_{ot}, x_{yt}, x_{mt})$. Using Lemma 21, the unique optimal successor to $\mathbf{x}^*_{t+1}$ is $\mathbf{x}^*_{t+2} = (x_{yt} + x_{mt}, x_{ot}, 0)$ and $\boldsymbol{\xi}^3(\mathbf{x}_t) = \{(x_{yt} + x_{mt}, r)\}$ for all selections $r \in R(x_{yt} + x_{mt})$. Hence for every selection $\xi^3(\mathbf{x}_t) \in \boldsymbol{\xi}^3(\mathbf{x}_t)$, $\xi^3_m(\mathbf{x}_t) = x_{yt} + x_{mt}$. Using Lemma 20, $1 - x_{mt} = x_{yt} > 0$ for all $x_{mt} < 1 - \bar{x}_m - \bar{r}$, so $x_{yt} + x_{mt} > x_{mt}$.

From lemma 23, $\mathcal{R}_4 \cap (\mathrm{bd}(K_3) \cap \mathrm{bd}(K_5))$ is a singleton. Denote this allocation as $\mathbf{x}^\dagger$ and note $x^\dagger_o = \bar{x}_m + \bar{r}$ by construction. Recall that $R(1 - \bar{x}_m - \bar{r}) = \{\bar{x}_m\}$ with unique selection $r(1 - \bar{x}_m - \bar{r}) = \bar{x}_m$, so $(1 - \bar{x}_m - \bar{r}, r(1 - \bar{x}_m - \bar{r})) = (1 - \bar{x}_m - \bar{r}, \bar{x}_m)$ and $(1 - \bar{x}_m - \bar{r}, \bar{x}_m + \bar{r}) \notin \mathcal{R}$ by the interiority of K_0. Therefore $x^\dagger_y > 0$.

The closure $\mathrm{cl}((\mathcal{R}_4 \cap K_5) \cup \{\mathbf{x}^\dagger\})$ is compact and for all $\mathbf{x} \in \mathrm{cl}((\mathcal{R}_4 \cap K_5) \cup \{\mathbf{x}^\dagger\})$, $x_y > 0$. Therefore, by compactness, there exists $\delta_2 > 0$ such that for all $\mathbf{x} \in (\mathcal{R}_4 \cap K_5) \cup \{\mathbf{x}^\dagger\}$, $x_y > \delta_2$.

Finally, for every $\mathbf{x}_t \in \mathcal{R}_4 \cap K_5$ and every selection $\xi^3(\mathbf{x}_t) \in \boldsymbol{\xi}^3(\mathbf{x}_t)$, we have $\xi^3(\mathbf{x}_t) = x_{yt} + x_{mt} > x_{mt} + \delta_2$. □

Corollary 26. *Beginning from an allocation $\mathbf{x} \in \mathcal{R}_4 \cap K_5$, the optimal trajectory will enter $K_2 \cup K_3$ in at most $N_2 = 3\left\lceil \frac{x^\dagger_m}{\delta_2} \right\rceil$ periods (where $x^\dagger_m$ and δ_2 were defined in Lemma 25).*

Proof. For each $\mathbf{x} \in \mathcal{R}_4 \cap K_5$, let $q(\mathbf{x}) = \left\lceil \frac{x^\dagger_m - x_m}{\delta_2} \right\rceil$. That is, x_m is no greater than $q(\mathbf{x})$ units of length δ_2 less than $x^\dagger_m$. Since $\mathbf{x}^\dagger \notin \mathcal{R}_4 \cap K_5$ and $0 \leq x_m < x^\dagger_m$, $3\left\lceil \frac{x^\dagger_m - x_m}{\delta_2} \right\rceil \leq 3\left\lceil \frac{x^\dagger_m}{\delta_2} \right\rceil = N_2$, it follows that $3q(\mathbf{x}) \leq N_2$.

Fix $\mathbf{x}_t \in \mathcal{R}_4 \cap K_5$ such that $q(\mathbf{x}_t) = 1$. By Lemma 25, choose any $\xi^3(\mathbf{x}_t) \in \boldsymbol{\xi}^3(\mathbf{x}_t)$, then $\xi^3_m(\mathbf{x}_t) > x_{mt} + \delta_2 \geq x^\dagger_m$ and $\boldsymbol{\xi}^3(\mathbf{x}_t) \subset K_2 \cup K_3$. So $\mathbf{x}_t$ optimally enters $K_2 \cup K_3$ in $3 \leq 3q(\mathbf{x}_t)$ periods.

Now fix $1 \leq n < \left\lceil \frac{N_2}{3} \right\rceil$ and assume that every $\mathbf{x}_t \in \mathcal{R}_4 \cap K_5$ with $q(\mathbf{x}_t) = n$ enters $K_2 \cup K_3$ within $3n$ periods. Suppose there exists $\mathbf{x}_t \in \mathcal{R}_4 \cap K_5$ such that $q(\mathbf{x}_t) = n + 1$. From Lemma 25, choose any $\xi^3(\mathbf{x}_t) \in \boldsymbol{\xi}^3(\mathbf{x}_t)$ such that $\xi^3_m(\mathbf{x}_t) > x_{mt} + \delta_2$, implying that $q(\xi^3_m(\mathbf{x}_t)) \leq n$ and $\xi^3(\mathbf{x}_t)$ enters $K_2 \cup K_3$ in at most $3n$ periods. Therefore $\mathbf{x}_t$ enters $K_2 \cup K_3$ in at most $3 + 3q(\xi^3_m(\mathbf{x}_t)) \leq 3 + 3n = 3(n+1)$ periods.

Therefore, for any $\mathbf{x}_t \in \mathcal{R}_4 \cap K_5$, $\mathbf{x}_t$ enters $K_2 \cup K_3$ in at most $3q(\mathbf{x}_t) \leq N_2$ periods, completing

the proof. □

4.3.9 Uniform Convergence

Proposition 4.2. *For the problem defined by conditions (5a)-(5e), there exists $N = 6 + N_1 + N_2$ such that the optimal trajectory starting from any $\mathbf{x}_0 \in \Delta^2$ converges to K_0 in at most N periods (N_1 is defined in Corollary 14 and N_2 is defined in Corollary 26).*

Proof. Consider each part of the simplex in turn. First, for all $\mathbf{x}_0 \in \Delta^2 \setminus \mathcal{R}$, the optimal trajectory transitions to $\mathcal{R}$ in at most 3 periods (Lemma 8). Now consider each segment of the ridge in turn.

From definition 8 and Proposition 4.1, for all $\mathbf{x} \in \mathcal{R}_2$, the optimal trajectory reaches K_0 in at most N_1 periods.

From 18c, $\mathcal{R}_3 \subset K_0$ so for all $\mathbf{x} \in \mathcal{R}_3$ the optimal trajectory is already in K_0 (reaches K_0 in 0 periods).

From definition 10, definition 18d, and Corollary 19, $\mathcal{R}_4 \subset K_2 \cup K_3 \cup K_5$. Note that for all $\mathbf{x} \in K_2$, the feasible transition set intersects K_1 (definition 6), and for any selection $\xi(\mathbf{x}_t) \in \boldsymbol{\xi}(\mathbf{x})$, we have $\xi(\mathbf{x}_t) \in K_1 \cup \mathcal{R}_2$, which optimally transitions to K_0 in at most $\max\{1, N_1\} = N_1$ periods. Therefore, for all $\mathbf{x} \in \mathcal{R}_4 \cap K_2$, the optimal trajectory enters K_0 in at most $1 + N_1$ periods.

