

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

The Dynamics of Perennial Crop Production and Processing

### Permalink

<https://escholarship.org/uc/item/70g9r15x>

### Author

Tregeagle, Daniel Trevellan

### Publication Date

2017

Peer reviewed|Thesis/dissertation

# **The Dynamics of Perennial Crop Production and Processing**

by

Daniel Trevellan Tregeagle

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David Zilberman, Chair

Professor Larry Karp

Professor David Romer

Professor Leo Simon

Spring 2017

# **The Dynamics of Perennial Crop Production and Processing**

Copyright 2017  
by  
Daniel Trevellan Tregeagle

## Abstract

The Dynamics of Perennial Crop Production and Processing

by

Daniel Trevellan Tregeagle

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor David Zilberman, Chair

Perennial crops are plants, used in agriculture, that can be harvested multiple times before replanting. They are beneficial in a variety of ways, producing food and fuel, and providing agronomic, environmental, and cultural benefits. Economic analysis of the production of perennials is complicated by the long life of these crops, and by the changing pattern of productivity over the crop's lifespan. This dissertation contains three studies looking at the implications of these two facts on perennial crop production and processing. The first chapter introduces perennial crops and the history of agricultural economists' attempts to study these issues. Chapters 2 and 3 analyze perennial crop production and processing in the context of biofuel production in Brazil. Chapter 4 presents the foundation of a unifying framework that can overcome limitations in the preceding two chapters.

Brazilian sugarcane production growth, being a key feedstock to sugar and ethanol production, exhibited a puzzling slow-down in the decade starting in 2010. In chapter 2 I investigate a noted, but unexplored, mechanism to explain this slow-down: the link between credit availability, sugarcane replanting, and sugarcane yield. Using secondary sources, I establish the plausibility of the conjecture that credit restrictions affected sugarcane farmer replanting decisions. To establish the link between replanting and yield, I develop a formal model to analyze the dynamics of yield after a change in the replant rate. I test this model econometrically using data from the South-Central region of Brazil, finding evidence mostly consistent with the model. The model is able to explain around one third of the variation in sugarcane yields, implying that, while this channel is important, it alone cannot explain the production slow-down.

Perennial crop field-biorefinery supply chains are necessary in the production of many agricultural products. In particular, the low-cost production of low-carbon biofuels, such as ethanol from sugarcane or cellulosic feedstocks, relies on minimizing the costs of production along a perennial crop field-biorefinery supply chain. In chapter 3 I develop and analyze an unexplored mechanism to reduce perennial crop field-biorefinery supply chain costs: adjusting the age-structure, and hence yield, of the perennial feedstock. I present comparative statics of this model, finding that smaller biorefineries are most likely to benefit from age-

structure endogenization. However, the results from a simulation of this model, calibrated to the sugarcane ethanol industry in the South-Central region of Brazil and comparing the cost-minimizing to the yield-maximizing age-structure, show that the cost-reductions from endogenization are small in this case (less than 1 percent cost reduction). Generally, the magnitude of the cost-reduction will depend on the growing pattern of the crop, the costs of growing and transporting the feedstock, and the reference age-structure.

In chapter 4 I adapt the theoretical framework of Mitra et al. (1991) to a two-age-class, finite horizon model, recasting their growth theoretic model into a form better suited for policy analysis. After introducing a simplified version of the model, and its general case, I determine necessary conditions on key parameters (relative productivity of mature trees, opportunity cost of land, and farmer patience) for each qualitative trajectory type in an arbitrary period. I use these conditions to develop a proposition about planting in the final period, and to analyze an example trajectory. Finally, I propose ways that this model can be used to answer several questions about perennial crop management.

# Contents

<b>Contents</b>	<b>i</b>
<b>List of Figures</b>	<b>iii</b>
<b>List of Tables</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 The Yield Dynamics of Perennial Crops: An Application to Sugarcane in Brazil</b>	<b>7</b>
2.1 Introduction . . . . .	7
2.2 Literature Review . . . . .	9
2.3 The Brazilian Sugarcane Industry . . . . .	11
2.4 Decomposing Sugarcane Production into Area and Yield Effects . . . . .	12
2.5 The 2008-09 Financial Crisis is a Prime Candidate for Explaining the Yield Variation . . . . .	16
2.6 Crop Renewal: The Link Between Credit and Yield . . . . .	19
2.7 The Yield Trajectory after a Change in the Replant Rate . . . . .	20
2.8 Empirical Strategy . . . . .	27
2.9 Data . . . . .	29
2.10 Results . . . . .	31
2.11 Discussion . . . . .	33
2.12 Conclusion . . . . .	36
<b>3 Optimal Balanced Age-Structure in a Perennial Crop Field-Refinery Supply Chain</b>	<b>39</b>
3.1 Introduction . . . . .	39
3.2 An Analytical Model of Orchard Age, Growing Region Area, and Biorefinery Capacity . . . . .	41
3.3 Cost Minimization and Simulation . . . . .	47
3.4 Discussion and Future Research . . . . .	55

3.5	Conclusion . . . . .	57
<b>4</b>	<b>The Economics of Perennial Orchards with Endogenous Age-Classes</b>	<b>59</b>
4.1	Introduction . . . . .	59
4.2	Two-Period Models . . . . .	64
4.3	A $T$ -Period Model . . . . .	68
4.4	Discussion and Future Work . . . . .	74
	<b>Bibliography</b>	<b>79</b>
<b>A</b>	<b>Appendix to Chapter 2</b>	<b>88</b>
A.1	Formulae for Calculating the Yield Change following a Discrete Change in the Replant Rate that Changes the Number of Active Age-Classes . . . . .	88
<b>B</b>	<b>Appendix to Chapter 3</b>	<b>90</b>
A.1	Proofs of Propositions and Lemmas . . . . .	90
A.2	Calibration of the Cost-Minimization Problem . . . . .	101
<b>C</b>	<b>Appendix to Chapter 4</b>	<b>105</b>
A.1	Lemmas and Proofs . . . . .	105

# List of Figures

2.1	Map of sugarcane production regions in Brazil (from UNICA (2015)) . . . . .	12
2.2	Sugarcane production, planted area, and yield series for Brazil from 1990 to 2014. . . . .	13
2.3	The three effects produced by the modified Venegas and Ruttan decomposition. Area $A$ represents the area effect; area $B$ represents the yield effect; and area $C$ represents the mixed effect . . . . .	14
2.4	Decomposition of yearly changes in sugarcane production into area driven changes, yield driven changes, and mixed effect changes (using equation 2.1). In 2004–05 there was minimal production change since an increase in area almost exactly offset a decline in yield. . . . .	15
2.5	Age yield relationship taken from Margarido and Santos (2012). Freshly planted canes provide no yield (year 0). . . . .	21
2.6	Diagrammatic representation of the dynamics of the quantity of land in each age-class. Each age-class is populated only from the previous age-class, except for the youngest age-class which collects all the land replanted in a particular year. Land in the oldest age-class stays in the oldest age-class unless replanted. . . . .	23
2.7	The change in yield $t$ years after a 1 percentage point increase in the replant rate. Graph generated using the São Paulo age-yield relationship from figure 2.5. . . . .	27
2.8	Average yield and replant rate across the 30 mesoregions in the sample. . . . .	31
2.9	Two views of the coefficients of the 6 models. Figure 2.9a shows the coefficient estimates relative to zero. Figure 2.9b shows the coefficient estimates relative to the predictions from figure 2.7b. . . . .	37
2.10	Actual and predicted yields for the 30 South-Central mesoregions in my sample using the 4-lag variant of the regression model. . . . .	38
3.1	Yield is increasing in $n$ while the marginal age class, $n$ , is more productive than the average of the $n$ -orchard, $y(n)$ . Here the age-yield relationship is a Hoerl function fitted to the Brazilian yield data from Margarido and Santos (2012). . . . .	44
3.2	Example isocost and isoquant curve for the cost minimization model calibrated to São Paulo, Brazil. Details of calibration are in appendix A.2. Note that the lower-cost set is non-convex—there is a local maximum of the isocost curve around $n = 1$ and a local minimum around $n = 4$ . . . . .	49



3.3	Cost minimizing orchard age and planted land area as biorefinery capacity is increased from 1 million tons to 36 million tons. $n_{MY} = 5.31$ . Details of calibration are in appendix A.2 . . . . .	52
3.4	Details of calibration are in appendix A.2 . . . . .	54
3.5	Truncated graph of cumulative percentage cost reductions from 100 000 draws from a uniform distribution with support described in table 3.2. The result from the simulation using the parameters calculated in appendix A.2 is shown with the dashed, red line. Note that the y-axis begins at 0.96, and approximately 1.3 percent of the results had a cost reduction of greater than 2 percent. . . . .	56
4.1	Binding and non-binding aging constraints with $f_m > 1$ . The horizontal axis is the land allocated to perennials, $x$ , and the vertical axis measures marginal benefit. Here I assume linear marginal benefit $u(f_n x_n) = b - a(f_n x_n)$ , where $a = 1$ and $b = 4$ . Also $\beta = 0.5$ . . . . .	76
4.2	Binding and non-binding aging constraints with $f_m < 1$ . The horizontal axis is the land allocated to perennials, $x$ , and the vertical axis measures marginal benefit. Here I assume linear marginal benefit $u(f_n x_n) = b - a(f_n x_n)$ , where $a = 1$ and $b = 4$ . Also $\beta = 0.5$ . . . . .	77
4.3	Four example trajectories generated by changing the time horizon and opportunity cost of land. Apart from price, the parameter values are the same as those used to generate figure 4.1, that is $u(f_n x_n) = b - a(f_n x_n)$ , with $a = 1$ and $b = 4$ , $f_m = 2$ , $\beta = 0.5$ . Additionally, $x_{y0} = 0$ and $S'(x_{m,T+1}) = 0$ . . . . .	78
A.1	The Hoerl function provides a poor fit for the observed sugarcane age-yield relationship. Fitting the Hoerl function to a modified dataset (increasing the observation for one-year-old canes) greatly improves the fit. I use the Hoerl function fitted to the adjusted data for the simulations in this study. . . . .	102

# List of Tables

2.1	Results from estimating equation (2.3) with the Brazilian dataset using 0–5 lags of replant rate . . . . .	32
3.1	Signs of comparative statics of $n^*$ and $L^*$ with respect to four key parameters. .	53
3.2	Support for cost minimization with random parameters . . . . .	55
4.1	Conditional signs of comparative statics of $x^*$ with respect to key parameters. .	66
4.2	The most relaxed possibility for each case in lemma C.1. The order of possibilities is Interior $\Rightarrow$ Boundary $\Rightarrow$ Origin $\Rightarrow$ Infeasible. Each element of this list is a superset of the elements to the right, but not vice versa. Row $i$ and column $j$ refers to case $(i,j)$ in lemma C.1 . . . . .	71
4.3	A table of possibilities, testing the consistency of the solution to the KKTs against the maintained assumptions on $\lambda$ 's and $\eta$ 's, showing the values the land allocation variables may take in each case. Row $i$ and column $j$ refers to case $(i,j)$ in lemma C.1 . . . . .	71
A.1	Parameter values used for cost minimization simulation . . . . .	101

I dedicate this dissertation to my parents, Anne and Peter. You've always been there for me.

## Acknowledgments

Thanks to the members of my dissertation committee for their service.

Thanks to my dissertation chair, David Zilberman, for encouragement, ideas, feedback, support, and discussion.

For assisting in the development of chapter 2 I thank Larry Karp, David Romer, Leo Simon, David Zilberman, and the participants of: the 2015 Giannini Agricultural and Resource Economics Graduate Student Conference, 2015 International Consortium of Applied Bioeconomy Researchers Conference, the 2016 Joint Western Agricultural Economics Association and Canadian Agricultural Economics Association Conference, and the 2016 Agricultural and Applied Economics Association.

For assisting in the development of chapter 3 I thank Peter Berck, Larry Karp, Seth Garz, David Romer, Emma Roos-Collins, Leo Simon, Justus Wesseler, Yang Xie, David Zilberman, and the participants of the 20th International Consortium of Applied Bioeconomy Research Conference for helpful comments and suggestions. Thanks to Leo Simon for sharing code for random parameter draws.

For assisting in the development of chapter 4 I thank thank Larry Karp, Gordon Rausser, Leo Simon, David Zilberman, and the participants of the UC Berkeley Ag Policy Seminar

Thanks to Michael Harris for encouraging my ambition to study for a PhD in the US. Thanks to Jim Wilen for encouraging my interest in dynamic optimization, and supporting my applications to study in the US.

Thanks to my officemates Yang Xie, Becca Taylor, Seth Garz, Jimmy Lu, and Rita Du.

Thanks to Richard and Margit Roos-Collins, my “Berkeley parents”.

And finally, thanks to Emma Roos-Collins, whose love, patience, support, and organization helped me bring this dissertation to fruition.

# Chapter 1

## Introduction

Perennial crops are plants, used in agriculture, that can be harvested multiple times before replanting. In contrast, annual crops can be harvested only once. There are many types of perennial crop. Fruit trees (for fresh and processed fruit, or juice) are a salient example, producing the type of crop one might most easily associate with perennial crops. Perennials, however, provide a much wider range of crops. There are fruit trees whose fruit is used for commodity crops, such as cocoa and coffee. Perennial crops also include perennial grasses, used for food, such as sugarcane, or for biofuel, such as sugarcane, *miscanthus*, or switchgrass. Perennial crops may also include plants used for decoration, like roses and other flowering shrubs.

An alternative way to conceptualize perennial crops is as ‘not-annual’ crops, which emphasizes the great variety in lifespans and number of harvests exhibited by perennial crops. Perennials may be grown for as few as three years, as is the case for strawberries in the New England region of the United States (Handley, 2011), or as many as 3000 years. According to Rackham and Moody’s book *The Making of the Cretan Landscape* (cited in Riley (2002)) the Monumental Olive Tree of Vouves, in Crete, has been dated between 2000–4000 years old, and is still bearing olives. There is even great variety in bearing lifespan within a single species. Sugarcane, for example, in Kenya is commonly grown for the first two harvests only, while in Swaziland over 20 harvests have been recorded on the best soils (James, 2004).

Devadoss and Luckstead (2010), building on the enumeration by French and Matthews (1971), identify four key differences between annual and perennial crops:

- (1) perennial crop supply is vastly different from the annual crop supply because
- (2) trees are a long-term investment;
- (3) perennial crops have long gestation intervals between initial plantings and first harvest;
- (4) once trees start yielding, there is an extended period of productivity, and then gradual decline in production; and
- (5) after trees reach their final productivity decline, they are removed.

This highlights a crucial component to understanding the economics of perennial crops, the *age-yield relationship*, which is the yield of the crop as a function of the number of years since its planting. The general shape of the age-yield relationship has three phases: the

establishment phase (weakly increasing yield), the peak phase (constant, maximal yield), and the senescence phase (weakly decreasing yield) (Mitra et al., 1991). Depending on the crop, age-yield relationships can be more complicated, due to the presence of alternate-bearing behavior of the crop, or to interactions between different age-classes of trees (e.g. old trees casting shade on young trees).

Perennial crops provide substantial value to humanity in a variety of different ways. First, they provide food. In the US, the value of production of fruit and tree nuts in 2014 was 30 billion USD, accounting for around 15 percent of the gross value of US agricultural crop production (USDA, 2016b,a). Tree nut value was around \$10b, led by almonds, walnuts, and pistachios (USDA, 2016a). The fruit production value was led by grapes, apples, strawberries<sup>1</sup>, and oranges.

Second, perennial crops provide fuel. Much contemporary work in perennials is being done in the biofuels context (see, for example, the work of the Energy Biosciences Institute, and chapters 2 and 3).

Third, perennial crops provide agronomic benefits relative to annuals. “Compared with annual counterparts, perennial crops tend to have longer growing seasons and deeper rooting depths, and they intercept, retain, and utilize more precipitation... Perennial crops require fewer passes of farm equipment and less fertilizer and herbicide” (Glover et al., 2010; Wallace, 2000, p. 1638). Perennials and annuals can be mixed together to extract even more water, the different root depths minimizing interference between crop types.

Fourth, perennial crops can provide valuable ecological and environmental services. “Greater root mass reduces erosion risks and maintains more soil carbon compared with annual crops” (Glover et al., 2010, p. 1638). The erosion control benefits of perennial crops have been known for some time. In his 1929 book *Tree Crops*, Smith advocated using perennial tree crops as sources of food for humans and feed for animals, especially on hilly regions prone to erosion, thereby providing not only sustenance for humans and animals, but also agronomic benefits from improved future soils, and aesthetic benefits too. Smith’s ideas were not well accepted at the time, and he was pessimistic about the future prospects of his ideas, partly because he was advocating *new* crop varieties: “And Secretary of Agriculture Jardine told me that his department has no time for such new things as honey locusts, that they were busy with the bugs and bites and blights of crops already established. Such is the scientific side of this democracy!” (Smith, 1929, p. 95). His pessimism proved accurate: “For a variety of reasons, though, his [Smith’s] ideas were not implemented to any great degree” (Molnar et al., 2013, p. 457). However, the recent research efforts of Glover et al. (2010) to commercialize perennial varieties of widely accepted grain crops may overcome some of the resistance to widespread use of perennial crops in commodity agriculture. “With similar commitments [to biofuel perennials] for developing food-producing perennial grains, we estimate that commercially viable perennial grain crops could be available within 20 years” (Glover et al., 2010, p. 1639).

---

<sup>1</sup>Strawberries are grown as an annual crop in California (California Strawberry Commission, 2017), but are often grown as a perennial in colder climates. The grapes, apples, and oranges are perennials.

Finally, perennial crops provide cultural value, having been cultivated by humans since early in our history. For example, humans have been cultivating olives since at least the early bronze age (5000–6000 years ago) (Vossen, 2007), and of the biblical “seven species”—wheat, barley, grapes, figs, pomegranates, olives, and date honey—five are perennial crops (Berry et al., 2011). Additionally, apples have been intertwined with the history of European settlement in the United States. Johnny Appleseed exploited the non-bearing period of apples, by running ahead of the frontier and planting apple trees. By the time the settlers arrived, he was ready to sell them apple trees, providing food and a feedstock for alcohol to the new arrivals (Pollan, 2001).

Perennial crops are interesting from a methodological perspective because perennials are durable assets, complicating the normative and positive analysis of their production. “Specifically, lags and dynamic processes appear to be at the heart of understanding large-animal livestock and perennial crop production problems.” (Just and Pope, 2001, p. 706).