For $\mathbf{x} \in \mathcal{R}_4 \cap K_3$, the feasible transition set is a subset of $K_2 \cup K_4$ and has a non-empty intersection with K_2 (definition 8) so for any selection $\xi(\mathbf{x}_t) \in \boldsymbol{\xi}(\mathbf{x})$, $\xi(\mathbf{x}_t) \in K_2$ (Lemma 19). From the reasoning above, any such $\xi(\mathbf{x}_t) \in K_2$ enters K_0 in at most $1 + N_1$ periods. Therefore, for all $\mathbf{x} \in \mathcal{R}_4 \cap K_3$, the optimal trajectory enters K_0 in at most $2 + N_1$ periods.

For $\mathbf{x} \in \mathcal{R}_4 \cap K_5$, the optimal trajectory reaches $\mathcal{R}_4 \cap (K_2 \cup K_3)$ in at most N_2 periods (Corollary 26). From the reasoning above $\mathbf{x}$ enters K_0 in at most $\max\{1 + N_1, 2 + N_1\} + N_2 = 2 + N_1 + N_2$ periods. Therefore, for all $\mathbf{x} \in \mathcal{R}_4$, the optimal trajectory enters K_0 in at most $\max\{1 + N_1, 2 + N_1, 2 + N_1 + N_2\} = 2 + N_1 + N_2$ periods.

From definition 7 and definition 18a, $\mathcal{R}_1 \subset K_1 \cup K_2 \cup K_3$. For $\mathbf{x} \in \mathcal{R}_1 \cap K_1$, the optimal trajectory reaches K_0 in one period (Lemma 9). For $\mathbf{x} \in \mathcal{R}_1 \cap K_2$, the optimal trajectory reaches K_0 in at

most N_1 periods (Lemma 14). For $\mathbf{x} \in \mathcal{R}_1 \cap K_3$, $\boldsymbol{\xi}(\mathbf{x}) \subset \mathcal{R}_4$ by definition 9 and corollary 17. As shown above, the optimal trajectory starting at $\mathbf{x} \in \mathcal{R}_4$ transitions to K_0 in at most $2 + N_1 + N_2$ periods. Therefore, for $\mathbf{x} \in \mathcal{R}_1 \cap K_3$, the optimal trajectory reaches K_0 in at most $3 + N_1 + N_2$ periods.

Therefore, the optimal trajectory enters K_0 in at most $6+N_1+N_2$ periods, completing the proof. □

4.4 Numerical analysis: Trajectories and comparative statics for the three-age-class model

To analyze the three-age-class model numerically, we employ the running horizon algorithm detailed in Appendix A.1. We generate approximation sequences of length $S = 25$ periods, sufficient for trajectories to reach and demonstrate steady-state cycle behavior. The finite horizon length varies with the discount factor: $T = 25$ periods for most cases, but $T = 55$ periods when $\beta = 0.95$ to ensure adequate time for convergence.

4.4.1 Example Trajectory

Figure 9 illustrates the optimal trajectory beginning from an initial allocation of all young trees, demonstrating how an arbitrary starting condition converges to the optimal cycle region. This example assumes that $\beta = 0.75$ and $\mathbf{f} = (1, 3, 2)$

In the initial period, the orchard consists entirely of young trees. Following optimal aging, period one transitions to all mature trees. The 'first' decision occurs in period two, where the grower faces the choice between allowing all mature trees to age naturally versus implementing early replacement for consumption smoothing. The optimal solution involves early replacement of just over one-third of the mature trees before they reach their Faustmann age, while retaining just under two-thirds as old trees. This early replacement decision reflects the economic trade-off between current period production losses and improved future age structure for consumption smoothing.

The subsequent transition shows the orchard moving to a state with only young and mature trees in period three, as all old trees have been harvested. By period four, the optimal policy maintains this

structure temporarily before the trajectory enters the optimal three-period cycle region in period five. The convergence occurs within six periods, so N_1 and N_2 are zero in this case.

This trajectory pattern demonstrates the economic mechanism underlying optimal orchard management: the grower willingly sacrifices short-term production to achieve long-term stability within the cycle region. The early replacement of mature trees, while costly in terms of immediate foregone output, generates the optimal age distribution for sustaining the consumption-smoothing cycle indefinitely.

4.4.2 Numerical Comparative Statics

Figure 10 presents comparative statics examining how the cycle region responds to changes in economic and biological parameters. These results provide a numerical analog to the theoretical comparative statics derived for the two-age-class model.

Panel A shows how the discount factor β influences the optimal cycle region. As β increases from 0.5 to 0.95, the cycle region contracts around the even-aged allocation. This contraction reflects the stronger preference for consumption smoothing among more patient growers. When $\beta = 0.5$, representing relatively impatient preferences, the cycle region encompasses a large portion of the state space, allowing for considerable production variation depending on initial conditions. Conversely, with $\beta = 0.95$, the cycle region becomes concentrated near the even-aged orchard $(1/3, 1/3, 1/3)$, indicating that patient growers optimally maintain nearly constant production levels. The economic intuition follows from our theoretical analysis: higher discount factors increase the relative importance of future utility, making consumption smoothing more valuable relative to current period optimization.

Panel B shows how the productivity of old trees affects optimal cycle patterns. When $f_o = 2$, creating a decline from mature to old tree productivity, the cycle region exhibits moderate amplitude. As f_o increases to 2.5, the cycle region expands. By $f_o = 3$, where old trees are as productive as mature trees, the cycle region extends to the boundary of the simplex. As the yield differences decrease between the mature and old trees, the production variation generated by the age class

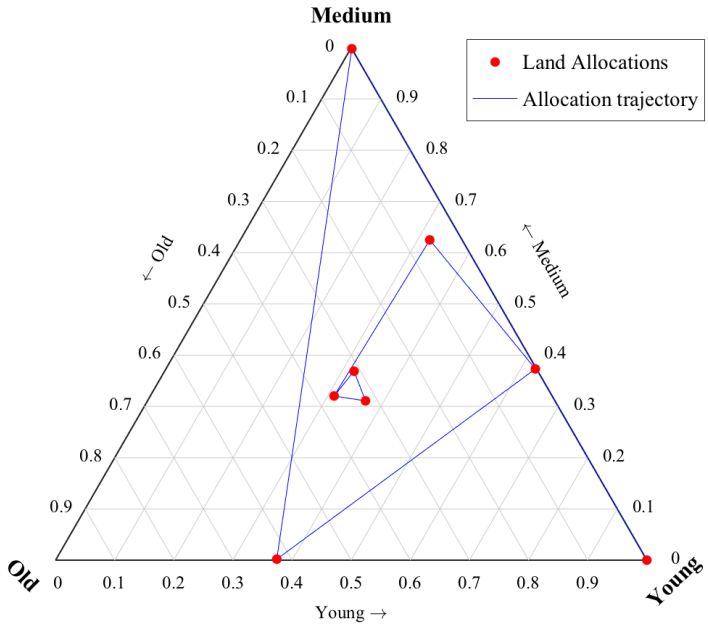
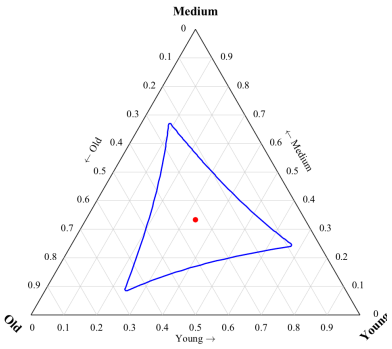
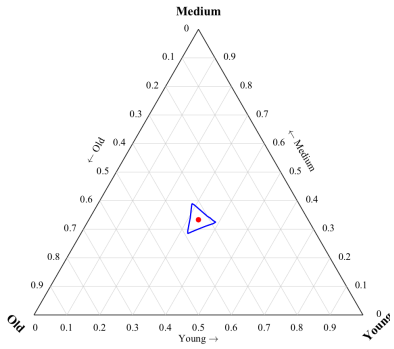


Figure 9: Example trajectory starting from all young trees with $\beta = 0.75$, $f = (1, 3, 2)$

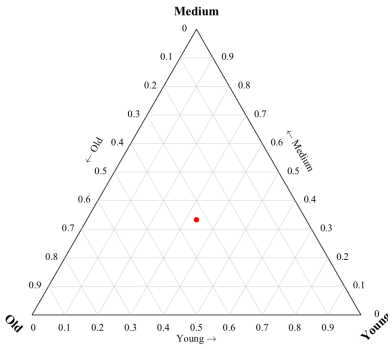
diminishes, pushing the balance toward the preference for current utility.



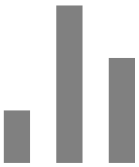
(a) $\beta = 0.5$



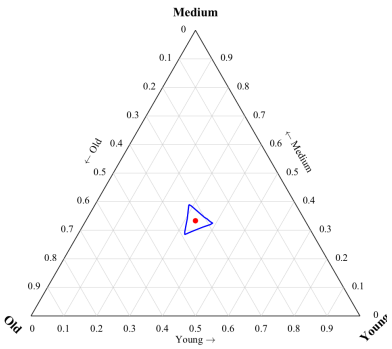
(b) $\beta = 0.75$



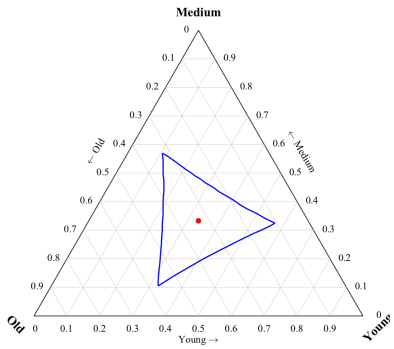
(c) $\beta = 0.95$



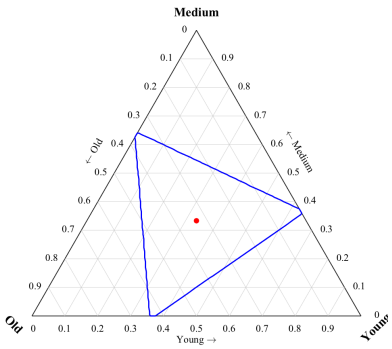
A. Varying β at $f_o = 2$.



(d) $f_o = 2$



(e) $f_o = 2.5$



(f) $f_o = 3$

B. Varying f_o at $\beta = 0.75$.

Figure 10: Combined cycle region examples (top rows) and their directly paired yield curves (bottom rows). All panels use $f_y = 1$ and $f_m = 3$.

5 Discussion

In their seminal work, Mitra et al. (1991) established that under monotonically increasing yields (where older trees always produce more than younger ones), optimal programs never converge to a stationary orchard,[10] rather they permanently oscillate away from any steady state. Moreover, they proved a neighborhood turnpike theorem showing that, as the discount factor approaches unity, optimal paths remain within some neighborhood of the stationary optimal orchard.

Mitra, Ray and Roy also provide a hump-shaped yield example (Example 5.2) that produces infinitely many stationary orchards. Yet the two dynamic theorems with strictly concave utility they prove—Propositions 5.2 and 5.3—track how an orchard moves relative to just one even-aged orchard. They do not identify the set of states an orchard may revisit under non-monotonic yields, nor do they say whether the motion eventually settles into a particular pattern. Consequently, the long-run form of optimal sequences with hump-shaped yields remains unresolved.

We resolve this gap. With a three-age-class, hump-shaped yield profile we (i) identify a precise cycle region K_0 that generates optimal three-period cycles; and (ii) prove that every optimal path, for every discount factor in $(0, 1)$, reaches K_0 after at most N periods. Thus our convergence result is stronger than a neighbourhood turnpike: it applies for all β, gives exact convergence to a set of optimal cycles (not just proximity to a point), and provides a finite-time bound. Because K_0 exists under hump-shaped yields, our findings extend the earlier insight that even-aged age structures are not uniquely optimal to a wider class of perennial-crop models.

Importantly, there is no contradiction between our findings and those of Mitra et al. (1991). The different convergence properties emerge directly from the different biological assumptions about yield patterns. While Mitra et al. (1991) provided a neighborhood turnpike result that applies generally, they showed non-convergence specifically under monotonically increasing yields. Our contribution is demonstrating finite-time convergence to optimal cycles under hump-shaped yields, contrasting with their non-convergence result for the monotonic case.

[10] Towards the end of their paper, Mitra et al. use "forest" when referring to orchards. Here, we refer to their results in terms of orchards.

Results in forestry suggest that optimal cyclic sequences may not be a general feature of point-input, point-output capital theory models. It remains to be seen whether these results carry over to the point-input, flow-output setting. Our model assumes a fixed total land area with the decision problem centered on allocating this area among different age classes. This approach contrasts with Salo and Tahvonen (2004), who examine endogenous land allocation between forestry and alternative uses. They show that cyclical patterns may vanish when land can be freely converted between uses (as land moves in and out of forestry to smooth production), while adding conversion costs reintroduces optimal cycles and path-dependent equilibria.

However, this concern is less relevant for established orchard operations for several reasons. First, orchards involve substantial sunk investments in specialized infrastructure including irrigation systems, drainage, soil amendments, and tree establishment costs that are orchard-specific and cannot be recovered through temporary conversion to alternative uses. Second, unlike annual crops, orchard establishment requires multiple years before achieving full productivity, making temporary entry and exit economically prohibitive. Third, even if growers could temporarily cease harvesting to smooth production (analogous to leaving land fallow), they would still face the fundamental age-structure optimization problem since the trees continue aging and the optimal management of existing tree cohorts remains unchanged. Additionally, Fabbri et al. (2015) has shown that cyclical solutions optimal in discrete-time forestry models may not persist in continuous-time formulations. However, our discrete-time approach aligns naturally with the seasonal nature of orchard harvesting.

These findings have direct applications for perennial crop managers. Understanding that even-aged age structures, while optimal, are not uniquely so provides flexibility in transition planning, allowing growers to adapt replanting strategies to their operational constraints. Furthermore, our analysis of cycle amplitude responses to economic parameters enables growers to anticipate how changes in discount rates or relative productivity between age classes will affect optimal orchard structures. This knowledge can inform long-term planning and help growers adapt to changing economic conditions without sacrificing optimality. Moreover, these results suggest policies aimed at reducing production volatility should focus on planning for the economic consequences of unavoidable cycles

rather than attempting to eliminate cycles entirely.

6 Conclusion

We have extended the theoretical framework for optimal management of age-structured orchards by characterizing transition dynamics in the two-age-class model and developing a novel three-age-class model accommodating non-monotonic yields. Our work builds upon the seminal contribution of Mitra et al. (1991), who pioneered the analysis of orchards as a canonical example of point-input, flow-output capital theory.