Agricultural economists have been writing about perennial crop supply issues since at least the 1960’s. The majority of these studies have used positive, econometric methods for their analysis. Far fewer have used numerical or theoretical analyses. The modern framework used in the econometrics of perennial crop supply can be traced to a paper by French and Bressler (1962), who introduced separate planting and removal relationships for trees, a conceptual distinction followed (and sometimes expanded upon) by subsequent studies. They applied this framework to lemon trees in California, finding evidence of cyclicity in production and prices. French and Matthews (1971) expanded on the planting and removals framework by generalizing the new plantings and acreage equations, and introducing rational expectations, applying their framework to US asparagus production. Rausser (1971) developed a theoretical foundation for the estimation of perennial crop supply response, and estimated it using data from orange growers in California and Arizona. A particular novelty of Rausser’s approach was the incorporation of the influences of tax and ownership structures on growers’ behavior.

In the 1980’s the two papers of Bellman and Hartley (1985) and Trivedi (1986) outlined a general, dynamic programming approach to perennial crop production, carefully specifying the perennial crop production process. These frameworks were designed to provide guidance for econometric studies, or to be solved computationally. Using numerical simulation Knapp (1987) found that price, production, and consumption of alfalfa in California converge to the long-run equilibrium in 1-2 years (4 years in the most extreme case), but acreage exhibited persistent cycles.

Knapp and Konyar (1991) and Kalaitzandonakes and Shonkwiler (1992) used a Kalman filter to recover estimates of the unobserved age-structure of a perennial crop. Knapp and Konyar (1991) studied alfalfa in California, and Kalaitzandonakes and Shonkwiler (1992) studied grapefruit in Florida. Knapp and Konyar (1991) found that under naive expectations, convergence to long run equilibrium occurs after 4-5 years. Under quasi-rational expectations, there is convergence to equilibrium too, but it takes longer (10-20) years, suggesting that the cyclical behavior of earlier studies may be due to short-run shocks, rather than an inherent characteristic of the industry.

Burger and Smit (1997) provided a literature survey of production and supply studies of

perennial crops. Further, they developed a simple theory for modeling the replanting decision for perennials, based on the net present value of a single tree. They provide an application of their model to find the minimum optimal replanting age for rubber trees in India. Wessler (1997) also provides a method for calculating the value of a single tree, or even-aged orchard, showing a formula for value starting from an arbitrary age.

Gotsch and Wohlgenant (2001) presented a framework for simulating the expected welfare effects of a technical change in a perennial crop with an application to Cocoa in Malaysia. They observed that the standard welfare economics framework needs to be adjusted for perennials, incorporating: dynamic supply response to prices; biological lags in production; and vintage effects to account for variation in the productivity of the tree by age.

Devadoss and Luckstead (2010) used the French and Matthews (1971) model to estimate the apple supply response function for Washington state. Brady and Marsh (2013) used owner-field level data from Washington state to determine whether changes in perennial supply are largely driven by changes in the intensive or the extensive margin. They introduced entry and exit to the econometric model of perennial crop supply, allowing a farmer to completely exit perennial production, rather than just make marginal adjustments to an existing acreage. Feinerman and Tsur (2014) studied stochastic interruptions to an orchard production cycle (in this case a drought). They only considered the case of a uniform aged orchard.

Franklin (2012) developed a computational model for wine grape production in South Australia, relaxing the assumption of perfect capital markets, and including stochastic water supply. He found cycles in consumption and planting in a model where household income depended solely on perennial production. Including annual production, however, led to damped oscillations in area and consumption, eventually converging to a steady state.

A common thread through many of these papers is whether the market for perennials exhibits cyclical behavior, and if so, whether the cycles persist or dampen in the long-run. Tisdell and De Silva (1986) discuss in broad terms the desirability of the even-aged orchard in the long run, where an equal share of land is devoted to each age-class, but they do not demonstrate the optimality of the even-aged orchard in an optimization framework. Further, they raise the issue of how to optimally approach a steady-state (regardless of what the steady-state is), noting:

The problem of how optimally to achieve the steady state age-distribution of the crop needed to minimize fluctuations in annual yield has not been discussed. There would be a number of alternative paths to this steady state path. The question of whether to phase such a path in over one cycle or several cycles is likely to depend upon the existing age composition of the crop (Tisdell and De Silva, 1986, p. 248).

I am aware of two papers that explicitly consider the theoretical, optimal, long-run, age-structured steady-state: the studies by Mitra et al. (1991) and Wan (1993). They find that both even-aged orchards and long-run cycles may be optimal, depending on the initial conditions. Further, Mitra et al. (1991) finds, under a restrictive condition on the age-yield



relationship, that an orchard starting from an age-structure different from an even-aged orchard will never converge to the even-aged orchard, but rather to a cycle in a neighborhood around the even-aged orchard. I further discuss the implications and results of these two papers in chapter 4.

In this dissertation I explore several issues the economics of the production of perennial crops. In chapters 2 and 3 I examine the dynamics and age-structure of perennial crops in the context of sugarcane production for ethanol in Brazil. In chapter 4, I lay the foundation for a general model of optimal, age-structured perennial crop management decisions, building on the framework introduced by Mitra et al. (1991).

Brazilian sugarcane production growth, being a key feedstock to sugar and ethanol production, exhibited a puzzling slow-down in the decade starting in 2010. In chapter 2, I investigate a noted, but unexplored, mechanism to explain this slow-down: the link between credit availability, sugarcane replanting, and sugarcane yield. Using secondary sources, I establish the plausibility of credit restrictions having affected sugarcane farmer replanting decisions. To establish the link between replanting and yield, I develop a formal model to analyze the dynamics of perennial crop yield after a change in the replant rate. I test this model econometrically using data from the South-Central region of Brazil, finding evidence mostly consistent with the model. The model is able to explain around one third of the variation in sugarcane yields, implying that, while this channel is important, it alone cannot explain the production slow-down.

Perennial crop field-biorefinery supply chains are necessary in the production of many agricultural products. In particular, the low-cost production of low-carbon biofuels, such as ethanol from sugarcane or cellulosic feedstocks, relies on minimizing the costs of production along a perennial crop field-biorefinery supply chain. In chapter 3, I develop and analyze an unexplored mechanism to reduce perennial crop field-biorefinery supply chain costs: adjusting the age-structure, and hence yield, of the perennial feedstock. I present comparative statics of this model, finding that smaller refineries are most likely to benefit from age-structure endogenization. However, the results from a simulation of this model, calibrated to the sugarcane ethanol industry in the South-Central region of Brazil and comparing the cost-minimizing to the yield-maximizing age-structure, show that the cost-reductions from endogenization are small in this case (less than 1 percent cost reduction). Generally, the magnitude of the cost-reduction will depend on the age-yield relationship of the crop, the costs of growing and transporting the feedstock, and the reference age-structure.

In chapter 4 I adapt the theoretical framework of Mitra et al. (1991) to a two-age-class, finite horizon model, recasting their growth theoretic model into a form better suited for policy analysis. After introducing a simplified version of the model, and its general case, I determine necessary conditions on key parameters (relative productivity of mature trees, opportunity cost of land, and farmer patience) for each qualitative trajectory type in an arbitrary period. I use these conditions to develop a proposition about planting in the final period, and to analyze an example trajectory. Finally, I propose ways that this model can be used to answer several questions about perennial crop management.

Chapter 4 ties together the analytical frameworks from chapters 2 and 3. Chapter 2

studied perennial crop dynamics without an explicit optimization framework. Chapter 3 included an explicit optimization framework, but assumed away the dynamics of planting and growing by restricting attention to changes between even-aged orchards. Chapter 4 lays the foundation for a model that includes all three components—optimization, age-structure, and dynamics.

# Chapter 2

## The Yield Dynamics of Perennial Crops: An Application to Sugarcane in Brazil

### 2.1 Introduction

Sugarcane production in Brazil, being the key feedstock of Brazil's ethanol industry, expanded rapidly in the 2000's, leading Brazil to become a major producer of ethanol. This was an explicit policy goal of Brazil's government at the time, with President Luiz Inacio Lula da Silva declaring that Brazil wanted to become the "Saudi Arabia of Biofuel" (Globo, 2007).

Brazilian biofuel is a particularly valuable liquid fuel in the fight against climate change, with Crago et al. (2010) estimating that its life-cycle carbon emissions are about half those of corn ethanol. Furthermore, ethanol is an important component of Brazil's economy, accounting for around 2.3 percent of its GDP in 2010 (Valdes, 2011).

However, to the surprise of observers, Brazilian sugarcane production growth slowed at the end of the decade, stagnating from the period between 2010 and 2014, even reversing in 2011. Also illustrative of this slow-down is investment in sugarcane processing capacity, where, in the mid-2000's when the industry was booming, a net of 27 new sugarcane-processing mills opened in 2008, but from 2011 to the present a net of ten mills have been closing each year (UNICA, 2014). In this chapter I contribute to the explanation of the slow-down of Brazilian sugarcane production growth by focusing on a mechanism that has been often mentioned in passing, but not yet explored in detail.

I decompose production changes into area and yield effects and find that the production change from 2010 onwards was yield driven. What happened to yields? Why did Brazil's average sugarcane yield drop by approximately 10 tons/ha in 2011? I have collected evidence from journalistic and industry reports that the low yield may be a consequence of the 2008-2009 global financial crisis. Newspapers and observers noted that during this period of slow-down sugarcane growers and mills struggled to obtain credit, the region faced bad

weather, and the average age of production increased (Cohen, 2012; Crooks and Meyer, 2011; da Silva, 2016; Ewing, 2013a,c,d, 2014; Moreira, 2015). For example, the *The Economist* (2012) stated that “Poor weather, and cash-strapped growers delaying their replanting after the 2008 credit crunch, have recently squeezed production.”

Replanting sugarcane provides the link between credit availability, average age, and yield. Average age of a perennial crop is linked to its yield, since perennial crops, after they are established, generally become less productive as they become older. Replanting sugarcane is an expensive operation, and is accounted for as a capital expense, rather than an annual operating expense: “The high cost of establishing plant cane should be accounted as a capital investment” (James, 2004, p.109). In Brazil access to credit became more scarce for farmers, who, lacking access to the necessary funds, chose to delay replanting their sugarcane fields, hence increasing field age, lowering yields.

This chapter has two parts, the first presenting the link between credit and replanting, and the second presenting the link between replanting and yield. Because of a lack of data on credit availability, I rely on informal and qualitative sources such as newspaper reports to make the link between credit and replanting. I then present a formal model and empirical analysis of the link between sugarcane replanting and yield .

My model allows me to predict the future trajectory of yield in response to a change in the replant rate. The model is applicable to a wide-variety of perennial crops, allowing for an arbitrary number of age-classes, and an arbitrary yield in each age-class. The choice of replant rate is exogenous to the model, allowing me to avoid specifying the mechanism by which credit constraints affect the renewal rate, which is clearly important but beyond the scope of this chapter.

Using the model, I analyze the yield trajectory between steady states of the system that results from both a marginal and a discrete change in the renewal rate. I am able to determine, in general, the number of lags of the renewal rate that affect current yield, and whether the equilibrium yield after a replant rate shock will be higher or lower than the old equilibrium; the relationship can be non-monotonic.

I use these predictions to develop an econometric model that I then use to test the theory, and to predict the effect replant rates had on sugarcane yields. I used yield data from Brazilian Institute of Geography and Statistics, and data on area replanted from the CANASAT project, a remote sensing effort led by the Brazilian National Institute for Space Research from 2005 to 2013. My results are generally consistent with the predictions of the theoretical model, but changes to the replant rate only explain about one third of the yield variation over the sample period, leading to the conclusion that changes in the replant rate are an important, but incomplete, component of the explanation of Brazil’s sugarcane production slow-down.

Although I do not to verify it quantitatively, the story linking credit constraints and yield suggests an important issue in the analysis of perennial crop yields, namely that they can be adversely affected for several years after a shock to any factor that affects the replant rate.

The chapter is organized as follows. Section 2.2 reviews related literature. Section 2.3 provides background and context for the Brazilian sugarcane ethanol industry. Section 2.4

decomposes sugarcane production changes into area and yield effects, identifying yield as the primary determinant of production since 2010. Section 2.5 provides an overview of the effect of the 2008-09 financial crisis on Brazil and the sugarcane industry in particular. Section 2.6 provides a qualitative link between credit constraints and perennial crop replanting. Section 2.7 develops and analyzes a theoretical model of perennial crop yields as a function of the replant rate. Sections 2.8 through 2.10 present my empirical test of this model. Sections 2.11 and 2.12 discuss the results and conclude.

## 2.2 Literature Review

This chapter analyzes the effect of the credit crisis on Brazilian sugarcane production, emphasizing the relationship between credit, perennial crop replanting, and perennial crop yield dynamics. It is therefore related to several branches of existing literature, but, to my knowledge, this is the first study to bring these strands together.

The link between the credit market and sugarcane production has been noted by several authors, including many newspaper articles and industry reports (Cohen, 2012; Crooks and Meyer, 2011; da Silva, 2016; The Economist, 2012; Ewing, 2013a,c,d, 2014; Moreira, 2015). A quote from the book by de Moraes and Zilberman (2014, p. 195) succinctly summarizes this link:

However, after this period of rapid growth, there was a reduction in the production of both sugarcane (observed production of sugarcane was 559 million t) and ethanol (production was reduced to 22.7 billion L in the 2011-2012 harvest season). Several factors may explain this reversal: increased costs of production (through higher wages and land prices), the global financial crisis, lack of credit, climate and weather changes, indebtedness of productive units (and, consequently, the postponement of investment for renovation of sugarcane plantations), and loss of competitiveness of hydrous ethanol compared to gasoline.

These sources, however, only mention the link between credit, replanting, and production in passing, neither elaborating on nor quantifying the link.

The report by Mendonça et al. (2013) comes closest to this study in elaborating the effects of the credit crisis on the Brazilian sugarcane sector. Unlike this study, however, they rely mainly on textual analysis, and focus more on the changes in the corporate structure of the industry and the effects on sugarcane laborers. Also, like de Moraes and Zilberman (2014) they only briefly mention, and do not elaborate on, the need for replanting to maintain yields.

Agricultural credit has a long history of study in economics, and a longer history of use in agriculture. The second volume of the American Economic Review included an article titled *Agricultural Credit in the United States* (Kemmerer, 1912), which includes a delightful quote, attributed to King Louis XIV of France: “Credit supports agriculture, as the cord supports the hanged.” The question of whether credit hinders, as King Louis implies, or helps farmers is beyond the scope of this study, but clearly agriculture has been tied to credit

since at least the dawn of the modern era. Barry and Robison (2001) survey the agricultural economics literature and find that there is generally little evidence for “widespread, chronic” credit rationing in developed countries, but credit rationing by commercial lenders may be greater in developing countries, where both lenders and borrowers have questionable viability. Additionally, “in times of financial stress, credit may become more constraining as borrowers’ creditworthiness weakens.”

What of the relationship between credit constraints and agricultural productivity? Guirkinger and Boucher (2008) present evidence on the effect of credit constraints on the productivity of farmers in Peru, finding that credit constrained farmers are less productive than unconstrained farmers. Likewise Dong (2012) studies farmers in rural China, finding that removing credit constraints can increase agricultural productivity by up to 75 percent. Briggeman et al. (2009) compare credit constrained to unconstrained farmers in the US, finding that the presence of a credit constraint decreases the value of production for sole proprietorships by \$39 000. These studies, like the majority of similar studies before them (see introduction of Guirkinger and Boucher (2008) for overview), take a dual approach, measuring productivity in terms of farm revenues and/or costs. The study by Ali et al. (2014), looking at farmers in Rwanda, is unusual in taking a primal approach to productivity, measuring farm productivity in terms of yield (output per hectare), and finding that exogenously lifting credit constraints can increase yields by at least 17 percent.

The studies cited in the previous paragraph focus on the productivity of annual production, whereas this study explicitly focuses on the dynamic productivity effects associated with a perennial crop. There is a much smaller literature on the effect of credit on perennials, which is surprising since the durability of perennials places them closer, conceptually, to capital than to variable inputs (James, 2004). Bocquého and Jacquet (2010) simulate the decision of French farmers whether to adopt the perennial grasses switchgrass and *miscanthus*. They find that the presence of liquidity constraints reduces the quantity of perennials in the optimal crop mix. Similarly, Miao and Khanna (2015) analyze a two-period model of perennial crop choice, finding that the presence of a binary credit constraint reduces the land allocated to perennials.

There is a substantial literature on perennial supply response in agricultural economics (Rausser, 1971; French and Matthews, 1971; Knapp, 1987; Trivedi, 1986; Knapp and Konyar, 1991; Kalaitzandonakes and Shonkwiler, 1992; Devadoss and Luckstead, 2010), focusing on structural estimation of production response, with separate equations for planting and removal. These models acknowledge the importance of age-structure in perennial crop production, and the structural modeling approach attempts to keep track of age-structure—often difficult due to data limitations. These papers, however, focus on the response of supply to changes in input and output prices, rather than changes in agricultural credit.

## 2.3 The Brazilian Sugarcane Industry

In the 2014/15 harvest year, Brazil produced 532 million tones of sugarcane, processed into 35.5 million tons of sugar (of which 24.2 million tons were exported) and 28.4 billion liters of ethanol (of which 1.4 billion liters were exported) (UNICA, 2015). This harvest was grown on 10.9 million hectares of land, a small fraction of Brazil’s 330 million hectares of arable land, but a more sizable fraction of its 60 million hectares of cultivated land. Brazil is by far the world’s largest producer of sugarcane, producing a greater mass of sugarcane in 2015 than the next 6 largest producing countries combined.

The sugarcane sector plays a substantial role in Brazil’s economy. In 2015, the sugarcane sector’s revenue was greater than US\$70 billion, which is around 3.5 percent of Brazil’s GDP. The sector brought US\$10.2 billion into Brazil from export earnings. Just over 1 million workers are directly employed by the sugarcane sector, which is just under 1 percent of Brazil’s labor force (UNICA, 2015).

Brazilian sugarcane is processed into either sugar or ethanol. For a liquid fuel, Brazilian sugarcane ethanol has particularly low carbon emissions, with Crago et al. (2010) estimating that, on an energy equivalent basis, it reduces carbon emissions by 74 percent relative to gasoline, and its life-cycle emissions are about half that of corn ethanol.

In 2015, 91 percent of the area planted with sugarcane in Brazil was in the south-central region, and 9 percent was in the north-east. Although the north-east is the oldest growing region in Brazil, with cultivation dating back to the 1500s, the growth of the industry in modern times has been centered in the south-central growing region.<sup>1</sup> Figure 2.1 shows the sugarcane producing regions in Brazil. The centers of the south-central and north-eastern growing regions are over 2000 kilometers from the border of the Amazon rainforest.

In 2015, 98 percent of the sugarcane grown in the south-central region comes from 6 states: Goiás, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Paraná and São Paulo. São Paulo is by far the largest producer, responsible for 60 percent of sugarcane production. The next largest producing state, Minas Gerias, accounts for 11 percent of production.<sup>2</sup>

Sugarcane is a perennial grass, usually grown in rotations of 4-8 years, that is harvested and sent to local mills for processing into sugar or ethanol. Harvesting takes places between April and December, the dry season, and the sucrose content of the cane reaches a maximum in August and September. Mechanized harvesting is replacing manual harvesting, eliminating the need to burn the cane. A single machine can harvest up to 800 tons of can in a single day.

After it is cut, sugarcane is highly perishable, needing to be processed in a mill as fast as possible to avoid losing sugar content. Most cane is collected from fields close to the mill—the average distance is 22 kilometers in São Paulo—to the mill, and is delivered less than 24 hours after harvesting.

At the mill, the sugarcane stalks are crushed. The resulting fiber, along with some cane straw, is burned to produce electricity, while the juice is purified and processed into sugar

---

<sup>1</sup><http://www.unicadata.com.br> Accessed: 31 Dec, 2016

<sup>2</sup><http://www.unicadata.com.br> Accessed: 31 Dec, 2016

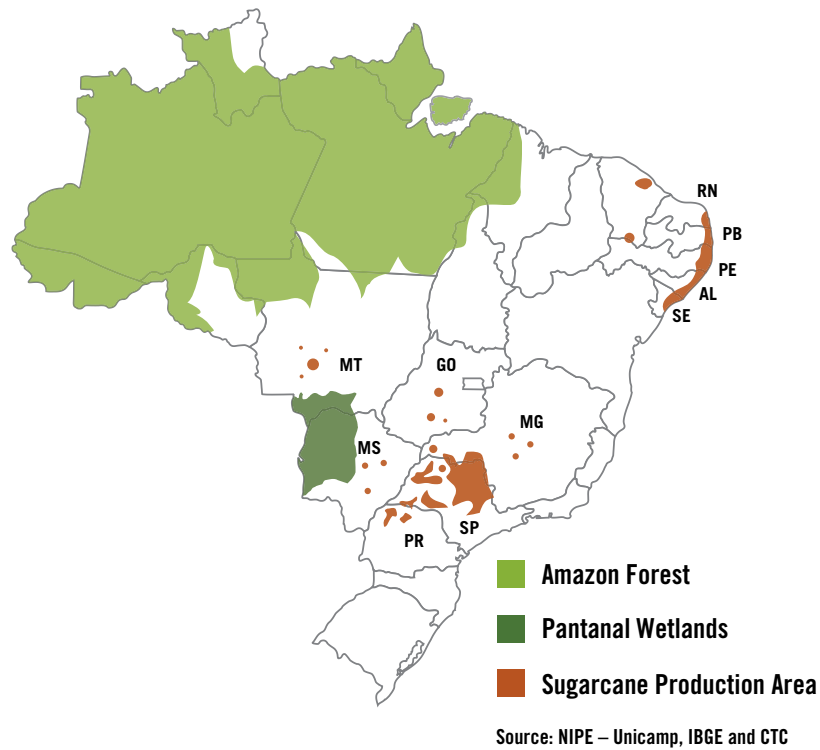


Figure 2.1: Map of sugarcane production regions in Brazil (from UNICA (2015))

and/or ethanol, depending on the configuration of the mill and the market conditions at the time. In 2015 there were 369 sugarcane mills operating nationwide (UNICA, 2016).

## 2.4 Decomposing Sugarcane Production into Area and Yield Effects

Around 2010, Brazil's sugarcane production ended a decade long period of steady growth, entering a period of relative lethargy. This period was heralded by a decline in production of around 10 percent in 2011, relative to 2010. What accounts for this change in trend? In particular, how much of this change can be attributed to changes in area, and how much to changes in yield?

Visually inspecting the area and yield panels in figure 2.2 suggests that area growth has been the main driver of overall production growth, but that yield deviations bear more responsibility for production pattern after 2010.

I can be more precise, however, in attributing changes in production to changes in area and yield. By definition, production is the product of area and yield,  $Q = Y \times A$ . Taking the



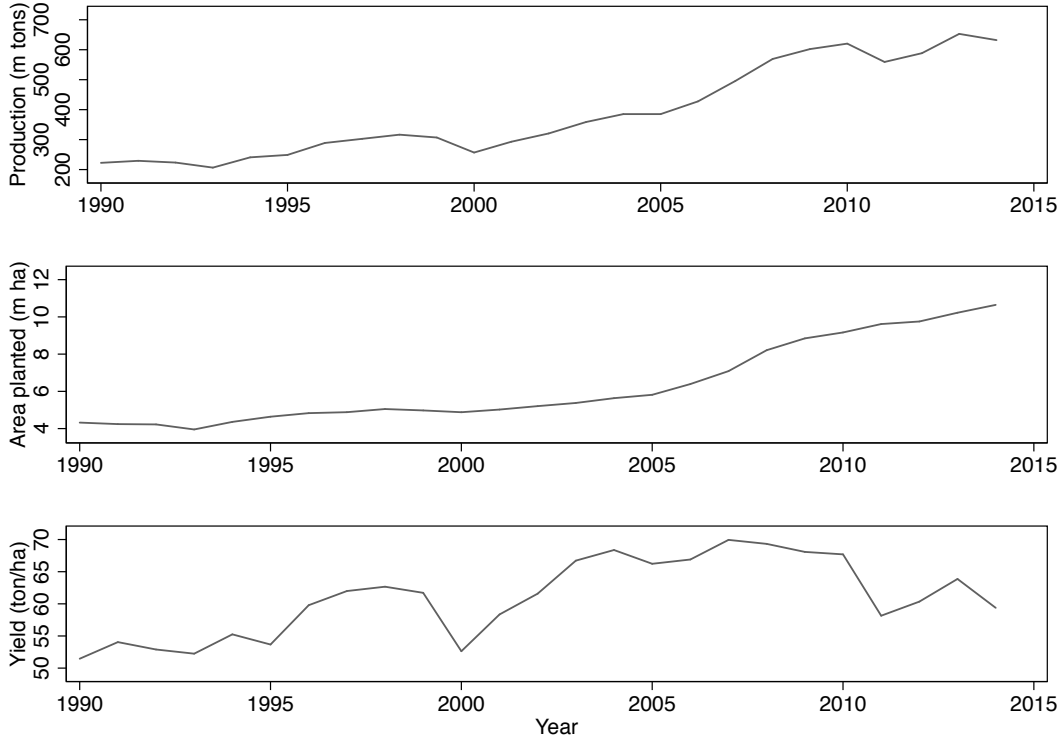


Figure 2.2: Sugarcane production, planted area, and yield series for Brazil from 1990 to 2014.

total derivative of this identity to attribute changes in quantity to changes in area and yield (Babcock, 2015), that is,  $dQ = dA \cdot Y + dY \cdot A$ . This identity is true for infinitesimally small changes in the variables.

To actually calculate this decomposition it must be discretized. Alauddin and Tisdell (1986) propose an extension of the decomposition method proposed by Venegas and Ruttan (1964), where changes in production are the sum of three effects

$$\begin{aligned}
 \Delta Q &= \underbrace{(\Delta A)Y}_{\text{Area Effect}} + \underbrace{(\Delta Y)A}_{\text{Yield Effect}} + \underbrace{(\Delta A)(\Delta Y)}_{\text{Mixed Effect}} \\
 (2.1) \quad Q_t - Q_0 &= (A_t - A_0)Y_0 + (Y_t - Y_0)A_0 + (A_t - A_0)(Y_t - Y_0)
 \end{aligned}$$

Figure 2.3 shows these three effects graphically. The area effect is the change in production attributable to changes in area, holding yields constant. It is represented by area  $A$ . The yield effect is the change in production attributable to changes in yield, holding area constant, represented by area  $B$ . Finally, the mixed effect is that contribution to production where both area and yield changes simultaneously, allowing the average yield of additional land to be different from the average yield of the existing land, represented by area  $C$ .

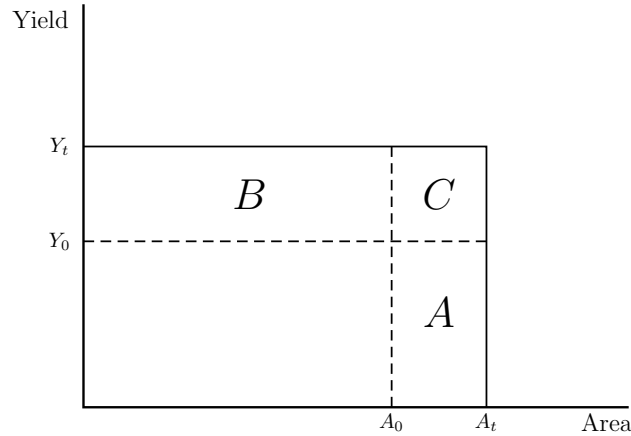


Figure 2.3: The three effects produced by the modified Venegas and Ruttan decomposition. Area  $A$  represents the area effect; area  $B$  represents the yield effect; and area  $C$  represents the mixed effect

Figure 2.4 shows the results of decomposing Brazilian sugarcane production into area, yield, and mixed effects using equation (2.1) from 1990 to 2015. The discretization in equation (2.1) decomposes the change over an arbitrary length of time ( $t$  periods). In this case the time step is a single year.

I exclude the result for 2005 from the graph since there was practically no change in production between 2004-2005 (nearly two orders of magnitude smaller than the next smallest production change), where an increase in area was almost exactly offset by a decline in yield. Such a small production change led to a small denominator when the decomposition shares were normalized and a distracting outlier when placed on the graph.

I also changed the sign of the effects to correspond to the sign of the production change. For example, in 2011-12 production declined by 10 percent from the year before. This decline was driven by a large decline in yield, and partially offset by an increase in area. The decomposition, as written in equation (2.1), would place a positive sign on the yield effect (because it was moving in the same direction as the production change) and a negative sign on the area effect. In my opinion it is more intuitive that negative changes in production should correspond to negative signs on the effects, so I report the negative of the calculated effect for those years when the production change is negative.

Looking at the results of the decomposition, I see three distinct periods. First, from 1990–91 to 2003–04, the effect of area and yield is relatively equal, with neither effect dominating the production trajectory. In the second period, from 2004–05 to 2010–11, I see a distinct decoupling between area and yield changes. During this period production growth is driven almost entirely by area growth, and the contribution of yield to growth is small, or slightly negative. Also observe during this period that area driven production growth increases from 2003–04 to 2009–10 after which the effect size declines. The third period, from 2011–12 to 2014–15, is a period of highly variable yield effects. During this period the magnitude of the

yield effects dominate the area effects, and I see unusually large negative yield contributions in 2011-12 and 2014-15. Area driven growth is positive during this period, but mostly continues the decreasing trend started in 2009-10. Throughout the entire time horizon, the mixed effect plays an insubstantial role in explaining changes in production.

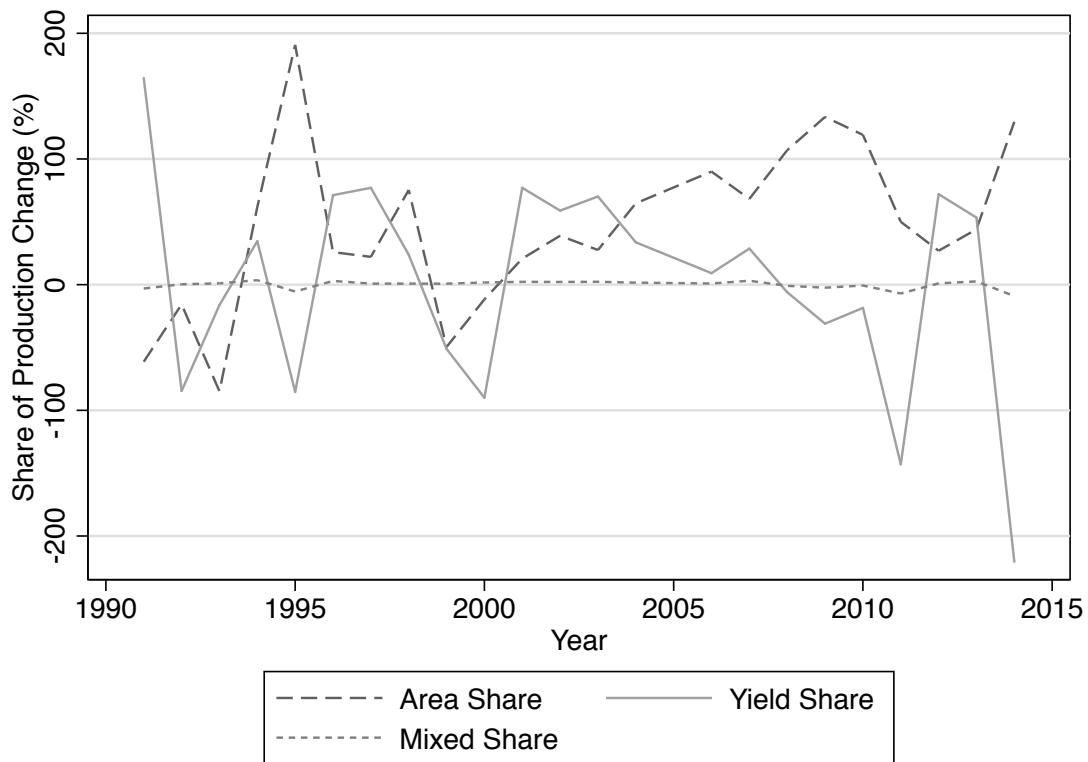


Figure 2.4: Decomposition of yearly changes in sugarcane production into area driven changes, yield driven changes, and mixed effect changes (using equation 2.1). In 2004-05 there was minimal production change since an increase in area almost exactly offset a decline in yield.

The task of the remainder of this chapter is to explore an explanation for the yield swings after 2009-10, yield being the main production driver during the period when production growth slowed.

## **2.5 The 2008-09 Financial Crisis is a Prime Candidate for Explaining the Yield Variation**

### **The credit crisis prevented renewal and expansion of sugarcane**

An important factor explaining the recent performance of the Brazilian sugarcane industry is the credit shortage caused by the 2008-09 credit crisis. Credit is a key determinant of sugarcane production. Carlos Sperotto, the vice president of Brazil's National Agriculture Confederation said in 2009 "The amount of financing will determine the size of the crop... lower financing means smaller crops" (Cortes, 2009). Credit also plays a key role in the construction and maintenance of sugarcane refineries.

During the credit crisis, the availability of credit to farmers was restricted substantially. This caused flow-on effects to the expansion and replanting of sugarcane fields, and the construction and operation of refineries. In this section I will explain how the 2008 credit crisis moved to Brazil, and how its effects permeated the economy. Then I show reasons why this initially financial problem led to real impacts on the sugarcane sector.

### **The effect of the credit crisis on Brazil**

The credit crisis in Brazil began with a sell-off of shares and currency (The Economist, 2008). As fears of the credit crisis spread, transnational corporations removed their capital from emerging markets, returning it to their domestic markets. They also increased their rate of profit repatriation to strengthen their domestic balance sheets (Filho, 2011). This flight of foreign capital put downward pressure on both the exchange rate of the Brazilian Real, and the price of shares on the domestic share market (de Mendonça and Deos, 2013).

The flight of capital caused serious losses for holders of currency derivatives. In the years leading up to the crisis, Brazilian corporations had been accepting dollar put options from banks in exchange for lower funding costs (Canuto, 2008; de Mendonça and Deos, 2013). By the start of the financial crisis, Brazilian corporations were exposed to around US\$37 billion worth of currency derivatives, whose value would fall if the Real depreciated (de Mendonça and Deos, 2013). The Real's sharp fall against the dollar caused losses for holders of exchange rate derivatives predicated on the continued rise of the real against the dollar (The Economist, 2008). The losses from these derivatives led to fears about further falls in other derivatives, and uncertainty about which assets would be affected (Canuto, 2008).

As foreign capital was repatriated credit became more scarce in Brazil, with debtors and potential borrowers turning to domestic sources. Even as the demand for domestic credit increased, its supply decreased as Brazilian banks increased their preference for liquidity (de Mendonça and Deos, 2013). Due to the combination of these forces, during the year of 2008, the average lending interest rate in Brazil increased from 40 to 55 percentage points.

These changes in the financial markets were largely channeled into the real economy through two sectors: industrial production, and exporters of manufactured goods. Real GDP fell in the last quarter of 2008 and the first quarter of 2009, with the output gap closed by the

end of 2009 (Canuto, 2010). This fleeting fall and rapid recovery of GDP was made possible by the policy responses by Brazilian institutions, discussed in the next section.

That the financial crisis affected Brazil as much as it did surprised analysts and observers of Brazil's economy (Filho, 2011). Before the crisis Brazil had a low debt to GDP ratio and, although it had been growing fast, at the time of the crisis it was around 40 percent of GDP—a much lower figure than in developed countries (The Economist, 2008). Observers had believed that Brazil was 'decoupled' from the developed world and would fare well during the financial crisis. However they became more pessimistic after the collapse of Lehman brothers in the USA (Sobreira and de Paula, 2010).

### **Policy responses to the credit crisis**

Responding to the crisis, Brazilian policy makers used four main tools. First, the central bank eased reserve requirements to create liquidity. Second, public banks created additional lines of credit to offset the freeze in private credit. Third, the central bank reduced official interest rates. And fourth, the central bank supported the exchange rate of the Real.

The Central Bank of Brazil (BCB – *Banco Central do Brasil*) relaxed reserve requirements for banks in order to increase the money supply and encourage lending. Taken together, the changes to reserve requirements injected around US\$115 billion into the economy, just shy of 6 percent of GDP (Canuto, 2008; de Mendonça and Deos, 2013; Sobreira and de Paula, 2010). In addition to relaxing the reserve requirements, the BCB also moved to increase confidence in the banking system by increasing its guarantee on time deposits from R\$60 000 to R\$20 million per depositor per bank (de Mendonça and Deos, 2013).