Returning to our central research questions, we can now provide clear answers: (1) Regarding optimal replacement age, we demonstrate that it depends critically on the yield curve and discount factor, with the even-aged three-age-class orchard being optimal under specified conditions, but also showing multiple other optimal solutions exist within the cycle region. (2) On whether growers should replace all trees simultaneously, our analysis reveals that partial replacement is typically optimal, creating a cyclical age structure that maximizes value over time. (3) Concerning the steady-state and convergence, we prove that although the even-aged orchard remains a stationary optimum, optimal dynamics are typically cyclical. In the two-age-class model, every optimal sequence is contained within a two-period cycle region (which includes the even-aged orchard) in at most 2 periods. In the three-age-class model, we derive a similar result showing that all optimal sequences are contained within a three-period cycle region after a finite number of periods.

the optimal steady-state is a cyclical process rather than a fixed point, and we establish uniform convergence to this cycle in at most two periods for the two-age-class model and characterize the finite number of periods (at most N) for the three-age-class model.

Our complete characterization of transition dynamics—proving convergence to optimal cycles in at most two periods for the two-age-class model and in at most N periods for the three-age-class model—represents a significant advancement over previous work that either omit transitions or provided only asymptotic results. Additionally, our work corrects an omission in Wan (1993), whose solution applies when the aging constraint binds but does not include cases where young trees are optimally replaced early.

While Mitra et al. (1991) demonstrated that for monotonically increasing yields, optimal programs never converge to a stationary optimal forest, we complement their findings by showing that with non-monotonic yields, optimal programs converge to a cycle region in finite time. Their neighborhood turnpike result established proximity to the even-aged orchard when the discount factor is sufficiently close to one. In contrast, our convergence theorem is stronger—establishing exact convergence to a set of optimal cycles for all relevant discount factors, and within a finite number of periods.

Although our results establish strong convergence properties for perennial crops under our modeling assumptions, we acknowledge that these findings may not fully generalize across all settings. As noted in our discussion, research by Salo and Tahvonen (2004) and Fabbri et al. (2015) suggests that optimal cycles may be sensitive to assumptions about land conversion flexibility and continuous versus discrete time frameworks. Future work could explore whether their results also apply to the orchards context.

6.1 Declaration of competing interests

The authors have no competing interests to declare

6.2 Funding Sources

Tregeagle's research is supported by the Research Capacity Fund (HATCH), project award no. 1024582, from the U.S. Department of Agriculture's National Institute of Food and Agriculture.

6.3 Research Data

No empirical data was used in this study. Code for the generation of figures in this article is available from the corresponding author on request.

6.4 Declaration of generative AI and AI-assisted technologies in the writing process.

During the preparation of this work the authors used ChatGPT from OpenAI and Claude from Anthropic in order to improve the clarity and flow of the manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

de la Fuente, A. (2000). *Mathematical Methods and Models for Economists*. Cambridge: Cambridge University Press.

Devadoss, S. and J. Luckstead (2010, March). An analysis of apple supply response. *International Journal of Production Economics 124*(1), 265–271.

Fabbri, G., S. Faggian, and G. Freni (2015). On the mitra-wan forest management problem in continuous time. *Journal of Economic Theory 157*, 1001–1040.

FAO (2024). *World Food and Agriculture – Statistical Yearbook 2024*. FAO.

Feinerman, E. and Y. Tsur (2014). Perennial crops under stochastic water supply. *Agricultural Economics 45*(6), 757–766.

Franklin, B. (2013). *The Dynamics of Irrigated Perennial Crop Production With Applications to the Murray-Darling Basin of Australia*. Ph. D. thesis, UC Riverside.

Luckstead, J. and S. Devadoss (2024). Optimal Timing of Removal and Planting of Perennial Crops in Anticipation of a New Cultivar. *Journal of Agricultural and Resource Economics 49*(3), 411–435.

Mitra, T. (2000). *Introduction to Dynamic Optimization Theory*, Volume 11, pp. 31–108. Berlin, Heidelberg: Springer Berlin Heidelberg.

Mitra, T., D. Ray, and R. Roy (1991). The economics of orchards: An exercise in point-input, flow-output capital theory. *Journal of Economic Theory 53*(1), 12–50.

Mitra, T. and H. Y. Wan (1985). Some Theoretical Results on the Economics of Forestry. *The Review of Economic Studies 52*(2), 263–282.

Mitra, T. and H. Y. Wan (1986). On the Faustmann solution to the forest management problem. *Journal of Economic Theory 40*(2), 229–249.

Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica 32*(1/2), 122–136.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press.

Rockafellar, R. T. and R. J. B. Wets (1998). *Variational Analysis*. Grundlehren Der Mathematischen Wissenschaften. Berlin, Heidelberg: Springer.

Salo, S. and O. Tahvonen (2002). On Equilibrium Cycles and Normal Forests in Optimal Harvesting of Tree Vintages. *Journal of Environmental Economics and Management 44*(1), 1–22. They invoke Theorem 4.6 in Stokey and Lucas to ensure the existence of an optimal solution.

Salo, S. and O. Tahvonen (2003). On the economics of forest vintages. *Journal of Economic Dynamics and Control 27*(8), 1411–1435.

Salo, S. and O. Tahvonen (2004). Renewable Resources with Endogenous Age Classes and Allocation of Land. *American Journal of Agricultural Economics 86*(2), 513–530.

Siegle, J., G. Astill, Z. Plakias, and D. Tregeagle (2024). Estimating perennial crop supply response: A methodology literature review. *Agricultural Economics 55*(2), 159–180.

Stokey, N. L., R. E. Lucas, and E. C. Prescott (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.

Wan, H. (1993). A Note on Boundary Optimal Paths. In R. Becker, M. Boldrin, R. Jones, and W. Thomson (Eds.), *General Equilibrium, Growth, and Trade*, pp. 411–426. Academic Press.

A Appendices

A.1 The running horizon algorithm

The running horizon algorithm approximates the solution to an infinite horizon dynamic problem by calculating the solution to a sequence of finite horizon problems. This description is adapted from Franklin (2013) and Salo and Tahvonen (2004). The relevant parts of their papers are included in the appendix.

We wish to find a solution to the non-linear programming problem

$$V(\mathbf{x}_0) = \max_{\{\mathbf{x}_t\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(c_t) \quad \text{s.t. constraints}$$

Numerically solving this problem directly requires approximating the value function, a computationally intensive and unstable process.

Alternatively, we can repeatedly solve the finite horizon analog of the infinite horizon problem to obtain an approximation of the infinite horizon solution. The finite horizon analog is given by:

$$V^T(\mathbf{x}_0) = \max_{\{\mathbf{x}_t\}_{t=1}^{T}} \sum_{t=0}^{T} \beta^t u(c_t) \quad \text{s.t. constraints}$$

Solving this problem is relatively straightforward, and can be directly implemented with a non-linear numerical optimization algorithm, e.g. the `fmincon` command in MATLAB.

The optimal sequence in this finite horizon problem is $\{\mathbf{x}_t^*(T, \mathbf{x}_0)\}_{t=1}^{T}$. The parentheses following $\mathbf{x}_t$ denote the dependence of the solution on the length of the finite horizon, T, and the initial condition $\mathbf{x}_0$.

The running horizon algorithm generates a vector of length S, approximating the first S terms of the infinite horizon problem, $\{\hat{\mathbf{x}}_t\}_{t=1}^{S} \approx \{\mathbf{x}_t^{(\infty, \mathbf{x}_0)}\}_{t=1}^{S}$. The choice of S may affect the accuracy of the approximation, since errors in the approximation in the early terms will propagate to the later terms. For the orchard management problem, we want to pick an S large enough for the infinite

horizon approximation to reach the steady state.