The Brazilian government directed the large public banks to operate in the credit markets to offset the private banks' preference for liquidity (Filho, 2011). From 2008-2010 the private sector did not increase its credit outstanding as percentage of GDP, and it was this reluctance to lend that the government wished to offset through the public banks (Bonomo et al., 2014). The government created lines of credit through the National Bank for Economic and Social Development (BNDES – *Banco Nacional de Desenvolvimento Econômico e Social*), Bank of Brazil (BdB – *Banco do Brasil*), and the Federal Savings Bank (CEF – *Caixa Econômica Federal*). The BNDES and BdB funded investment expenses and working capital. Additionally, the BCB and the CEF provided credit to finance exports (Filho, 2011).

The BCB lowered both long and short term interest rates. Between December 2008 and September 2009 the BCB lowered the SELIC rate (overnight lending rate) from 13.75 to 8.75 percent (Filho, 2011). In mid-2009 the National Monetary Council (comprised of members of the government and the head of the BCB) lowered the long term interest rate (TJLP) from 6.25 percent to 6 percent. The TJLP is the rate used by the BNDES for granting long-term investment loans.

The BCB supported the Real's exchange rate, selling US\$14.5 billion in the spot market to mitigate the effect of the Real's devaluation (Filho, 2011; de Mendonça and Deos, 2013). The BCB was also active in the exchange derivatives market, offering swaps in which it assumed liability positions in dollars (de Mendonça and Deos, 2013).

## **The effect of the credit crisis on the Brazilian sugarcane industry**

Unlike the relative robustness of the Brazilian economy as a whole, the sugarcane sector was in a far more fragile state at the start of the financial crisis. The sugarcane industry had greatly increased its indebtedness levels between the 06/07 and 07/08 harvest years. In 06/07 industry debt stood at almost 50 percent of revenue, composed mostly of long-term debt. This level was the highest it had been that decade, but it was not unusually high. In 07/08, however, indebtedness levels skyrocketed to almost 125 percent of revenues, and the short-term component of this debt rose to 40 percent of revenues, almost to the level of total debt the year before. This placed the sugarcane industry in a highly vulnerable position, ready to be toppled by the strong financial headwinds that would hit the financial markets at the end of 2008 (Moreira, 2015).

The advent of the financial crisis affected both the processing and growing of sugarcane. The acceleration of processing plant openings ceased and plants began to close. “The 2008-2009 crisis combined the impossibility of accessing resources for debt rollovers with losses from investments in exchange rate derivatives. Many plants went under, which intensified the process of acquisitions by and mergers with transnational corporations” (Mendonça et al., 2013). At the growing level, the rapid expansion of the previous years slowed and the percentage of the sugarcane crop being replanted began a substantial decline from its peak of 22 percent in the 08/09 growing year down to a minimum of 7 percent in the 10/11 growing year (UNICA, 2014). The drop-off in replanting was driven by lack of access to capital (Ewing, 2013b). “Investments have slowed in the expansion of cane production, one of the most capital intensive aspects of the business” (Ewing, 2014).

Credit, for Brazilian sugarcane growers and processors, is important for investing in replanting, which maintains the productivity of the crop. Recall the quote from the vice president of Brazil’s National Agriculture Confederation, Carlos Sperotto: “the amount of financing determines the size of the crop...lower financing means smaller crops” (Cortes, 2009). Following the credit crisis “companies stopped investing in, for example, the renewal of sugarcane plantations, crop handling and fertilization, which are needed to maintain high production levels” (Mendonça et al., 2013).

Farmers receive credit from a mixture of private and public lenders. Farmers receive credit for purchases of machinery from its manufacturers, John Deere and Case New Holland, for example. They also receive financing from commercial banks and the BNDES. I do not have data on the exact composition of funding sources, but for illustrative purposes, funding for the Brazilian soybean industry during the 07/08 harvest year was 45 percent from agribusinesses, 30 percent from banks, and 25 percent from the farmers themselves (Wheatley, 2008). During the financial crisis, the BdB provided funding to offset the withdrawal of private funding from banks and agribusiness (Barbosa, 2011).

As early as the end of 2008 it was clear that farmers were having trouble obtaining credit, and that it would affect future production (The Economist, 2008). Agribusinesses, a major source of farm finance, were restricting their lending to farmers (Wheatley, 2008). For example, two large traders, Cargill Inc. and Bunge Ltd., cut their credit to farmers to 25

percent in 2009, down from 52 percent the previous year (Cortes, 2009).

Farmers were not able to take full advantage of even the credit that was available from private sources: “out of R\$49 billion made available to finance this year’s [2008] planting...just R\$18.5 billion was actually lent to farmers nationwide” (Wheatley, 2008). The expansion of the growing region and the mechanization of harvesting had left farmers highly leveraged, leaving banks reluctant to lend. In a typical example, Emerson Spigoso, a farmer in Mato Grosso state, could not access new loans, having already used his collateral to purchase combine harvesters. “He bought one combine harvester in 2001 for R\$220 000 and another in 2003 for R\$280 000. Because of the accumulated interest, the debt on his two machines has grown to R\$800 000, while the machines’ value has fallen to less than half that amount” (Wheatley, 2008).

Similarly, processing mills closed due to lack of finance (Brough and Ewing, 2014). “Liquidity problems, stemming largely from the global financial crisis that caught the industry over-leveraged from its expansion, forced more than 40 mills to shut down over the past few years. This took more than 30 million metric tons of Brazil’s 690 million metric ton crushing capacity” (Ewing, 2013b).

## 2.6 Crop Renewal: The Link Between Credit and Yield

The previous section illustrates that Brazilian agriculture, and sugarcane in particular, experienced tumult during the period of the global financial crisis. But how can the impacts of this shock on the production and yield of the sugarcane sector be quantified?

I am not able to directly observe a link between credit constraints and yield, since it would require data which I do not have. It would require firm-level data on credit constraints, input use, replanting decisions, and yield. In the absence of this data I take another approach, making use of aggregate data that I can access, relying on the relationship between credit availability, sugarcane replanting, and yield. I develop the first link in this chain, the effect of a change in credit availability on sugarcane replanting, in this section and develop the second link, which has non-straightforward dynamic consequences, in the next section.

One impact of a credit constraint on the sugarcane industry is a reduction in investment in sugarcane replanting, since, for the majority of farmers, credit is necessary for replanting. Sugarcane replanting is a capital intensive operation, generally requiring financing, and indeed this was the case during the credit crisis. Unable to access credit, farmers responded by foregoing planned replanting operations, choosing instead to further harvest their aging canes (Ewing, 2013b). Following the credit crisis “companies stopped investing in, for example, the renewal of sugarcane plantations, crop handling and fertilization, which are needed to maintain high production levels” (Mendonça et al., 2013).

This behavioral response to a reduction in credit availability is supported by economic theory. Briggeman et al. (2009) provide a conceptual model illustrating the effect of a credit

constraint on farm household production. Their model shows that if the credit constraint is binding, the household's level of production will be less than a similar household not facing a credit constraint. While their model does not explicitly include investment in farm capital (as sugarcane would be), it illustrates the more general point that the existence of credit constraints can lead to choices of economic variables below their unconstrained optima.

The events in Brazil are consistent with banks rationing credit to farmers. "Out of R\$49 billion made available to finance this year's [2008] planting...just R\$18.5 billion was actually lent to farmers nationwide" (Wheatley, 2008). Stiglitz and Weiss (1981) describe a model where imperfect information creates an excess of credit demanded over credit supplied. In their model there exists a *bank optimal* interest rate, which maximizes the return on lending to the bank, given the bank's inability to perfectly screen the creditworthiness of potential borrowers. Banks will not offer loans to potential borrowers above this interest rate because the expected marginal return on these loans is negative. In Brazil the high proportion of potential bad borrowers with unusually high debt/revenue ratios (Moreira, 2015), combined with the banks' increased preference for liquidity in response to the credit crisis (de Mendonça and Deos, 2013), is likely to have led to a lowering of the bank optimal interest rate, and thus preventing farmers from obtaining fresh credit, or refinancing their existing loans, even though they may have been willing to pay a higher interest rate to obtain the loans.

The occurrence of credit rationing during the financial crisis is also consistent with the findings of the agricultural finance literature. Barry and Robison (2001) survey this literature and find that there is generally little evidence for "widespread, chronic" credit rationing in developed countries, but credit rationing by commercial lenders may be greater in developing countries, where both lenders and borrowers have questionable viability. Additionally, "in times of financial stress, credit may become more constraining as borrowers' creditworthiness weakens".

Accepting now the possibility that Brazilian sugarcane farmers experienced credit constraints, and that they delayed their replanting operations as a consequence, what implications does this have for yields? In the next section I develop a theoretical model of perennial crop yields as a function of the current age-structure and the history of replanting decisions. Using this model I can predict the effect of a decline in the replanting rate on current and future sugarcane yields.

## 2.7 The Yield Trajectory after a Change in the Replant Rate

Given my focus on replanting rates as a proxy for credit constraints, I need to know how a change in the replant rate affects yields, both immediately and in the future. In this section I develop a model of the dynamics of perennial crop yields as a function of replant rates, and develop sharp predictions on the sign and magnitude of future yields given a marginal increase in current replant rate.



First some preliminaries on the yield system I will be studying. Perennial crops, such as sugarcane, can be grown and harvested for multiple years before they need to be replanted. Over their lifespan, the yield of the crop changes with time. I call this pattern the *age-yield relationship*. Following Mitra et al. (1991) I can characterize the age-yield relationship into three phases: the establishment phase (increasing yield), the peak phase (constant, maximal yield), and the senescence phase (decreasing yield). The particular age-yield relationship will vary depending on the crop, the growing location, the farm management practices, pest load, temperature, and water availability, among other factors.

To illustrate the idea of an age-yield relationship, I present an example for the Alta Mogiana region<sup>3</sup> of São Paulo state (Margarido and Santos, 2012), shown in figure 2.5. The establishment phase occurs in the year of planting (year 0). The peak occurs in the first year after planting and lasts for only one year. The senescence phase begins in the second year after planting and continues until the 6<sup>th</sup> year. Since Brazilian sugarcane tends to be renewed by or before its 6<sup>th</sup> year, I am not aware of data on the age-yield relationship for Brazilian sugarcane for higher years. For Brazilian sugarcane, the first age-class is non-yielding since the plant is still establishing itself. “On average, the first cut is carried out one year-and-a-half after planting” (Margarido and Santos, 2012).

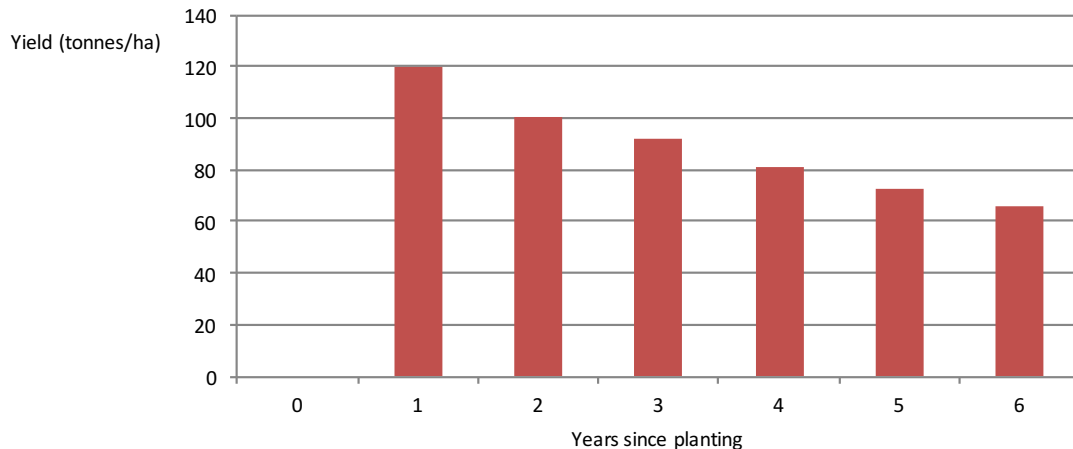


Figure 2.5: Age yield relationship taken from Margarido and Santos (2012). Freshly planted canes provide no yield (year 0).

Margarido and Santos (2012) identify the key features of sugarcane yield dynamics, mainly that the yield trajectory will be non-monotonic in response to a change in the renewal rate:

It is important to point out that after large decreases in planting or in renovation, there is a significant increase in total production in the next year, but drastic

---

<sup>3</sup>The Alta Mogiana region is in the north-east of São Paulo state. It is located within the Ribeiro Preto mesoregion, which is included in my dataset.

reduction in the second year, because of two factors: i) part of the first cut cane (1/7), which is used for seedlings, is not used for sowing, and therefore, it is added to the next growing season; ii) because of renovation itself, which if it is not carried out, increases the cutting area in the following year. (Margarido and Santos, 2012, p. 12)

Although they identify this key feature, they leave several questions unanswered. What happens in subsequent years? What will be the new equilibrium level of production? In order to create an econometric model of the effect of renewal rates on sugarcane yield I must have answers to these questions—both to correctly specify the model, and also to provide testable hypotheses. In the remainder of this section I develop a general model of yield trajectories as a function of changes in the renewal rate. This model uses an exogenous renewal rate—it is not determined by an optimization model. The model is applicable for any perennial crop, and I apply it to a stylized representation of Brazilian sugarcane to generate testable hypotheses for this specific case.

## The dynamics of age-structure

Before considering the dynamics of the yield of a perennial crop I must first consider the dynamics of its age-structure. Age-structure is the division of the plants in a growing region into different age-classes. I study the simplest model of age-structure dynamics, where there is a fixed plot of land (size normalized to 1) divided into sub-plots of different ages, e.g. 20% of the land may be freshly planted cane, 10% of the land is 1-year-since-planting cane, and so on.

Let  $x_{st}$  be the quantity of land allocated to age-class  $s$  in year  $t$ . Under the natural dynamics of this system (i.e. without human intervention) the canes will enter the next oldest age-class next year,<sup>4</sup> that is,  $x_{s,t} = x_{s-1,t-1}$ . Following Mitra et al. (1991) and Salo and Tahvonen (2004) I assume the existence of some oldest age-class,  $S$ , creating  $S + 1$  age-classes in total (freshly planted cane is denoted by  $x_0$ ). This makes the analysis tractable by making the age-class state space finite. It is also reasonable—the oldest age-class could simply be a zero yield class for plants that are dead or non-yielding from old age.

Now I introduce the possibility of replanting, which is a human choice. Replanting a perennial crop means to replace an old plant with a fresh seed, seedling, or cutting. When replanting  $s$ -year-old plants, land is moved from age-class  $s$  to age-class 0. When considering a single replanting decision, this implies two linked dynamic equations:  $x_{0,t} = r_{st}$  and  $x_{st} = x_{s,t-1} - r_{st}$ . Not all land allocated to a single age-class needs to be replanted at once, and replanting happens at the start of a period, with yield being realized at the end of that period. Due to the physical nature of land, the replant choice variable is constrained to be between 0 and  $x_{st}$ .

---

<sup>4</sup>I am assuming away any loss between years, e.g. due to weather damage, or pest damage etc

Combining the natural and artificial dynamics of the system and summing up across all age-classes yields the following system of dynamic equations, which is illustrated in figure 2.6:

$$\begin{aligned}
 x_{0,t} &= \sum_{s=0}^S r_{st} \\
 x_{s,t} &= x_{s-1,t-1} - r_{st} \text{ for } 1 \leq s < S \\
 x_{S,t} &= x_{S-1,t-1} + x_{S,t-1} - r_{St}
 \end{aligned}$$

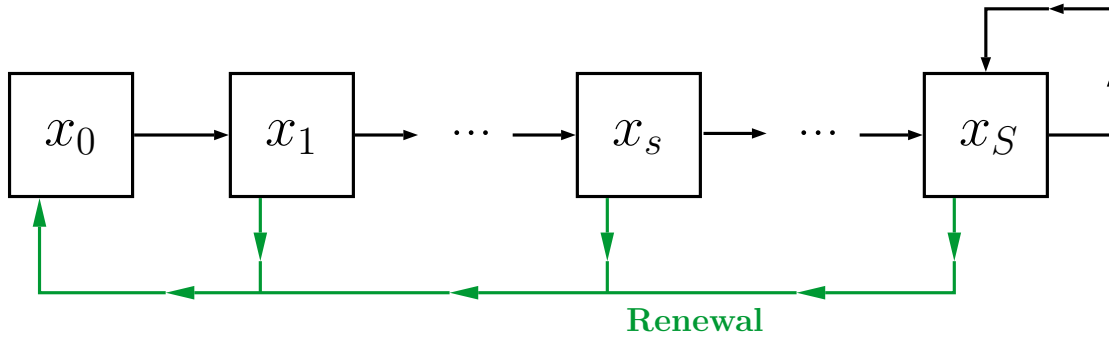


Figure 2.6: Diagrammatic representation of the dynamics of the quantity of land in each age-class. Each age-class is populated only from the previous age-class, except for the youngest age-class which collects all the land replanted in a particular year. Land in the oldest age-class stays in the oldest age-class unless replanted.

Let  $x_{st}$  be an *active age-class* if  $x_{st} > 0$ . In principle, land from any age-class could be replanted, implying that without further restrictions there could be an age-structure with active and inactive age-classes interleaved among each other. Mitra et al. (1991, section 3) demonstrate that old age-classes will be replanted in preference to young age-classes if the crop follows a single-peaked age-yield relation—like the one shown in figure 2.5.

The active-age contiguity result of Mitra et al. (1991) allows me to study the dynamics of a stationary system in terms of the renewal rate. In a stationary system the state of the system remains unchanged from period to period (Mitra et al., 1991). Let  $\mathbf{x}_t$  be the vector of land allocations across all age-classes in period  $t$ . Then in a stationary system  $\mathbf{x}_t = \mathbf{x}_{t+1}$ . To achieve this state, a constant fraction of the land must be renewed each year.

**Proposition 2.1** *In a stationary system a constant fraction of the land must be renewed each year.*

**Proof.** Since  $\mathbf{x}_t = \mathbf{x}_{t+1}$ , it follows that  $x_{1t} = x_{1,t+1}$ . Using the equation of motion for land in the first age-class to write this in terms of replanting decisions I get  $\sum_{s=1}^S r_{s,t-1} = \sum_{s=1}^S r_{st} \forall t$ . Hence the aggregate quantity of land replanted in each period must be constant in a stationary system. ■

In a stationary system there will be equal quantities of land allocated to all but the oldest active age-class, i.e.  $\mathbf{x}_t = \mathbf{x}_{t+1}$  AND  $x_{0t} = x_{1t} = \dots = x_{s-1,t} \geq x_{st}$ .

Let  $R$  be the replant rate, that is the fraction of the land that is renewed at the start of the year. For each replant rate  $R \in [0, 1]$  there exists a corresponding stationary system, denoted  $\mathbf{x}(R)$ , defined as:

$$(2.2) \quad \mathbf{x}(R_t) = \begin{cases} \text{for } \lceil \frac{1}{R} \rceil < S & \begin{cases} x_{st} = R & \text{for } s < \lceil \frac{1}{R} \rceil \\ x_{st} = 1 - R(\lceil \frac{1}{R} \rceil) & \text{for } s = \lceil \frac{1}{R} \rceil \\ x_{st} = 0 & \text{otherwise} \end{cases} \\ \text{for } \lceil \frac{1}{R} \rceil \geq S & \begin{cases} x_{st} = R & \text{for } s < S \\ x_{st} = 1 - RS & \text{for } s = S \end{cases} \end{cases}$$

where  $\lceil \cdot \rceil$  is the ceiling function. This characterization assumes a constant, unit area of land.

**Proposition 2.2** *If the replant rate is held constant at  $\bar{R}$ , then an arbitrary plantation will reach the stationary state described by equation (2.2) in at most  $\min(\lceil \frac{1}{\bar{R}} \rceil, S)$  periods.*

**Proof.** Start with a system in an arbitrary state. Let the replant rate be set to  $\bar{R}$  at the start of period  $t = 0$ . Thus  $x_{0,0} = \bar{R}$ . In each subsequent period  $x_{0t}$  will be set to  $\bar{R}$ . Hence after  $\min(\lceil \frac{1}{\bar{R}} \rceil, S)$  periods the fraction of land in each of the age-classes 0 to  $\min(\lceil \frac{1}{\bar{R}} \rceil, S) - 1$  will be equal to  $\bar{R}$ , and since I am assuming a constant quantity of land, age-class  $\min(\lceil \frac{1}{\bar{R}} \rceil, S)$  must contain  $1 - \bar{R}(\min(\lceil \frac{1}{\bar{R}} \rceil, S))$  units of land. This corresponds to the stationary-state in equation 2.2. ■

## Stationary-state yield after a change in the replant rate

Given this dynamic yield system, what happens to the stationary-state yield after a one-off, persistent change to the replant rate?

As equation (2.2) shows, when the replant rate is changed it is possible that the number of active age-classes changes too. If the replant rate increases sufficiently, the older active age-classes will become inactive, and, inversely, if the replant rate decreases sufficiently, previously inactive age-classes will activate. For this analysis, I only consider small, i.e. marginal, changes to the replant rate. In the case of a marginal increase<sup>5,6</sup> in the replant rate it is not possible for the number of active age-classes to decrease, since for any  $R \in [0, 1]$  there exists an  $\varepsilon > 0$  such that  $\lceil \frac{1}{R+\varepsilon} \rceil = \lceil \frac{1}{R} \rceil$ .

<sup>5</sup>It is possible for a marginal decrease in the replant rate to increase the number of active age-classes, but only if  $\lceil \frac{1}{R} \rceil = \frac{1}{R}$ . The set of such  $R$  has Lebesgue measure zero, and can thus be neglected for all practical purposes.

<sup>6</sup>See appendix A.1 for the case with non-marginal increases in the replant rate large enough to decrease the number of active age-classes.

**Proposition 2.3** *Equilibrium yield increases after an increase in the renewal rate if and only if*

$$\frac{f_0 + f_1 + \dots + f_{s-1}}{s} - f_s > 0$$

**Proof.** Equation (2.2) implies that all but the oldest age-classes will necessarily have  $R$  units of land allocated to them, and the oldest age-class will contain  $1 - R(\min(\lceil \frac{1}{R} \rceil, S))$  units of land, which allows me to rewrite the yield equation as a function of the renewal rate:

$$\begin{aligned} \text{yield} &= f_0 x_0 + f_1 x_1 + \dots + f_s x_s \quad \text{Where } s \text{ is the oldest active age-class} \\ \text{yield} &= f_0 R + f_1 R + \dots + f_{s-1} R + f_s (1 - R s) \end{aligned}$$

where  $f_i$  is the productivity of age-class  $i$ . The set  $\{f_0, \dots, f_s\}$  is the age-yield relationship.

The derivative with respect to  $R$  represents the change in stationary-state yield with respect to a change in the renewal rate.

$$\frac{d \text{yield}}{dR} = f_0 + f_1 + \dots + f_{s-1} - f_s s$$

This expression is positive if and only if  $\frac{f_0 + f_1 + \dots + f_{s-1}}{s} - f_s > 0$ . ■

That is, an increase in the replant rates increases stationary-state yield if the average productivity of all but the oldest age-class is greater than the productivity of the oldest age-class, or, alternatively, if having more land allocated to the oldest age-class reduces the average yield.

## Yield trajectory after a change in the replant rate

It is not enough to know the change in stationary-state yield from a marginal change in the replant rate, since to specify an econometric model I need to know the trajectory that yield takes to reach the new stationary-state. Proposition 2.2 says that the new stationary-state will be reached in at most  $s$  periods. Hence, for each of those periods ( $0 \leq t \leq s$ ) does yield,  $y_t$ , increase or decrease relative to the yield before the change,  $y_{-1}$ ?

**Proposition 2.4** *The change in yield  $t$  years after an increase in the replant rate, relative to the yield prevailing before the change,  $y_{-1}$ , is given by*

$$\frac{d(\Delta \text{yield}_{t,-1})}{dR} = \sum_{i=0}^t (f_i - f_s)$$

**Proof.** At the beginning of period  $t = 0$ , let the replant rate change from  $R$  to  $R'$  and let  $\Delta R = R' - R$ . The yield  $t - 1$  years after the renewal rate change is given by:

$$\text{yield}_{t-1} = f_0 R' + \dots + f_{t-1} R' + f_t R + \dots + f_{s-1} R + f_s (1 - R(s-t) - R' t)$$

Similarly, after  $t$  years, the yield will be given by:

$$yield_t = f_0R' + \dots + f_{t-1}R' + f_tR' + \dots + f_{s-1}R + f_s(1 - R(s - (t + 1)) - R'(t + 1))$$

The change in yield from  $t - 1$  to  $t$  ( $yield_t - yield_{t-1} = \Delta yield_t$ ) is given by:

$$\Delta yield_{t,t-1} = f_tR' - f_tR + f_s(1 - R(s - (t + 1)) - R'(t + 1)) - f_s(1 - R(s - t) - R'(t))$$

Simplifying and collecting like terms gives:

$$\begin{aligned} \Delta yield_{t,t-1} &= f_t(R' - R) - f_s(R' - R) \\ &= \Delta R(f_t - f_s) \end{aligned}$$

Hence,  $\frac{d yield_{t,t-1}}{dR} = \lim_{\Delta R \rightarrow 0} \frac{\Delta yield_{t,t-1}}{\Delta R} = \frac{\Delta R(f_t - f_s)}{\Delta R} = (f_t - f_s)$ .

The net change  $t$  years after a change in the replant rate is the sum of these year-to-year marginal changes

$$\frac{d(\Delta yield_{t,-1})}{dR} = \sum_{i=0}^t \frac{d(\Delta yield_{i,i-1})}{dR} = \sum_{i=0}^t (f_i - f_s)$$

■

## Application to Brazilian Sugarcane

I use the formulae developed in proposition 2.4 to get qualitative and quantitative predictions about the effect of a marginal change in the replant rate on Brazilian sugarcane yields.

In the Brazilian case,  $f_0 = 0$  and  $f_1 > f_2 > \dots > f_s > f_0$ . Thus

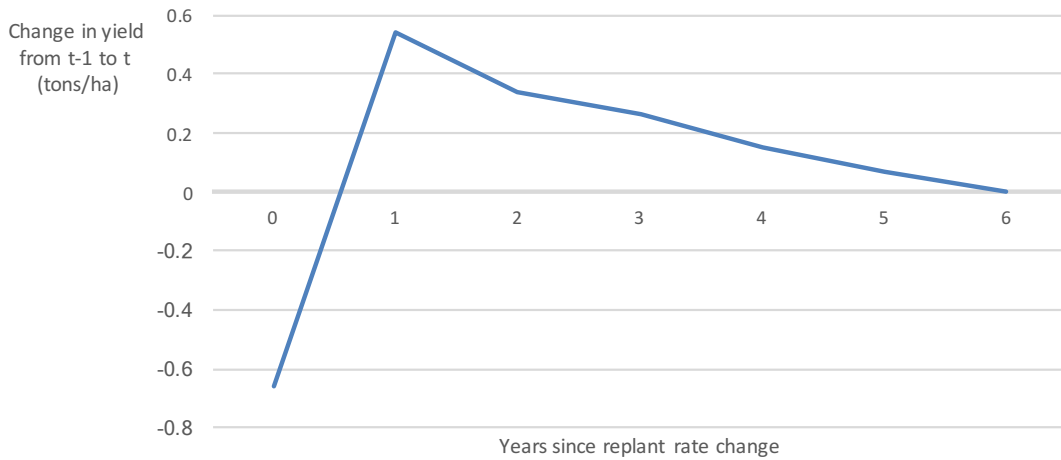
$$\frac{d(\Delta yield_{0,-1})}{dR} = f_0 - f_s < 0$$

and

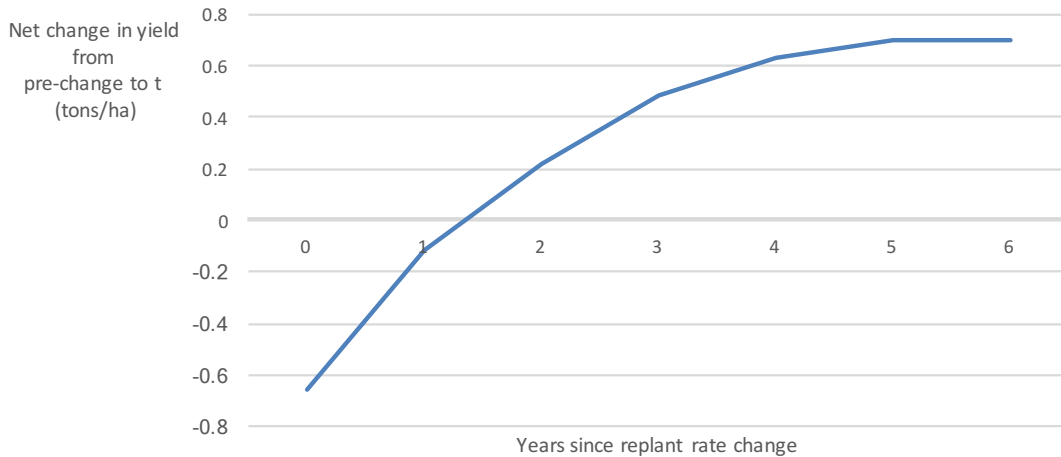
$$\frac{d(\Delta yield_{t,t-1})}{dR} = f_t - f_s > 0, \quad \forall t \text{ such that } 0 < t < s$$

Figure 2.7a presents these year-on-year changes using the Margarido and Santos (2012) age-yield relationship, showing the qualitative shape predicted above, with the first year-on-year change being negative, and the remainder being positive, each positive change being smaller than the last.

Figure 2.7b shows the *net* change in yield  $t$  years after a change in the replant rate, relative to the yield before the change. For Brazilian sugarcane, the change trajectory is a concave, monotonically increasing function of time since the change, with the same-year effect negative, the one-year effect slightly negative, and the subsequent effects positive until the new stationary-state is reached 5 years after the change, stabilizing the yield at its new level. The shape of the age-yield relationship determines the shape of this curve—the roughly zero net effect in the year following the replant rate increase is an artifact of the yield in the oldest age-class being roughly halfway between the yield of the first two age-classes.



(a) Marginal *year-to-year* changes in sugarcane yield  $t$  years since a change in the replant rate.



(b) Marginal *net* changes in sugarcane yield  $t$  years since a change in the replant rate.

Figure 2.7: The change in yield  $t$  years after a 1 percentage point increase in the replant rate. Graph generated using the São Paulo age-yield relationship from figure 2.5.

## 2.8 Empirical Strategy

To transition from the theoretical model to an empirical model, I need a change in perspective. The theoretical model explores the *future* impacts of a change to the *current* replant rate, while an empirical model is restricted to exploring the past. Hence in the empirical model I ask “in which previous year(s) could a change in the replant rate have affected the current

yield?”, changing to explaining current yield as a function of previous changes, or lags, of the replant rate.

The relationship between the change in the replant rate and its effect on current and future yields is given by proposition 2.4. For the econometric equation I examine the effect of current and past changes in the replant rate on the current yield. For example, consider a system where a current change in the replant rate only effects current yield, next year’s yield, and the yield two years from now. From the perspective of this year’s yield, it can only be affected by replant rate shocks from this year, last year, and two years ago. In this econometric specification I seek to capture the effect of these years’ changes in the replant rate on current yield, all else being equal.

The regression equation is

$$(2.3) \quad y_{it} = \sum_{l=0}^L \beta_l \text{ReplantRate}_{i,t-l} + \alpha \mathbf{X}_{it} + v_i + u_{it}$$

This equation implies that the yield in region  $i$  in period  $t$  is a function of  $L$  lags of the replant rate, including the contemporaneous replant rate, a vector of region and time specific co-variates (namely total sugarcane growing area), a region specific fixed effect, reflecting unobservable, unchanging differences in yield across regions, and an idiosyncratic shock. I control for total area in the regression since the theoretical model included the assumption that total area was unchanging over time.

I need to use multiple lags of the replant rate in the regression equation. The lags represent the *net* changes in yield since the replant rate change, since the coefficient on any lag should be interpreted as *other lags held constant*.

As proposition 2.2 shows, a sugarcane plantation managed in the manner of the theoretical model in section 2.7 will take  $\min(\lceil \frac{1}{R} \rceil, S)$  years to reestablish a stationary state after a shock to the replant rate. For the study region in Brazil, the replant rate varies between 5.7 and 12 percent (see figure 2.8b), implying that the time to equilibrium, and hence the number of lags of replant rate that affect current yield, may be anywhere between 9 and 18 years if  $S$  is not binding. However, since I have observed no data suggesting that Brazilian sugarcane is cultivated beyond the 7<sup>th</sup> year, I speculate that  $S$  is binding, with a value of approximately 7.

Under the hypothesis that the theoretical model is correct, the sign predictions from figure 2.7b will hold for the econometric equation. I expect the lag of replant rate from  $t$  years ago to have the same impact on current yields as the impact of a change in the replant rate now on yields  $t$  years in the future. For example, if an increase in the replant rate now increases yield in 2 years time, I would expect an increase in the replant rate two years ago to increase current yield. In particular, I expect the coefficient on contemporaneous replant rate to have a negative effect on current yield, the coefficient on the first lag of replant rate to have a negative coefficient close to zero, and the coefficients on the remaining lags to be positive and increasing to a magnitude similar to the absolute value of the coefficient on the contemporaneous replant rate.



## Serial correlation in replant rates

Replanting rates may exhibit serial correlation. The serial correlation may be positive, where a low rate last year may be followed by a low rate this year due to a persistent shock, e.g. credit constraints spanning multiple years. Alternatively, the serial correlation may be negative, where a low replant rate last year leads to a high rate this year to compensate for the previous low rate. However, this is not necessarily an issue for my regression; an issue arises if the *idiosyncratic errors*,  $u_{it}$ , display autocorrelation.

To test for the presence of serial correlation in the idiosyncratic errors I perform the Wooldridge test of serial correlation for panel data models, as implemented for the STATA software package by Drukker (2003).

The test requires a panel data model of the form:

$$(2.4) \quad y_{it} = \alpha + \mathbf{X}_{it}\beta_1 + \beta\mathbf{Z}_i\beta_2 + \mu_i + \varepsilon_{it}$$

with  $i \in \{1, 2, \dots, N\}$  and  $t \in \{1, 2, \dots, T\}$ . Here  $y_{it}$  is the dependent variable;  $\mathbf{X}_{it}$  is a  $(1 \times K_1)$  vector of time varying observations;  $\mathbf{Z}_i$  is a  $(1 \times K_2)$  vector of time-invariant observations;  $\alpha$ ,  $\beta_1$ , and  $\beta_2$  are  $1 + K_1 + K_2$  parameters to be estimated;  $\mu_i$  is the unit-specific fixed effect; and  $\varepsilon_{it}$  is the idiosyncratic error. Consistent estimation of the parameters requires that the idiosyncratic errors be uncorrelated over time, i.e.  $\mathbb{E}[\varepsilon_{it}\varepsilon_{is}] = 0$  for all  $s \neq t$ .

Wooldridge's method of testing for correlation in the errors uses the residuals from a first-difference transformation of the panel model in equation (2.4), that is:

$$\Delta y_{it} = \Delta \mathbf{X}_{it}\beta_1 + \Delta \varepsilon_{it}$$

Wooldridge observes that the residuals from this regression must satisfy  $\text{Corr}(\Delta \varepsilon_{it}, \Delta \varepsilon_{it-1}) = -0.5$ , the case if the idiosyncratic errors are not serially correlated.

The STATA implementation of Drukker (2003) reports the  $p$ -value from a test of whether the coefficient from a regression of the residual on its lag is equal to  $-0.5$ , with the null hypothesis being that the coefficient is equal to  $-0.5$ . The  $p$ -values for the alternative lag specifications, estimated with the Brazilian data, are presented in table 2.1.

If the assumptions underlying the panel data model hold, particularly that serial correlation is either not-present, or adequately controlled by the use of clustered standard errors, I can interpret the  $\beta$  coefficients as follows:  $\beta_l$  represents the marginal effect of a one unit increase in the replant rate  $l$  years ago on yields in period  $t$ , all else being equal.