The first term of the approximation vector is the first term of the solution to the T period problem, starting from $\mathbf{x}_0$. That is, $\hat{\mathbf{x}}_1 = \mathbf{x}_1^*(T, \mathbf{x}_0)$. The second term of the approximation vector is generated by solving the T period problem, using $\hat{\mathbf{x}}_1$ as the initial condition, and taking the first term of the solution to this finite horizon problem. That is, $\hat{\mathbf{x}}_2 = \mathbf{x}_1^*(T, \hat{\mathbf{x}}_1)$. The remaining $S-2$ terms of the approximation vector are generated in a similar manner. That is, $\hat{\mathbf{x}}_s = \mathbf{x}_1^*(T, \hat{\mathbf{x}}_{s-1})$.

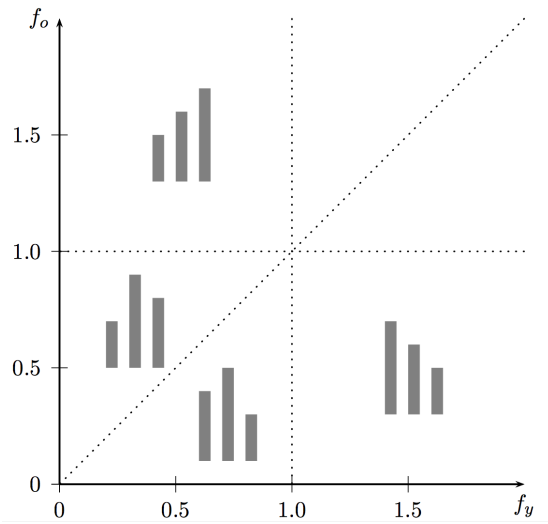
This algorithm works because the first period solution to the finite horizon problem approaches the first period solution to the infinite horizon problem as the finite horizon approaches infinity.

$$\lim_{T\to\infty} \mathbf{x}_1^{(T,\mathbf{x}_0)} = \mathbf{x}_1^{(\infty,\mathbf{x}_0)}$$

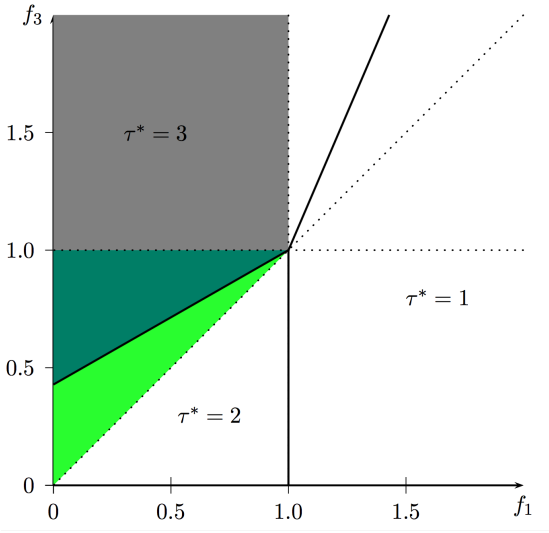
This occurs because, with positive discounting, the behavior of the finite horizon problem at the terminal time has a diminishing effect on the first period choice as the horizon is extended.

A.2 Comparison of yield curves between this model and Proposition 5.2 in Mitra et al. (1991).

Figure 12a is a graphical representation of the set of three-age-class yield curves. The yield of mature trees is normalized to one, with the relative yield of young trees on the horizontal axis, and the relative yield of all trees on the vertical axis. Figure 12b shows a specific case of the general optimal replacement age figure (Figure 2), highlighting the set of three-age-class yield curves applicable to the asymptotic Turnpike theorem in Mitra et al. (1991) (Proposition 5.2), represented by the gray square, and the set of yield curves that can be analyzed using propositions 4.1 and 4.2 in our three-age-class model. However, the analysis of Mitra et al. (1991) is applicable to yield curves with arbitrarily many age-classes, so long as yields are (weakly) monotonically increasing, whereas our analysis is restricted to 3-age-class yield curves.

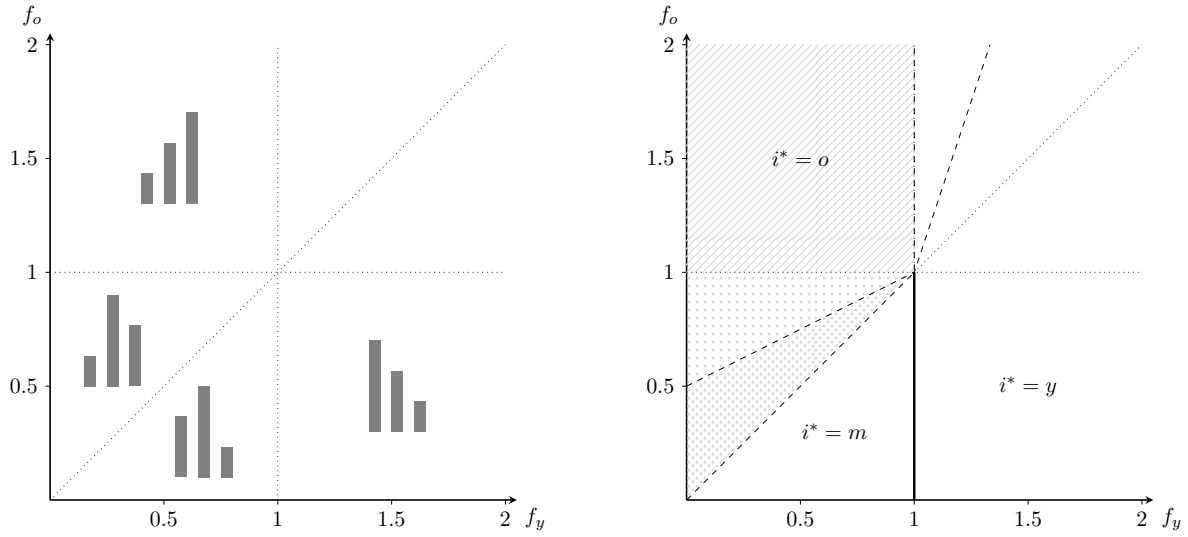


(a) The space of three-age-class yield curves, with f_m normalized to 1.



(b) The Faustmann age as a function of the yield curve. The gray square represents the yield curves studied by Mitra et al. (1991). The light green ($i^* = o$) and dark green ($i^* = o$) regions represent the non-monotonic yield curves that can additionally be analyzed by our model.

Figure 11: A comparison of yield curves studied in this model and Proposition 5.2 in Mitra et al. (1991).



(a) The space of three-age-class yield curves, with f_m normalized to 1.

(b) The Faustmann age as a function of the yield curve.

Figure 12: A comparison of yield curves studied in this model and Proposition 5.2 in Mitra et al. (1991). In Figure (b), the lined square represents the three-age-class yield curves admissable in their Proposition 5.2. The dotted regions represent the non-monotonic yield curves that can additionally be analyzed by our model. The extent of the $i^* = 3$ region depends on the discount factor (see Figure 2).

A.3 Proofs

A.3.1 Proof of proposition 3.1

Proof of proposition 3.1 (page 14). First, we will show that if a two-period cycle is optimal, then $\beta \leq \frac{u'(c_{t+1})}{u'(c_t)} \leq \frac{1}{\beta}$ (necessity).