## 2.9 Data

To measure the effect of changes in the replant rate on yields I created a dataset of sugarcane planted area, replanted area, and yields from the 2005/06 to the 2013/14 growing year in 30 mesoregions of the south-central sugarcane growing region of Brazil.

Mesoregions are a statistical (but not administrative) subdivision of Brazilian states. Created by the Brazilian Institute of Geography and Statistics<sup>7</sup> (IBGE – *Instituto Brasileiro de Geografia e Estatística*), the mesoregions attempt to subdivide the states into regions with similar “social processes”, conditioned by their “natural setting” and the degree of “communication and place network”. There are 136 mesoregions in Brazil. I focus on the South-Central region of Brazil, comprised of the states: Espírito Santo, Goiás, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Paraná, Rio de Janeiro, Rio Grande do Sul, Santa Catarina, and São Paulo. There are 74 mesoregions in the south-central region. In this dataset I used the states Goiás (GO), Mato Grosso (MT), Mato Grosso do Sul (MS), Minas Gerais (MG), Parana (PR), and São Paulo (SP), which accounted for over 99 percent of sugarcane production in the south-central region of Brazil in the 2014/2015 growing year.

Data for quantity of sugarcane produced, yield, and planted area were downloaded from the IBGE website on 4 Jan, 2017. The IBGE data included the planted area (hectares), production (tons), and average yield (kilograms/ha, which I converted to tons/ha), for each mesoregion in the South-Central region, by year. Each mesoregion in Brazil has a unique identifying code. These codes were not included in the downloaded data from IBGE, although the codes are available on their website. I manually added the mesoregion codes to the downloaded data.

The data on area replanted was obtained from the CANASAT project,<sup>8</sup> run by the Brazilian National Institute for Space Research (INPE – *Instituto Nacional de Pesquisas Espaciais*). The CANASAT project uses satellite data to classify sugarcane growing regions into one of four classes: ratoon, canes that are growing from established rootstock; expansion, area freshly converted from non-sugarcane use; under-renovation, canes that have been replanted, but not yet harvested; and renovated, the first harvest of freshly replanted canes. I downloaded this land use data for each of the six main sugarcane growing states in the south-central region. The data was provided at the municipality level, an IBGE statistical division two levels smaller than the mesoregion. Using a dictionary linking municipality codes to mesoregion codes, provided by Peter Johannessen<sup>9</sup>, I calculated the totals of the four classes of land use in each mesoregion.

I merged these two datasets in STATA, dropping any mesoregions that did not have a positive total-cultivated area each year, resulting in a balanced panel of 270 observations across 30 mesoregions and 9 years, running from harvest year 2004/05 to harvest year 2013/14. I dropped 2004/05 because the area replanted was not reported in all states except São Paulo, since that was the year monitoring began for those states. In 2013/14 the total production from these 30 mesoregions was 668 million tons. Total production in Brazil that year was 768 million tons. So these 30 mesoregions represent 87 percent of Brazil’s total sugarcane production in 2013/14.

---

<sup>7</sup>[http://www.ibge.gov.br/english/geociencias/geografia/default\\_div\\_int.shtm?c=1](http://www.ibge.gov.br/english/geociencias/geografia/default_div_int.shtm?c=1) Accessed: 10 Mar, 2017

<sup>8</sup><http://www.dsr.inpe.br/laf/canasat/en/tables.html> Accessed: 25 August, 2014

<sup>9</sup><http://www.peterjohannessen.com/data/> Accessed: 5 Jan 2017

## 2.10 Results

Figure 2.8a shows the area-weighted average yield and figure 2.8b shows the area-weighted percent replanted across the 30 mesoregions in the sample for the years 2005 to 2013. For each year in the sample, a weight was assigned to each mesoregion, representing the proportion of total cultivated area that mesoregion provided over the entire sample, that is:

$$\text{weight}_i = \frac{\sum_t \text{area}_{it}}{\sum_t \sum_i \text{area}_{it}}$$

This weighting scheme assigns more importance to the mesoregions that accounted for a larger share of the land use.

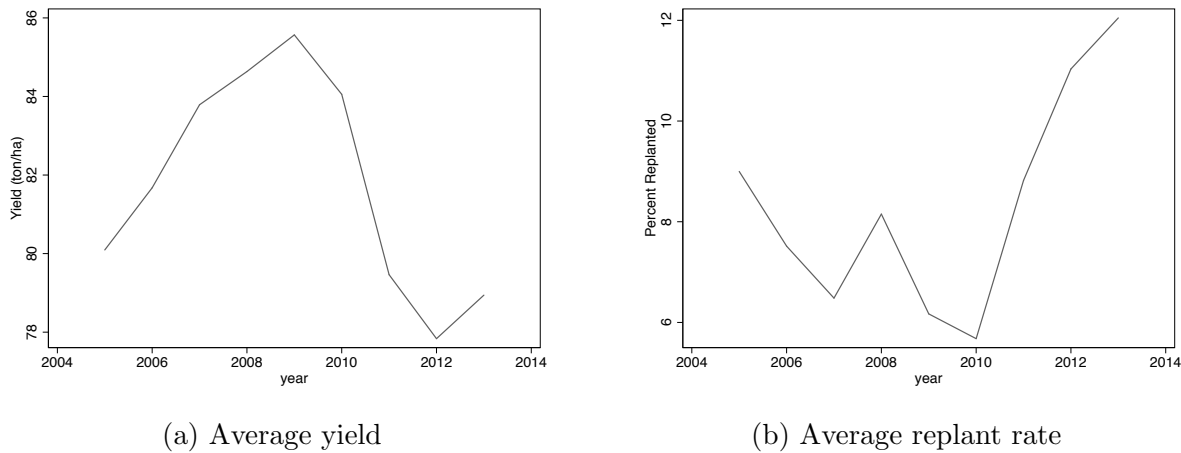


Figure 2.8: Average yield and replant rate across the 30 mesoregions in the sample.

In the figures there is a roughly inverse relationship between yield and percent replanted. Yield increased from 2005 to 2009 from a minimum of 80 tons/ha to a maximum of 85 tons/ha. From 2009 to 2012 yields declined to 78 tons/ha before recovering slightly in 2013. Inversely, for replant rates, I see percent replanted declining to a minimum in 2010, before rapidly increasing from 2011 to 2013. The two years with the lowest replant rate occurred in 2009 and 2010, the years most affected by the credit crisis, and before the government allocated funds to field renewal, beginning in Jan, 2012 (Mendonça et al., 2013). Although 2009 and 2010 were the two lowest years in the sample, they were not much lower than the renewal rate in 2007, and about 2.5 percentage points lower than the mean of 2005-2008. Renewal rates began increasing in 2011, before the government rescue package was launched.

Table 2.1 shows the results from six alternative specifications of the regression equation (equation 2.3), and figure 2.9a provides a graphical representation of the replant rate coefficient estimates presented in table 2.1.

	No Lag	1 Lag	2 Lags	3 Lags	4 Lags	5 Lags
% Replanted	-0.5968***	-0.5384***	-0.5335***	-0.5280***	-0.4841***	-0.1630
% Replanted - Lagged one year		-0.0441	-0.0110	0.0061	0.0586	-0.1575
% Replanted - Lagged two years			0.1995*	0.3557***	0.4076**	0.5309***
% Replanted - Lagged three years				0.0140	0.1872	0.1402
% Replanted - Lagged four years					0.2506	0.5154
% Replanted - Lagged five years						0.4998*
Area Planted (1000 ha)	-0.0046	-0.0095*	-0.0169*	-0.0279	-0.0264	-0.0228
Mean of Dep Variable	81.78	82.00	82.04	81.75	81.17	80.07
R-squared	0.18	0.22	0.30	0.35	0.35	0.44
Autocorrelation <sup>†</sup>	0.000	0.000	0.000	0.000	0.001	0.288
<i>N</i>	270	240	210	180	150	120

Standard errors clustered at the mesoregion level

<sup>†</sup> *p*-values of Wooldrige serial correlation test where  $H_0$ : No serial correlation (see Drukker (2003))

Mesoregions weighted by their average share of cultivated area

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 2.1: Results from estimating equation (2.3) with the Brazilian dataset using 0–5 lags of replant rate

Combining the observed renewal rates with proposition 2.2 implies that equilibrium yield would be reached in this system after between 9 and 18 years (or fewer if  $S < 9$ ). Unfortunately there are 9 years in the sample, so there is a trade-off between the theoretical desirability of more lags, and the resulting loss in sample size that comes from including them. Including no lags gives me 270 observations to work with, while including 5 lags reduces the number of observations to 120. Hence I present multiple variations on the model to show how the coefficient estimates vary with the number of lags.

I include 6 models, varying the number of lags,  $L$ , from 0 to 5. In the models with 0 to 4 lags, the Wooldridge autocorrelation test results imply I can reject the null hypothesis of no autocorrelation at the 1 percent significance level, indicating the presence of autocorrelation in these models. To control autocorrelation in all 6 models, I calculated clustered standard errors, with the mesoregion used as the unit of clustering. Clustering the standard errors allows for an arbitrary correlation structure within the cluster, accommodating the autocorrelation detected by the Wooldridge test. Clustering at the mesoregional level still maintains the assumption of independence of errors between the mesoregions. However, since I expect some of the farmers' replant rate choices to be influenced by state- and national-level factors (e.g. the credit crisis), this independence assumption is unlikely to hold in practice. Hence I can only take the standard error estimates as a lower bound—the actual error is likely to be larger.

The coefficient on the contemporaneous replant rate is negative in all the models, and significant at the 1 percent level in all but the 5 lag model. In most cases it is around -0.5, implying that a 1 percentage point increase in the replant rate decreases yields by approximately 0.5 tons/ha in the same year. In almost all the models the coefficients on the lags are positive, the exceptions being the coefficient on the first lag in the 1, 2, and 5 lag models. In each of these cases the coefficient is not significantly different from zero. The coefficient on the first lag is a tight zero in each of the 5 models that include it. The coefficient on the second lag is positive in 3 of the 4 models that include it. Coefficients on lags 3–5 have larger standard errors, although the results suggest that the coefficient magnitudes are either constant or returning to zero after the peak at 2 lags.

In each model the coefficient on area planted was negative, but also statistically indistinguishable from zero at a 5 percent significance level. The  $R^2$  values for the models ranged from 0.18 to 0.44, naturally increasing as more lags were added. The higher lag models (models 2–5) explained around one third of the variation in sugarcane yield during the sample, implying that there are other, omitted factors that play an important role in explaining sugarcane yields in the south-central region of Brazil. The  $R^2$  results reported from the regressions are the *within*  $R^2$  values.

## 2.11 Discussion

Figure 2.9b shows the theoretical prediction from figure 2.7b and the estimated coefficients for each of the 6 regression models tested. Comparing the regression coefficient estimates

from the 6 models to the theoretical predictions I see a striking consistency between them. Generally, the theoretical prediction is within the 95 percent confidence interval for most of the coefficients from most of the models. For the first three coefficient estimates (no lag, 1st lag, and 2nd lag) the theoretical prediction is within the 95 percent confidence interval of all but one of the the coefficient estimates, the exception being the no-lag coefficient from the 5-lag model, which is lying closer to zero than the theoretical prediction would place it. For the first three coefficients, their point estimates are generally higher than the theoretical prediction, although the prediction lies with in the 95 percent confidence interval. For the second three coefficients, the results are less strong, with the theoretical predictions falling outside, or close to the edge of, the 95 percent confidence intervals of the estimated coefficients. In each case the theoretical prediction is higher than the point estimate for each of the coefficients. A possible reason for the greater discrepancy between the predictions and the estimates for the higher lags is the smaller sample sizes that each of these models used. Adding an additional lagged variable reduces the sample size by 30 observations. So the no-lag model has 270 observations, while the higher lag models have only 180–120 observations to work with, reducing the precision of the estimates.

Figure 2.10 compares the actual average yield across the 30 mesoregions of the sample against the predicted yield from the 4-lag variant of the model. This graph shows that changes in the replant rate explain a substantial share of the yield variation, but clearly not all of it. This is reflected by the  $R^2$  value of 0.35 for the 4-lag model. Although my model does not seek to directly explain the yield share variation in section 2.3, it is plausible for the model to explain a substantial fraction of the variation in the yield share. Because of the indirect nature of the measurements, the best I can conclude is that the replant rate channel influenced a substantial share of the yield variation, but I cannot quantify the actual size of the share.

Figure 2.10 was generated using the Margarido and Santos (2012) age-yield relationship shown in figure 2.5. I assumed that the final age-class was 6-years-since planting cane ( $S = 6$ ) since data was not available for older age-classes. If the oldest age-class was older than 6 and less productive than 6-year-since-planting cane, the initial decline would be smaller and the remaining lags would be higher. Essentially it would shift upward the entire *net* change curve (an upward shift of figure 2.7b).

Including lagged values of the replant rate reduces the sample size. Including 4 lags prevents me from making predictions towards the beginning of the sample. The prediction begins in 2009. Overall, the prediction roughly captures the yield drop from 2009 to 2012, but does not completely capture the magnitude of the decline. Actual yields at the peak in 2009 (85.6 ton/ha) were higher than the prediction (83.6 ton/ha) by 2 tons/ha. Likewise, actual yields at the trough in 2012 (77.8 ton/ha) were lower than the prediction (79.2) by 1.4 tons/ha.

I chose the 4-lag model to make the prediction because it is the only one of the higher lag models (3–5 lags) that is consistent with the theoretical predictions for each of its coefficients. The 3-lag model’s coefficient on the 3-lag variable is significantly different from the theoretical prediction, while the 5-lag model’s coefficient on the no-lag coefficient is significantly different

from the theoretical prediction.

Although this is the best prediction I can produce with this data, it may not be completely accurate. In particular, the average age-yield relationship across the region may be different. Also, the model allows the age-yield relationship to vary across the regions only by a mesoregion specific scalar, i.e. all mesoregions have an age-yield relationship with the same relative differences between the age-classes, but the level of all the age-classes is shifted up or down by a common factor. If the age-yield relationships across the mesoregions have different relative differences between the age-classes, the regression equation only captures the average of these individual age-yield relationships. This makes predictions from the regression model valid for the sample region as a whole, but not for specific, individual mesoregions.

As discussed in section 2.6, this analysis cannot determine the quantitative effect of the credit crisis on the sugarcane industry. In that section I established the possibility of the credit crisis affecting yield, and in subsequent sections examined changes in the replant rate as a plausible channel for the effect. To quantitatively examine the effect of the credit crisis I would need unit level measures of credit constraints. Briggeman et al. (2009) provide one method for obtaining such data, a direct survey of farmers. This is left for future work.

My theoretical analysis treats yield as a function of replant rate, all else being equal. This assumption may not hold for the econometric analysis. The econometric analysis studies the effect of a replant rate change, holding constant the total cultivated area, other lags of replant rate (for those included in the model), and mesoregion specific fixed effects, such as soil quality. However there are other variables that may affect the yield that I do not control for. Some of the omitted variables include weather, input use, harvesting method (if it changes during the sample period), sugarcane variety, and pest damage. If any of these variables are unchanging over time, they will be captured by the mesoregion fixed effects. The components that are changing over time may bias the coefficient estimates, if they are systematically correlated with the replant rate.

I identified the replant rate as one channel for the credit crisis to affect sugarcane yields, but it is not necessarily the only channel. Any component of the growing process that both relies on credit and affects yields could conceivably create this link. For example, I would expect fertilizer use to be influenced by credit availability and to influence yields. As Mendonça et al. (2013) said: “[during the credit crisis] companies stopped investing in, for example, the renewal of sugarcane plantations, crop handling and fertilization, which are needed to maintain high production levels.” Unless there are inputs to production that increase yield when credit is restricted, I expect these omitted variables to reduce yields in the absence of credit. So the regression results and yield predictions represent a lower bound on the magnitude of the effect of the credit crisis on sugarcane yields.

The theoretical model does not account for changes in the area of the sugarcane growing region. This is a possible source of specification error since additions or subtractions to the total area can affect the age-structure of the region. For example, consider a 1 ha expansion of a 100 hectare growing region, which will then move through the age-classes following the dynamics of system. This is not equivalent to a 1 percent increase in the replant rate since the expansion increases the size of the growing region. The specification of the

econometric equation includes total area as a control, but this is an imperfect measure of the effect of expansion since it does not track the movement of the new land through the age-structure—it assumes a shift of the entire age-yield relationship, rather than accounting for the non-monotonic relationship I identified earlier. Examining the exact effect of, and the specification error caused by, expansion is left for future work.

An additional source of error in the theoretical model is that I assume all replanting is done simultaneously at the start of the period, and then all harvest occurs simultaneously at the end of the year. In reality harvest occurs as a flow throughout the harvest season, which occurs from April through November in the South-Central region, and replanting can take place from October to July the year too.

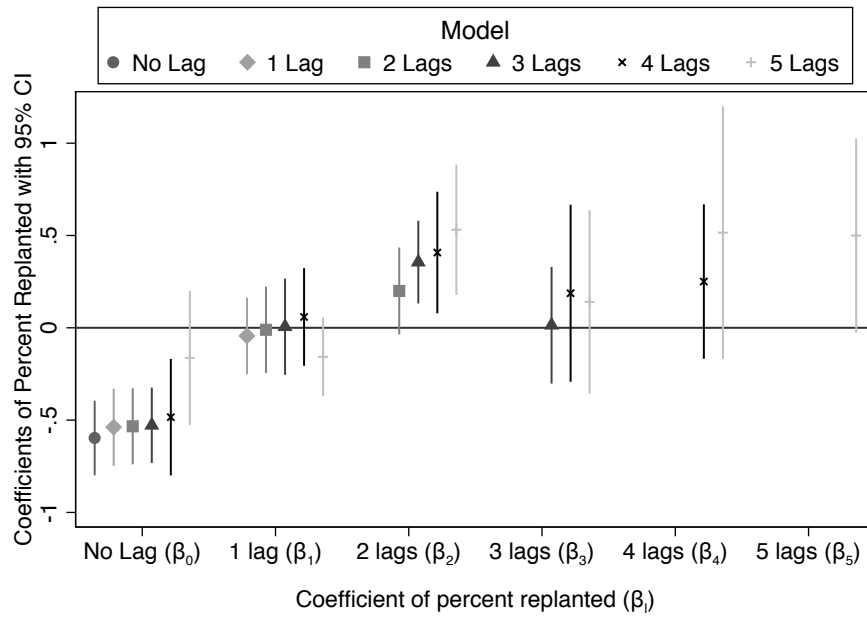
## 2.12 Conclusion

I show that the financial crisis had a plausible mechanism, field replanting, to cause the slowdown in Brazil’s sugarcane production, and that replant rate changes explain about one third of the variation in yield from 2009 to 2013, implying that it played a substantial, but incomplete, role in Brazil’s sugarcane production slowdown. Further, I developed a model that predicts future perennial crop yields as a function of changes in the replant rate—it is a general model, applicable to a wide variety of perennial crops.

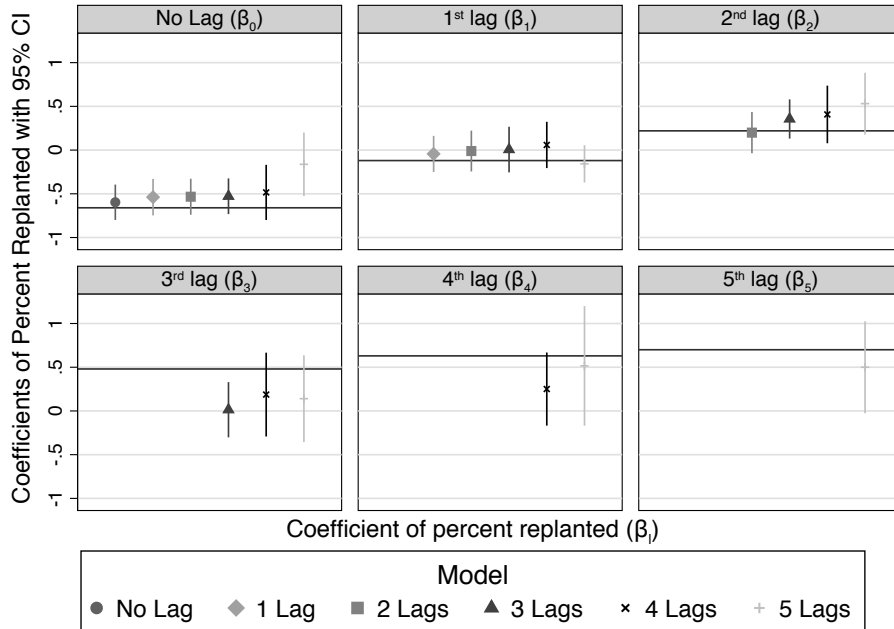
My analysis was limited along three main dimensions. First, due to data limitations, I was not able to quantitatively determine the effect of the credit crisis on sugarcane production. Second, I did not include weather in the econometric model, preventing a comparison of the relative effects of weather and changes in the replant as possible sources of yield variation. And third, the model does not explicitly account for expansions of sugarcane’s cultivated area, another source of new canes with the potential to affect future yields.

I leave for future work the task of quantifying the link between credit constraints and sugarcane replanting, as well as an extension of the theoretical model to explicitly include the effect of expansion of the cultivated area on the trajectory of future yields.





(a) Coefficient estimates and 95 percent confidence intervals from all 6 models



(b) Coefficient estimates compared to the theoretical predictions in figure 2.7b

Figure 2.9: Two views of the coefficients of the 6 models. Figure 2.9a shows the coefficient estimates relative to zero. Figure 2.9b shows the coefficient estimates relative to the predictions from figure 2.7b.

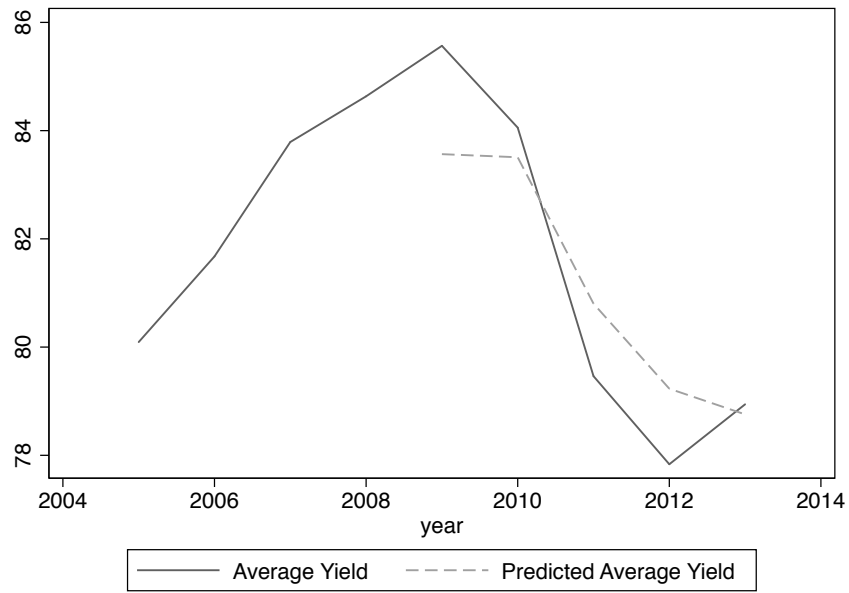


Figure 2.10: Actual and predicted yields for the 30 South-Central mesoregions in my sample using the 4-lag variant of the regression model.

## Chapter 3

# Optimal Balanced Age-Structure in a Perennial Crop Field-Refinery Supply Chain

### 3.1 Introduction

I show that the production costs of a perennial crop field-biorefinery supply chain can be reduced by optimizing the age-structure—the proportion of plants of different ages—of the crop.

A perennial crop field-biorefinery supply chain is an agricultural production system that begins with a feedstock being grown by a perennial crop. The feedstock is then harvested and transported to a local biorefinery where it is processed into a commodity and sold to downstream buyers. Imagine an apple orchard where the apples are picked, shipped to a local processing facility, and then sold downstream as canned apples and applesauce. In this chapter I focus on the example of ethanol and sugar production from a sugarcane feedstock. The key features to keep in mind are: the feedstock is grown by a perennial crop, that is, the crop yields feedstock for multiple years before needing to be replaced, so the yield of a unit of land is a function of the age-structure of the planted trees; and, the feedstock must be processed by a local biorefinery/processing plant before it becomes an easily tradable commodity.

There are numerous examples of perennial crop field-biorefinery supply chains providing economically important products. As the previous two examples illustrated, these products include both food and fuel. In the realm of food there are examples such as: wine, coffee, cocoa, palm oil, tinned or juiced tree fruit (e.g. citrus, pome fruit), and tree nuts.

Although this framework can be applied to many valuable food products, my focus in this chapter is on perennial field-biorefinery supply chains for biofuel production. Biofuels have great potential to assist efforts in climate change mitigation due to their lower life cycle carbon emissions in comparison to gasoline (Khanna and Crago, 2012). Biofuels are not

created equal in their carbon benefits. Feedstocks such as sugarcane and cellulosic ethanol, two of the lowest carbon biofuel feedstocks from a life cycle emission perspective (Crago et al., 2010), are perennial crops.

There is great interest in minimizing the costs of perennial crop field-biorefinery supply chains, especially in the biofuel context. The degree to which biofuels enter the global energy mix depends on the extent to which barriers to sustainable and efficient production can be overcome (De Meyer et al., 2014). The cost of the supply chain to convert biomass into energy is one of the most important barriers to adoption (De Meyer et al., 2014; Rentizelas et al., 2009). In addition to minimizing feedstock costs, biofuel supply chain optimization also aims to minimize the environmental impacts of the supply chain, and to ensure continuous feedstock supply (Mafakheri and Nasiri, 2014; Sharma et al., 2013). De Meyer et al. (2014) state that “to overcome all these barriers and uncertainties *biomass supply chain optimization* [emphasis in original] is essential.”

Feedstock supply has two key components: planted area, and yield per unit of land. Existing studies of biomass supply chain optimization focus on optimizing the planted area (De Meyer et al., 2014). These papers focus on the area and the location of land to grow the feedstock to supply a local biorefinery, or network of refineries. Mostly they hold yield per unit of land constant, although they may allow for heterogeneity between parcels of land. A notable exception to the fixed yield assumption is the study by Debnath et al. (2014) who solve for the cost-minimizing land area to feed a fixed biorefinery size when yields are subject to stochastic weather shocks.

These studies exclude the fact that the yield of perennial feedstock is a function of its age-structure. Economists have studied perennial age-structure optimization and have shown that the net present value of a perennial tree is a function of the discounted flow of benefits over the life of the tree and the tree replanting cost, e.g. (Wesseler, 1997). The optimal replacement rule, which states that the replacement age should be chosen to maximize the discounted sum of the per-period benefits over an infinite horizon (Mitra et al., 1991, prop. 3.1), is analogous to the Faustmann rule for optimal tree replacement in forestry, but in the case where there are only existence benefits and no direct benefit to cutting the tree (Brazee, 2007).

In this chapter I present a theoretical model of a perennial crop field-biorefinery supply chain with both endogenous planted area and age-structure choice. To my knowledge, the existing theoretical results and empirical studies do not incorporate the interrelationship between the age-structure and the biomass supply chain, hence this is the first study that incorporates an endogenous age-structure into a model of a perennial field-biorefinery supply chain.

This analysis focuses on a particular subset of possible age-structures: the balanced orchard. A balanced orchard has an equal proportion of land allocated to each age-class of tree (see Mitra et al. (1991) for a more general discussion of possible age-structures). For example, if there were only young and old trees, a balanced orchard would have half young trees and half old trees. The balanced orchard has the advantage of being the age-structure-induced-supply-variation minimizing age-structure (Tisdell and De Silva, 1986). This is an

explicit goal of biomass supply chain optimization (Mafakheri and Nasiri, 2014; Sharma et al., 2013; Margarido and Santos, 2012; Debnath et al., 2015)

I use this model to show that optimal feedstock age-structure is older than the yield-maximizing age-structure (in contrast to Tisdell and De Silva (1986), who suggest using the maximum yield age-structure in an orchard-only context). I also show that optimal maximum orchard age declines with the scale of the biorefinery, and that perennial field-biorefinery supply chain costs can be reduced by endogenizing yield. I present analytical formulas which explain the trade-offs inherent between age and land. The majority of existing models of biofuel supply chain optimization are mathematical programming models, which, while useful for analyzing policy scenarios, are less useful for illuminating the general trade-offs along the supply chain.

I calibrate the model to the sugarcane ethanol industry in the Brazilian state of São Paulo. My simulations show results that are consistent with the model’s theoretical predictions, but I find that the cost savings from including the age-structure optimization are small. In this case fixing feedstock age at the maximum yield age is a good rule-of-thumb. The empirical relevance of the age-structure optimization channel depends on the particular supply chain in question.

### 3.2 An Analytical Model of Orchard Age, Growing Region Area, and Biorefinery Capacity

Consider a biorefinery of given size that is supplied a feedstock grown by a perennial crop in surrounding fields. Assume that this is a vertically integrated system where the biorefinery manager also controls the land and crop management for the feedstock. The manager’s problem is to minimize the cost of the feedstock to the biorefinery<sup>1</sup> by choosing how much land to use, and how the feedstock crop is grown on that land. I pose this as a static problem for the manager, which can be interpreted as the long-run, steady-state management strategy for the biorefinery. In this study I neither study the short run dynamics of the manager’s problem, nor the choice of biorefinery size in the first place.

The cost minimization problem facing the manager<sup>2</sup> is

$$\min \left[ \frac{\text{Farm gate}}{\text{feedstock costs}} + \frac{\text{Feedstock}}{\text{delivery costs}} + \frac{\text{Processing}}{\text{costs}} \right] \quad \text{such that} \quad \frac{\text{Feedstock}}{\text{production}} = \frac{\text{Biorefinery}}{\text{capacity}}$$

I discuss each of these components of cost in turn to develop a mathematical statement of the manager’s objective function.

---

<sup>1</sup>This is equivalent to maximizing the profit of the biorefinery with a given output price and no investment costs.

<sup>2</sup>Wright and Brown (2007) observe that there are three components to the cost of producing processed product: feedstock cost at the farm gate; feedstock delivery costs; and biorefinery operating costs.

## Feedstock production

The biorefinery requires feedstock for processing. Call the quantity of feedstock arriving at the biorefinery  $Q$ . What are the factors determining  $Q$ , and what is the cost of producing  $Q$ ?

Feedstock production  $Q$  is the product of land planted with trees,  $L$ , and per-unit land productivity,  $y$ , i.e.  $Q = yL$

### Land productivity, $y$

Since the feedstock is perennial, the age-structure of the orchard<sup>3</sup> determines its productivity. The productivity of a single tree varies over its lifespan, so the total productivity of the orchard is the weighted sum of the productivities of all the constituent trees. Let  $f(a)$ —called the age-yield-relationship—be the yield per land unit of  $a$ -year-old trees.

The age-structure of the orchard through time can exhibit many different trajectories (see Mitra et al. (1991) for more discussion), but I restrict this analysis attention to a special type of trajectory: the balanced orchard. In a balanced orchard the distribution of tree ages follows a uniform distribution from 0 to maximum tree age,  $n$  (Tisdell and De Silva, 1986). The density of each age of tree is thus  $\frac{1}{n}$ . I call this an  $n$ -orchard. I call an *older* orchard one with a larger  $n$ , and a *younger* orchard one with a smaller  $n$ .

The balanced orchard is the supply-variation minimizing steady-state age-structure. There are two reasons to focus on balanced orchards. First, balancing an orchard to minimize supply variation is frequently a direct management objective for perennial crop growers (Tisdell and De Silva, 1986; Margarido and Santos, 2012, for sugarcane). Second, it allows me to write a simple model that can focus directly on the trade-offs between age-structure, land, and biorefinery capacity, while avoiding the technical details of transition dynamics.

Focusing on steady state orchards, however, prevents me from analyzing the transition dynamics to the steady state, and between steady states, so this analysis must be considered a long run analysis. It is an open question whether or not an optimally managed perennial crop supply chain, starting from an arbitrary initial condition, has the balanced orchard as its long run steady state. The answer to this question is beyond the scope of this study, and I simply assume that the balanced orchard is a good approximation of the long run steady state.

---

<sup>3</sup>I use the term orchard here as a metonym for all perennial crops, including tree crops and perennial grasses.

## Allowable age-yield relationships

I impose the following conditions on the age-yield relationship to ensure a solution

$$(3.1) \quad f(a) \text{ is continuous}$$

$$(3.2) \quad f(0) = 0$$

$$(3.3) \quad f(a) \text{ monotonically increases to a unique maximum and then monotonically declines}$$

$$(3.4) \quad \lim_{a \rightarrow \infty} a f(a) = 0$$

Assumption (3.1) aids analysis of the field-biorefinery optimization problem. Although continuity can only ever be an approximation of the true age-yield relationship, I consider it to be a reasonable assumption, and that it is a worthwhile price to pay to facilitate analysis. Assumption (3.2) requires that plants are non-yielding when they are planted. This is a reasonable assumption when considering the entire life-cycle of a plant since it would be most unusual for a plant to yield fruit (or sugar etc.) immediately when it first sprouts. This assumption may be more problematic when considering the problem from the perspective of the field-biorefinery manager only. It may be possible for the manager to buy saplings that are bearing fruit when he takes possession of them. I exclude this possibility. Assumption (3.3) is similar to a standard assumption in the perennial crop theory literature (e.g. Mitra et al. (1991)), but it is a little stronger, since it excludes the possibility that trees may have a maturity phase where they produce their maximum yield for several years in a row. This assumption, however, allows the age-yield relationship to become arbitrarily close to this case. Assumption (3.4) requires that the age-yield relationship approaches zero 'fast enough' as the age of the tree approaches infinity. In particular the assumption requires that the age-yield relationship approach zero faster  $\frac{1}{x}$ . This is clearly a reasonable assumption since the oldest known tree is around 5000 years old. However, this assumption imposes two important modeling restrictions on the analyst. First, the age-yield relationship cannot approach a positive constant. This may be an attractive assumption if the tree has a long period of relatively constant yield toward the end of its life. Second, it rules out age-yield relationships that approach zero too slowly. The analyst must check any candidate age-yield relationship against these assumptions before relying on the results of this study.

## Feedstock production from an $n$ -orchard

For an orchard with maximum age  $n$ , the yield of feedstock per unit of land is

$$y(n) = \frac{1}{n} \int_0^n f(a) da$$

There is a trade-off between marginal and average yield inherent in this function. Since  $f(a)$  is non-negative, the integral term is increasing in  $n$ . But an increase in  $n$  increases the number of age classes that the yield must be averaged across. Whether  $y$  is increasing or decreasing in  $n$  depends on the contribution of the marginal tree relative to the average at that  $n$ , which is shown by the derivative of  $y$  with respect to  $n$ .

$$\frac{dy}{dn} = \frac{1}{n} \left( f(n) - \frac{1}{n} \int_0^n f(a) da \right) = \frac{1}{n} \left( \underbrace{f(n)}_{\text{Yield of additional old tree}} - \underbrace{y(n)}_{\text{yield of } n\text{-orchard}} \right)$$

The terms in the parentheses are multiplied by  $\frac{1}{n}$  because the contribution of a single old tree (and every other age-class  $a < n$ ) declines with the number of age-classes in the  $n$ -orchard.

Let the maximum age that represents the age structure that maximizes yield be  $n_{MY}$ .

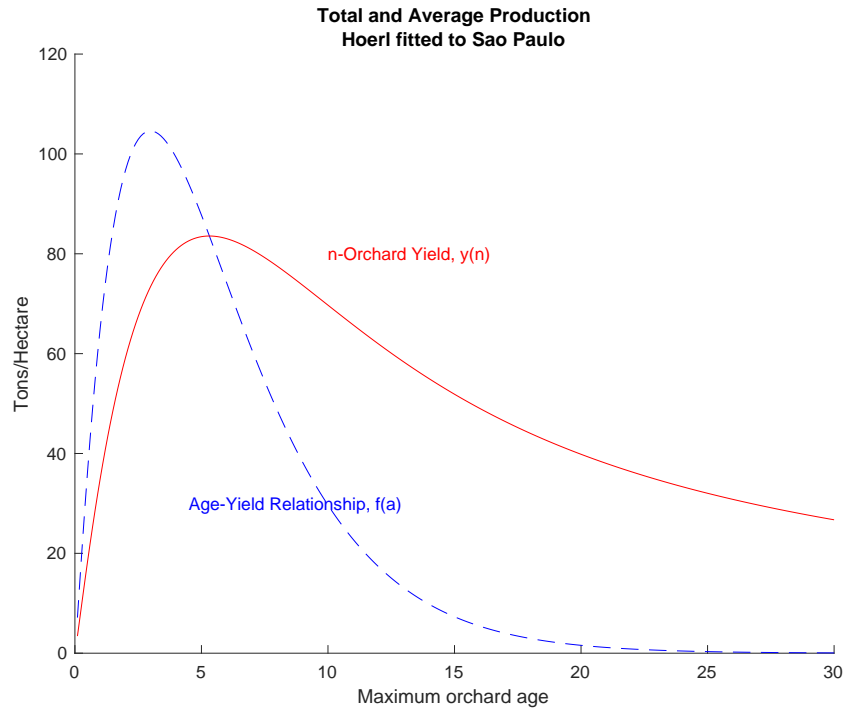


Figure 3.1: Yield is increasing in  $n$  while the marginal age class,  $n$ , is more productive than the average of the  $n$ -orchard,  $y(n)$ . Here the age-yield relationship is a Hoerl function fitted to the Brazilian yield data from Margarido and Santos (2012).