From equations (2b) and (2c) and assuming an interior allocation using all available land, we find an expression for λ_t:

$$\lambda_t = \beta(u'(c_{t+1})(f_o - f_y) - \lambda_{t+1})$$

Shifting this equation forward by one period gives an expression for λ_{t+1}:

$$\lambda_{t+1} = \beta(u'(c_{t+2})(f_o - f_y) - \lambda_{t+2})$$

Substituting the expression for λ_{t+1} into the expression for λ_t and assuming that there is a two-period cycle, i.e., $x_{yt} = x_{y,t+2}$, $x_{ot} = x_{o,t+2}$, gives:

$$\lambda_t = \beta(f_o - f_y)(u'(c_{t+1}) - \beta u'(c_t)) + \beta^2 \lambda_t$$

For the two-period cycle to be optimal, λ_t must be non-negative. Rearranging gives:

$$\lambda_t = \frac{\beta(f_o - f_y)}{1 - \beta^2}(u'(c_{t+1}) - \beta u'(c_t))$$

For this expression to be non-negative, we must have:

$$u'(c_{t+1}) - \beta u'(c_t) \geq 0 \Rightarrow \frac{u'(c_{t+1})}{u'(c_t)} \geq \beta$$

Similarly for λ_{t+1}:

$$u'(c_t) - \beta u'(c_{t+1}) \geq 0 \Rightarrow \frac{u'(c_{t+1})}{u'(c_t)} \leq \frac{1}{\beta}$$

Combining these inequalities:

$$\beta \leq \frac{u'(c_{t+1})}{u'(c_t)} \leq \frac{1}{\beta}$$

Thus, if a two-period cycle is optimal, this condition must hold.

Now, we will show that if $\beta \leq \frac{u'(c_{t+1})}{u'(c_t)} \leq \frac{1}{\beta}$, then a two-period cycle is optimal (sufficiency).

Suppose we have a land allocation sequence $\{x_{yt}, x_{ot}\}_{t=1}^{\infty}$ that follows a two-period cycle and satisfies $\beta \leq \frac{u'(c_{t+1})}{u'(c_t)} \leq \frac{1}{\beta}$ for all t. We need to show that there exists a sequence of non-negative Lagrange multipliers that, together with this land allocation, satisfy the KKT conditions.

Using the inequality $\frac{u'(c_{t+1})}{u'(c_t)} \geq \beta$, we define:

$$\lambda_t = \frac{\beta(f_o - f_y)}{1-\beta^2}(u'(c_{t+1}) - \beta u'(c_t))$$

Since $f_o > f_y$ and $u'(c_{t+1}) - \beta u'(c_t) \geq 0$, we have $\lambda_t \geq 0$.

Similarly, using $\frac{u'(c_{t+1})}{u'(c_t)} \leq \frac{1}{\beta}$, we have $\lambda_{t+1} \geq 0$.

With these non-negative values for λ_t and λ_{t+1}, the KKT conditions are satisfied, making the two-period cycle optimal.

Therefore, if $\beta \leq \frac{u'(c_{t+1})}{u'(c_t)} \leq \frac{1}{\beta}$, then a two-period cycle is optimal, completing the proof of both necessity and sufficiency. □

A.3.2 Proof of corollary 1

Proof of corollary 1 (page 14). For each period in a even-aged orchard, half of the land is allocated to young trees and half to old trees. Since the land allocation is the same every period, the production is the same every period $c_t = c_{t+1}$.

Applying this production path to the inequality in Proposition 3.1 gives

$$\beta \leq \frac{u'(c_t)}{u'(c_t)} \leq \frac{1}{\beta} \quad \Rightarrow \quad \beta \leq 1 \leq \frac{1}{\beta}$$

which is true for all $0 < \beta \leq 1$. Hence the even-aged orchard is a solution to the two-age-class orchard management problem. □

A.3.3 Proof of proposition 3.2

Proof of proposition 3.2 (page 18). Define the function $g(.)$ as

$$\begin{aligned} g(\phi;\ \beta, f_y, f_o, L) &= \frac{u'(c(\frac{L}{2} - \phi))}{u'(c(\frac{L}{2} + \phi))} - \beta = 0 \\ &= \frac{u'(\frac{L}{2}(f_y + f_o) + (f_o - f_y)\phi)}{u'(\frac{L}{2}(f_y + f_o) - (f_o - f_y)\phi)} - \beta = 0 \end{aligned}$$

Using the implicit function theorem, the partial derivative of ϕ with respect to any parameter α is

$$\frac{\partial \phi}{\partial \alpha} = \frac{-\frac{\partial g}{\partial \alpha}}{\frac{\partial g}{\partial \phi}}$$

Let $c(x_y) = f_y\, x_y + f_o\,(L - x_y)$. First, compute the denominator

$$\frac{\partial g}{\partial \phi} = \frac{(f_o - f_y)\left[u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi)) + u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi))\right]}{u'(c(\frac{L}{2}+\phi))^2}$$

$\frac{\partial g}{\partial \phi} < 0$ since $u''(.) < 0$. So $\text{sign}(\frac{\partial \phi}{\partial \alpha}) = \text{sign}(\frac{\partial g}{\partial \alpha})$.

Computing the sign of $\frac{\partial \phi}{\partial \beta}$ gives

$$\text{sign}(\frac{\partial \phi}{\partial \beta}) = \text{sign}(-1) \quad (< 0)$$

Computing $\frac{\partial \phi}{\partial f_y}$ gives

$$\frac{\partial \phi}{\partial f_y} = \frac{(\frac{L}{2}-\phi)u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi)) - (\frac{L}{2}+\phi)u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi))}{u'(c(\frac{L}{2}+\phi))^2}$$

$$\text{sign}\left(\frac{\partial \phi}{\partial f_y}\right) > 0 \Leftrightarrow (\frac{L}{2}-\phi)u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi)) - (\frac{L}{2}+\phi)u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi)) > 0$$

$$u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi)) > \frac{(\frac{L}{2}+\phi)}{(\frac{L}{2}-\phi)}u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi))$$

$$\frac{u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi))}{u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi))} < \frac{(\frac{L}{2}+\phi)}{(\frac{L}{2}-\phi)} \qquad \text{(Since } u''(.) < 0\text{)}$$

$$\frac{A(c(\frac{L}{2}+\phi))}{A(c(\frac{L}{2}-\phi))} < \frac{(\frac{L}{2}+\phi)}{(\frac{L}{2}-\phi)}$$

Where $A(c) = \frac{-u''(c)}{u'(c)}$ measures the preference for consumption smoothing.

Computing $\frac{\partial \phi}{\partial f_o}$ gives

$$\frac{\partial \phi}{\partial f_o} = \frac{(\frac{L}{2}+\phi)u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi)) - (\frac{L}{2}-\phi)u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi))}{u'(c(\frac{L}{2}+\phi))^2}$$

$$\begin{aligned}
\text{sign}\left(\frac{\partial \phi}{\partial f_o}\right) > 0 \Leftrightarrow & (\frac{L}{2}+\phi)u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi)) - (\frac{L}{2}-\phi)u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi)) > 0 \\
& u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi)) > \frac{(\frac{L}{2}-\phi)}{(\frac{L}{2}+\phi)}u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi)) \\
& \frac{u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi))}{u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi))} < \frac{(\frac{L}{2}-\phi)}{(\frac{L}{2}+\phi)} \qquad \text{(Since } u''(.) < 0\text{)}
\end{aligned}$$

$$\frac{A(c(\frac{L}{2}+\phi))}{A(c(\frac{L}{2}-\phi))} < \frac{(\frac{L}{2}-\phi)}{(\frac{L}{2}+\phi)}$$

Where $A(c) = \frac{-u''(c)}{u'(c)}$ measures the preference for consumption smoothing.

Computing $\frac{\partial \phi}{\partial L}$ gives

$$\frac{\partial \phi}{\partial L} = \frac{(f_y + f_o)\left(u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi)) - u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi))\right)}{2u'(c(\frac{L}{2}+\phi))^2}$$

$$\begin{aligned}
\text{sign}\left(\frac{\partial \phi}{\partial L}\right) > 0 \Leftrightarrow & u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi)) - u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi)) > 0 \\
& u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi)) > u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi)) \\
& \frac{u'(c(\frac{L}{2}+\phi))u''(c(\frac{L}{2}-\phi))}{u'(c(\frac{L}{2}-\phi))u''(c(\frac{L}{2}+\phi))} < 1 \qquad \text{(Since } u''(.) < 0\text{)}
\end{aligned}$$

$$\frac{A(c(\frac{L}{2}+\phi))}{A(c(\frac{L}{2}-\phi))} > 1$$

□

A.3.4 Proof of proposition 3.3

Proof of proposition 3.3 (page 20). Define ϕ, the maximum amplitude of the cycle region, as in equation (4). Assume that the land constraint binds every period, so we can define a land allocation as $\mathbf{x}_t = (L - x_{o,t}, x_{o,t})$, where $x_{o,t}$ is the area of land allocated to old trees in period t. Let $x_{o,t+1} = P(x_{o,t})$ be the function that returns the optimal area of old trees in period $t+1$ given the area of old trees in period t. Assume that it exists and is continuous.

We will consider the optimal transition for three regions in $x_{o,t} \in [0, L]$. Let region one be $[0, \frac{L}{2} - \phi)$, region two be $[\frac{L}{2} - \phi, \frac{L}{2} + \phi]$, and region three be $(\frac{L}{2} + \phi, L]$. We will construct the optimal transition function piecewise across these three regions

$$P(x_{o,t}) = \begin{cases} P_1(x_{o,t}) & \text{for } x_{o,t} \in [0, \frac{L}{2} - \phi) \\ P_2(x_{o,t}) & \text{for } x_{o,t} \in [\frac{L}{2} - \phi, \frac{L}{2} + \phi] \\ P_3(x_{o,t}) & \text{for } x_{o,t} \in (\frac{L}{2} + \phi, L] \end{cases}$$

We know from proposition 3.1 that $P_2(x_{o,t}) = L - x_{o,t}$. We will begin by showing that $P_3(x_{o,t}) = L - x_{o,t}$ as well, and then use this result to show that $P_1(x_{o,t}) = \frac{L}{2} - \phi$, thus finding the piecewise definition of $P(x_{o,t})$.

The optimal transition function will depend on the optimal value of the aging constraint multiplier, λ_t. If $\lambda_t > 0$, then the aging constraint binds, and $x_{o,t+1} = L - x_{o,t}$. On the other hand, if $\lambda_t = 0$, then $x_{o,t+1} \leq L - x_{o,t}$ and we will need to pin down its value.

We will show that for $x_{o,t}$ in region three, starting with $\lambda_t = 0$ implies there is no solution to the KKTs in period $t+1$. Therefore $\lambda_t > 0$ for $x_{o,t}$ in region three.

Recall the Euler equation

$$\lambda_t = \beta(u'(c_{t+1})f_o - u'(c_{t+1})f_y - \lambda_{t+1}))$$

Assume $x_{o,t} \in (\frac{L}{2} + \phi, L]$ and $\lambda_t = 0$. Therefore $\lambda_{t+1} = \beta u'(c_{t+1})(f_o - f_y) > 0$. Iterating the Euler

equation and using this value for λ_{t+1} gives an expression for λ_{t+2}

$$\lambda_{t+2} = (f_o - f_y)\left(u'(c_{t+2}) - \frac{1}{\beta}u'(c_{t+1})\right)$$

We must have $\lambda_{t+2} \geq 0$ to satisfy the KKT conditions.

$$\begin{aligned} & \lambda_{t+2} \geq 0 \\ \Rightarrow \quad & u'(c_{t+2}) \geq \frac{1}{\beta}u'(c_{t+1}) \\ \Rightarrow \quad & x_{o,t+2} < x_{o,t+1} < x_{o,t} \qquad \text{(with strict inequality because } \beta < 1\text{)} \end{aligned}$$

This is incompatible with $\lambda_t = 0$ and $\lambda_{t+1} > 0$, which imply that $x_{o,t+2} = L - x_{o,t+1} \geq x_{o,t}$. Therefore $x_{o,t} \in (\frac{L}{2} + \phi, L]$ implies $\lambda_t > 0$, and $P_3(x_{o,t}) = L - x_{o,t}$.

We now turn to characterizing $P_1(x_{o,t})$. For $x_{o,t}$ in region one, we begin by showing that assuming $\lambda_t > 0$ implies there is no solution to the KKTs in period $t+1$. We then find the optimal transition for $x_{o,t+1} \leq L - x_{o,t}$ given $\lambda_t = 0$.

For $x_{o,t}$ in region one, if $\lambda_t > 0$, then $x_{o,t+1} = L - x_{o,t}$, which is in region three. From our previous result, $x_{o,t+2} = P_3(x_{o,t+1}) = x_{o,t}$, which implies a cycle. However, by proposition 1, the only cycles consistent with the KKTs are those with $x_{o,t}, x_{o,t+1} \in [\frac{L}{2} - \phi, \frac{L}{2} + \phi]$. Therefore, $\lambda_t = 0$ in region one.

Finally, for $x_{o,t}$ in region one, let $\lambda_t = 0$, which implies $x_{o,t+1} \leq L - x_{o,t} \in$ (region three), so $\lambda_{t+1} > 0$. Further, from before, $x_{o,t+1}$ in region three, implies that $\lambda_{t+2} = 0$. Using the Euler

equation and iterating by one period we get an expression for λ_{t+2}:

$$
\begin{aligned}
& (f_o - f_y)\left(u'(c_{t+2}) - \frac{1}{\beta}u'(c_{t+1})\right) = 0 \\
\Rightarrow \quad & \frac{u'(c_{t+1})}{u'(c_{t+2})} = \beta \\
\Rightarrow \quad & c_{t+1} = (f_y(\frac{L}{2} - \phi) + f_o(\frac{L}{2} + \phi)) \\
\Rightarrow \quad & x_{o,t+1} = \frac{L}{2} + \phi
\end{aligned}
$$

So the only allocation of old trees in period $t+1$ consistent with the KKTs when $x_{o,t}$ is in region one is $x_{o,t+1} = \frac{L}{2} + \phi$. Hence $P_1(x_{o,t}) = \frac{L}{2} + \phi$, a constant. The land allocated to old trees in period $t+1$ is independent of the allocation of old trees in period t, so long as it is in region one. Any allocation in region one moves to the upper boundary of region two in the next period, and then remains in region two thenceforth.

The optimal transition rule is thus

$$
P(x_{o,t}) = \begin{cases} \frac{L}{2} + \phi & \text{for } x_{o,t} \in [0, \frac{L}{2} - \phi) \\ L - x_{o,t} & \text{for } x_{o,t} \in [\frac{L}{2} - \phi, \frac{L}{2} + \phi] \\ L - x_{o,t} & \text{for } x_{o,t} \in (\frac{L}{2} + \phi, L] \end{cases}
$$

as shown in figure 5. □

A.4 Notation used in the three-age-class model

$u(\mathbf{x})$ (also $u(c_t)$): Instantaneous utility of allocation $\mathbf{x}$. (Section 4, page 23)

$\mathbf{x}_t$: Allocation vector at time t; (x_{yt}, x_{mt}, x_{ot}). (Section 4, page 23)

$V(\mathbf{x}_t)$: Infinite-horizon value function evaluated at allocation $\mathbf{x}_t$. (Section 4, page 23)

x_{yt}: Area of young trees in period t. (Section 4, page 23)

x_{mt}: Area of mature trees in period t. (Section 4, page 23)

x_{ot}: Area of old trees in period t. (Section 4, page 23)

β: Single-period discount factor. (Section 4, page 23)

L: Total land area (normalised to 1). (Section 4, page 23)

c_t: Period-t harvest/consumption; $c_t := f_y x_{yt} + f_m x_{mt} + f_o x_{ot}$. (Section 4, page 23)

f_i ($i \in \{y, m, o\}$): Yield per unit land for age-class i. (Section 4, page 23)

$\mathbf{f} := (f_y, f_m, f_o)$: Yield vector. (Section **??**, page **??**)

λ_{1t}: Multiplier on $x_{m,t+1} \leq x_{yt}$. (Section 4, page 23)

λ_{2t}: Multiplier on $x_{o,t+1} \leq x_{mt}$. (Section 4, page 23)

ψ_t: Multiplier on the total-land constraint. (Section 4, page 23)

s_{it}: Multiplier on $x_{it} \geq -0$. (Section 4, page 23)

$\mathcal{F}(x_y, x_m, x_o) := (x_o, x_y, x_m)$: Pure biological ageing map; $\mathcal{F}^k$ is its k^{th} iterate. (Section **??**, page **??**)

Δ^2: Feasible-allocation simplex $\{(x_y, x_m, x_o) : x_y + x_m + x_o = 1,\ x_i \geq 0\}$. (Section **??**, page **??**)

i^*: Faustmann optimal replacement age for a single tree (equal to o here). (Section **??**, page **??**)

$\bar{\mathbf{x}} := (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$: Balanced allocation (one-third in each class). (Section **??**, page **??**)

$\mathcal{U} := \{\mathbf{x} \in \Delta^2 : \|\mathbf{x} - \bar{\mathbf{x}}\| < \varepsilon\}$: Neighbourhood of $\bar{\mathbf{x}}$ yielding local 3-cycles. (Section **??**, page **??**)

Δ_1, Δ_2: Yield-difference combinations used to sign λ_1 and λ_2. (Section **??**, page **??**)

K_0: Cycle region—allocations whose optimal path is a stationary three-period cycle. (Section **??**, page **??**)

$\Gamma(\mathbf{x}_t)$: Full feasible transition set from $\mathbf{x}_t$. (Section **??**, page **??**)

$F(\mathbf{x}_t, \mathbf{x}_{t+1}) := u(\mathbf{x}_t) + \beta u(\mathbf{x}_{t+1})$: Two-period (stage-return) payoff. (Section **??**, page **??**)

$\Gamma^r(\mathbf{x}_t)$: Restricted transition set when only mature trees may be replaced. (Section 2, page 34)

$\mathcal{M}(x_m) := \{\mathbf{x} \in \Delta^2 : 0 \leq x_o \leq 1 - x_m\}$: Iso-mature line at level x_m. (Section 3, page 34)

$R(x_m)$: Optimal old-tree share conditional on x_m; solves the ridge FOCs. (Section **??**, page **??**)

ψ_1, ψ_2: Complementary-slackness multipliers appearing in the ridge conditions. (Section **??**, page **??**)

$\mathbf{x}^R(x_m) := (x_m, R(x_m))$: Point on the ridge with mature share x_m. (Section **??**, page **??**)

$\mathcal{R}$: Entire ridge $\{\mathbf{x}^R(x_m) :\leq x_m \leq 1\}$. (Section 3, page 35)

$\mathcal{R}_1, \mathcal{R}_2, \mathcal{R}_3, \mathcal{R}_4$: Four mutually exclusive sub-segments of $\mathcal{R}$. (Section 4.3.5, page 47)

$\mathcal{R}_1$: The set of allocations on the ridge of the ridge whose mature elements are between $\tilde{x}_m$ and 1, formally, $\{\mathbf{x} \in \mathcal{R} : x_m \in (\tilde{x}_m, 1]\}$.

$\mathcal{R}_2$: The set of allocations on the ridge whose mature elements are between $\bar{x}_m$ and $\tilde{x}_m$, formally, $\{\mathbf{x} \in \mathcal{R} : x_m \in (\bar{x}_m, \tilde{x}_m]\}$.

$\mathcal{R}_3$: The set of allocations on the ridge whose mature elements are between $\bar{r}$ and $\bar{x}_m$, formally, $\{\mathbf{x} \in \mathcal{R} : x_m \in [R(\bar{x}_m), \bar{x}_m]\}$.

$\mathcal{R}_4$: The set of allocations on the ridge whose mature elements are between 0 and $\bar{r}$, formally, $\{\mathbf{x} \in \mathcal{R} : x_m \in [0, R(\bar{x}_m))\}$.

$\bar{\mathbf{x}}^c := (\bar{x}_m, \bar{r})$: Ridge point with that maximum mature share. (Section 4.3.6, page 51)

$\bar{x}_m := \max\{x_m : (x_m, x_o) \in K_0\}$: Maximum mature share inside K_0. (Section **??**, page **??**)

$\tilde{\mathbf{x}} := (\tilde{x}_m, R(\tilde{x}_m))$: Point where ridge meets $x_o = 0$. (Section 11, page 47)

$\tilde{x}_m$: Smallest x_m with $R(\tilde{x}_m) = 0$ (ridge touches $x_o = 0$). (Section 4.3.7, page 55)

K_i $(i \geq 1)$: Sets that can reach K_{i-1} in one step and are disjoint from earlier K_j. (Section 6, page 44)

$\text{bd}(K_{ij}) := \overline{K}_i \cap \overline{K}_j$: Boundary between K_i and K_j. (Section 4.3.6, page 49)

$\boldsymbol{\xi}(\mathbf{x})$: Optimal next-period allocation chosen from $\Gamma^r(\mathbf{x})$. (Section 16, page 37)

$\boldsymbol{\xi}m(\mathbf{x})$: Mature-tree component of $\boldsymbol{\xi}(\mathbf{x})$. (Section 13, page 51)

$\boldsymbol{\xi}^k(\mathbf{x})$: Allocation after k optimal transitions; $\xi^3(\mathbf{x}_t)$ and $\xi^3_m(\mathbf{x}_t)$ are the full allocation and its mature component after three steps. (Section 21, page 57)

N_1, N_2: Integer bounds on convergence time from $\mathcal{R}_2$ and from $\mathcal{R}_4 \cap K_5$. (Section 4.2, page 32)

$\delta^c, \delta_1, \delta_2, \delta^d$: Positive constants introduced in convergence lemmas. (Section 12, page 50)

$q(\mathbf{x}) := \left\lceil \dfrac{x_m - \bar{x}_m}{\delta_1} \right\rceil$: Integer distance of $\mathbf{x} \in \mathcal{R}_2$ above $\bar{x}_m$. (Section 4.3.6, page 52)