### Using a Hoerl function for $f(n)$

Haworth and Vincent (1977) undertook a study of the statistical modeling of perennial crops. In chapter 3 they discuss the merits of different functional forms for fitting perennial crop



age-yield data, arguing that the Hoerl function is the most appropriate function for modeling perennial crop yields.<sup>4</sup>

The Hoerl function is given by  $f(x) = ax^b e^{cx}$ . A desirable feature of this function is that its log is linear in parameters, which facilitates estimation by ordinary least squares. To model the properties of a perennial crop age-yield relationship I must assume  $a > 0$ ,  $b > 0$ , and  $c < 0$  (which is necessary and sufficient to satisfy assumptions (3.1)-(3.4)). I show this in lemma B.1 in appendix A.1). The age-yield function in figure 3.1 is an example of the Hoerl function. In this case the parameters have been fitted to age-yield data based on data for sugarcane grown in the Alta Mogiana region of São Paulo state, Brazil (Margarido and Santos, 2012) (see appendix A.2).

### Limitations of the Hoerl Function

The Hoerl function awkwardly fits the observed age-yield relationship for Brazilian sugarcane. The fitted Hoerl function peaks later than the observed yield peak, with the fitted Hoerl peak occurring around an age of 3.5 years, while the observed peak occurs at age 2 (see figure A.1 in appendix A.2).

Vincent and Haworth's argument in favor of the Hoerl function was based on data for apples, pears, peaches, and oranges, all trees with lifespans over 15 years. The non-bearing period of these trees is much longer than for sugarcane (3-5 years, compared to 1 year), and the rate of increase to maximum yield is much slower (maximum yield reached in 10-20 years, compared to 1 year for sugarcane).

It is possible, therefore, that some other functional form, satisfying the assumptions in section 3.2, is a superior choice for modeling sugarcane age-yield relationships. Searching for this relationship, however, is beyond the scope of this study, and the Hoerl function is sufficient for making the qualitative arguments in this chapter. Further, by making a simple adjustment to the age-yield data, I can obtain a far superior, but still plausible, fit with the Hoerl function.

### Replanting costs

Replanting costs are a substantial cost in perennial crop production. In an example sugarcane farm budget prepared for sugarcane growing in the South-Central region of Brazil, Teixeira (2013) found that replanting costs accounted for around 25 percent of the annual operating cost of a 5-orchard. If  $C_n$  is the cost of replanting trees on a unit of land,  $\frac{C_n}{n}$  is the annual replanting cost of an  $n$ -orchard, since only  $\frac{1}{n}$  of the trees are replanted each year.

---

<sup>4</sup>Specifically Haworth and Vincent (1977) compared quadratic, log-quadratic, log-reciprocal, Hoerl's function, the modified Gompertz function, and the generalized logistic function. While the generalized logistic function had the most flexible fit, the results from fitting the Hoerl function were indistinguishable from the generalized logistic. Further the Hoerl can be fitted with ordinary least squares, while the generalize logistic requires non-linear curve fitting techniques.

## Variable costs

I assume that the variable growing costs (net of replanting costs) do not depend on orchard age-structure. I am also assuming away the possibility for the manager to change yield by changing the use of variable inputs. Hence the variable cost of growing a feedstock on a unit of land is a constant,  $C_f$ . This includes variable inputs such as fertilizer, pesticides, and labor. It also includes the land rental rate.

## Total land, $L$

The other component determining total feedstock quantity is the area of land controlled by the manager,  $L$ . The choice of  $L$  determines how many units of land have a perennial feedstock orchard with yield  $y(n)$  on them. This determines total feedstock production,  $Q = y(n)L$ , and total feedstock growing costs,  $L(C_f + \frac{C_n}{n})$ .

## Delivery costs

The total quantity of land also affects the cost of transporting the feedstock from the farm gate to the biorefinery (Wright and Brown, 2007). Delivery costs are proportional to the quantity of feedstock multiplied by the average delivery distance.

The average delivery distance is increasing in the area of land around the biorefinery. In the case of a biorefinery surrounded by a circular growing region, following Overend (1982), the distance from the biorefinery to the furthest field is given by

$$L = \pi r_{max}^2 \Rightarrow r_{max} = \sqrt{\frac{L}{\pi}}$$

The area-weighted average delivery distance is  $r_{av} = \frac{2}{3}r_{max}$ . I express delivery costs as  $C_D y(n)L^\alpha$  (or  $C_D Q L^{\alpha-1}$ ), where  $\alpha > 1$ . Hence delivery costs are a convex function of growing region area. If the growing region is not circular,  $\alpha$  is not necessarily equal to  $\frac{3}{2}$ , but delivery costs still increase as a convex function of land (see appendix A.2 for full derivation).

I make a distinction between the area of land planted with orchards and the total area of the growing region. To allow for the possibility that some land in the growing region is used for other purposes, I allow the planted area to be a linear function of total growing region area,  $L = d \times A$  where  $A$  is the total growing region area, and  $0 < d \leq 1$ . The effect of this distinction on deriving the delivery cost function is shown in appendix A.2.

Bringing in another unit of land into the growing region increases both the quantity of feedstock produced, and the average distance all feedstock must be transported. The increase in feedstock quantity is linear (holding yield constant), and the increase in average delivery distance is proportional to a positive power of land, hence making the delivery cost function a convex function of growing region area.

## Processing costs

Since this analysis focuses on a static, deterministic setting, the biorefinery size can exactly match the level of feedstock production. Thus  $Q$  represents both the quantity of feedstock produced and the processing capacity of the biorefinery. Nguyen and Prince (1996) and Jenkins (1997) wrote two early, influential studies that suggest that plant operating costs are a concave function of biorefinery size. I thus write processing costs as  $C_p Q^\gamma$  where  $\gamma < 1$ .

## Objective function

Recall the manager's objective is to minimize the cost of feedstock production, given by:

$$\text{Feedstock costs} = \text{Farm gate feedstock costs} + \text{Feedstock delivery costs} + \text{Processing costs}$$

Using the notation and formulas developed in the previous section, I can rewrite the feedstock cost function mathematically

$$C(n, L) = \left[ \left( C_f + \frac{C_n}{n} \right) L + C_D y(n) L^\alpha + C_P (y(n) L)^\gamma \right]$$

where  $n$  is maximum orchard age,  $L$  is size of growing region,  $C_f$  are the growing costs independent of age-structure per unit of land,  $C_n$  are the age-structure dependent growing costs per unit of land,  $C_D$  is the delivery cost parameter,<sup>5</sup>  $\alpha$  is the measure of delivery cost convexity,  $C_P$  is the processing cost parameter, and  $\gamma$  is the measure of processing cost concavity.

## 3.3 Cost Minimization and Simulation

I now return to the manager's optimization problem, minimizing the costs of supplying a biorefinery of a given size ( $\bar{Q}$ ):

$$\min_{n,L} C(n, L) = \left[ \left( C_f + \frac{C_n}{n} \right) L + C_D y(n) L^\alpha + C_P (y(n) L)^\gamma \right] \quad \text{s.t. } y(n)L = \bar{Q}$$

Observe that I can substitute the biorefinery size into the expression for processing costs, leading them to become a constant relative to  $n$  and  $L$ , thereby reducing the cost minimization problem to one that only includes farm gate and delivery costs.

$$\min_{n,L} C(n, L) = \left[ \left( C_f + \frac{C_n}{n} \right) L + C_D y(n) L^\alpha \right] \quad \text{s.t. } y(n)L = \bar{Q}$$

The Lagrangian associated with this cost minimization problem is

$$\mathcal{L}(n, L, \lambda) = \left( C_f + \frac{C_n}{n} \right) L + C_D y(n) L^\alpha + \lambda (\bar{Q} - y(n)L)$$

---

<sup>5</sup>See appendix A.2 for details

## First order conditions

The three first order conditions for the cost minimization problem are

$$(3.5) \quad \frac{\partial \mathcal{L}}{\partial n} = \frac{-C_n L}{n^2} + C_D y'(n) L^\alpha - \lambda y'(n) L = 0$$

$$(3.6) \quad \frac{\partial \mathcal{L}}{\partial L} = (C_f + \frac{C_n}{n}) + \alpha C_D y(n) L^{\alpha-1} - \lambda y(n) = 0$$

$$(3.7) \quad \frac{\partial \mathcal{L}}{\partial \lambda} = \bar{Q} - y(n) L = 0$$

Equations (3.5) - (3.7) state that the marginal change in the Lagrangian function with respect to each of the choice variables must be zero at the optimum.

The left hand side of equation (3.5) shows how the Lagrangian function changes with respect to an increase in the orchard age. There are three components. The first component is the change in age-structure dependent costs (e.g. average replanting costs). This is always negative since the costs are averaged over more age-classes as  $n$  increases. The second component is the change in delivery costs due to the increase in orchard age. This can be either positive or negative depending on the sign of marginal yield,  $y'(n)$ . If marginal yield is negative, then an increase in orchard age reduces delivery costs since there is less feedstock to deliver. The third term is the penalty for violating the quantity constraint. If  $y'(n)$  is non-zero, a change in  $n$  changes the quantity of feedstock produced (since I am holding planted area constant). If the constraint was satisfied before the change, then it will now be violated after the change. Generally,  $\lambda$  represents the penalty for violating the constraint by a single unit (at the optimum it represents the change in costs due to a unit increase in biorefinery capacity). So the third term is the product of the per-unit penalty and the change in total feedstock quantity due to the increase in orchard age.

The left hand side of equation (3.6) shows how the Lagrangian function changes with marginal increase in the planted area. The three components have similar interpretations to the components of equation (3.5) except that now orchard age is being held constant. The first term is the marginal cost of growing feedstock on an additional unit of land. The second term is the marginal cost of delivery from the additional unit of land. This is an increasing function of total land due to the convexity of delivery costs. Again, the third term is the penalty for violating the capacity constraint, given the penalty per unit,  $\lambda$ .

## Isoquant and isocost curves

Figure 3.2 shows an example isocost and isoquant curve for the constrained minimization problem presented in the previous section. The isocost and isoquant functions are presented in  $(n, L)$  space, so both are functions of  $n$  here. The figure was generated with MATLAB

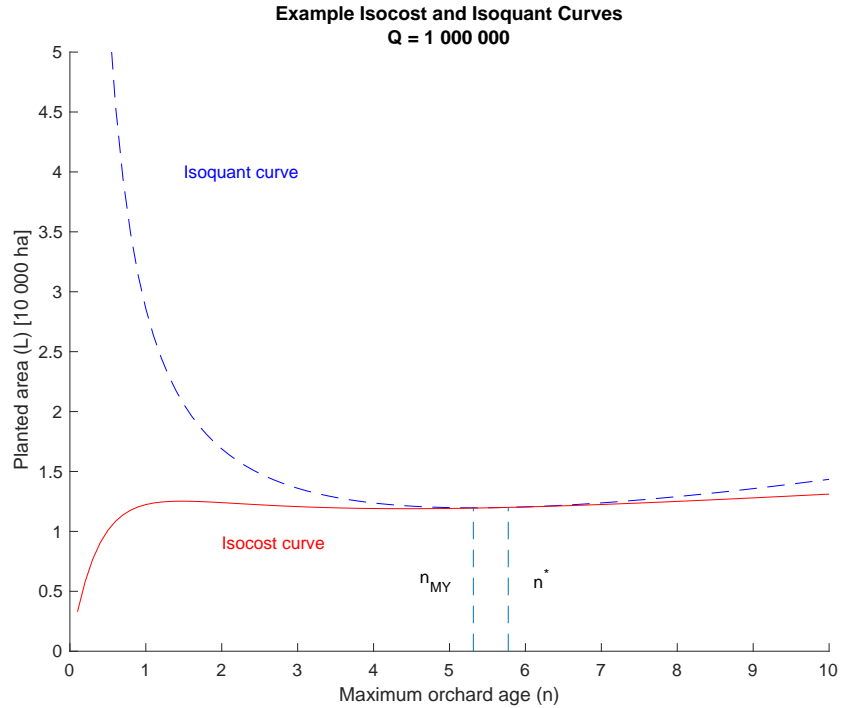


Figure 3.2: Example isocost and isoquant curve for the cost minimization model calibrated to São Paulo, Brazil. Details of calibration are in appendix A.2. Note that the lower-cost set is non-convex—there is a local maximum of the isocost curve around  $n = 1$  and a local minimum around  $n = 4$ .

using the parameterization described in appendix A.2. Although specific to the sugarcane ethanol industry in São Paulo state, Brazil, this figure displays all the qualitative features of an isocost and isoquant curve of the general constrained minimization problem, as the next sections establish.

Looking at figure 3.2, two issues arise regarding using the first order conditions to solve the constrained minimization problem. The first is that the domain of the isocost and isoquant function is unbounded to the right. The second is that the lower-cost set of the isocost curve is non-convex. These two issues mean that I cannot immediately invoke the usual sufficiency conditions for a convex optimization problem, which call for a bounded domain and a convex lower-contour set for the objective function and guarantee that a solution to the first order conditions is also a solution to the original optimization problem.

The first issue can be dispensed with almost immediately by noting that as  $n$  approaches infinity, yield approaches zero, requiring the growing region to increase without bound. The marginal cost savings from increasing  $n$  and reducing average fixed costs approach zero as  $n$  approaches infinity, whereas the marginal cost of expanding the growing region is always positive (and in fact increasing due to the convex delivery costs). Hence it cannot be optimal

to let  $n$  approach infinity, implying that for each parameter set there must be some upper bound beyond which the optimal  $n$  can never be found.

The second issue, a non-convex lower-cost set is dealt with by propositions 3.1–3.3.

### **Ruling out the optimality of $n \leq n_{MY}$**

**Proposition 3.1** *The optimal maximum orchard age,  $n^*$  must be strictly greater than the maximum orchard age that maximizes yield,  $n_{MY}$ , i.e.  $n^* > n_{MY}$ .*

This proposition shows that production costs can never be minimized while orchard yield is an increasing function of maximum orchard age.

In the presence of fixed replanting costs, the optimal maximum orchard age is never less than the orchard age that corresponds to peak yield. This is because for each  $y(n)$  below the peak, there is an  $n$  that generates this level of  $y(n)$  to the left of the peak and to the right of the peak. The  $n$  to the right of the peak has a lower average replanting cost, since replanting costs are dispersed over a larger number of age-classes, and thus is always the lower cost choice.

Since  $n_{MY}$  occurs to the right of the local minimum of the isocost curve, I have eliminated the most notable non-convexity from the domain. Note, however, that the isocost curve to the right of  $n_{MY}$  is convex, so the lower-cost set in this region is still non-convex. The next sections establish that a solution to the first order conditions exist in this region, and that this solution also solves the cost-minimization problem.

### **The existence of a solution to the cost minimization problem**

**Proposition 3.2** *Given assumptions (3.1)-(3.4), a solution,  $n^*$ , to the cost minimization problem exists such that  $n^* \in (n_{MY}, \infty)$  and  $\varepsilon_{y(n^*)} > -1$ , where  $\varepsilon_{y(n)}$  is the elasticity of the yield function with respect to age.*

The intuition behind this proof is that the derivatives of the isocost and isoquant functions must be equal somewhere on the set  $(n_{MY}, \infty)$ . At  $n_{MY}$  the slope of the isocost function is greater than the slope of the isoquant function. Conversely, as  $n$  approaches infinity, the slope of the isoquant approaches a positive constant while the slope of the isocost function approaches zero. By continuity there must be some intermediate point where the slopes are equal.

The Hoerl function satisfies the sufficiency condition (shown in lemma B.1), so for this family of functions a solution to the simulated cost minimization problem exists.

### **Does $n^*$ minimize costs?**

A sufficient condition for  $(n^*, L^*)$  to (locally) minimize costs is that the Lagrangian be convex at this point. A sufficient condition for the Lagrangian to be convex is that the determinant

of the bordered Hessian be negative at the candidate point. Evaluating the determinant of the bordered Hessian leads to the following proposition.

**Proposition 3.3** *The condition*

$n^3 \left( y'(n)^2 ((\alpha - 3) \alpha C_D L^{\alpha-1} + 2\lambda) - y(n) y''(n) (\lambda - C_D L^{\alpha-1}) \right) + 2 C_n (n y'(n) + y(n)) > 0$   
*when evaluated at  $(n^*, L^*)$  is a sufficient condition for  $(n^*, L^*)$  to solve the cost-minimization problem.*

All of the simulation results presented in this chapter satisfy the condition in proposition 3.3. This was verified numerically in MATLAB.

## The behavior of $n^*$ and $L^*$ with respect to changes in biorefinery capacity

How does the optimal growing region size and maximum orchard age change as the given biorefinery size changes? I answer this question by finding and analyzing the derivatives  $\frac{dn^*}{dQ}$  and  $\frac{dL^*}{dQ}$ .

**Proposition 3.4** *As biorefinery size increases, the optimal orchard age decreases, i.e.  $\frac{dn^*}{dQ} < 0$ .*

As the biorefinery size increase, the optimal maximum orchard age decreases, and approaches the maximum yield age. For  $n > n_{MY}$ , a decrease in  $n$  increases the yield of the feedstock growing region. Since a fixed quantity of feedstock must be produced, higher yield allows the growing region to be marginally smaller. I have shown that it is always worthwhile to marginally boost yields as the biorefinery size increases. The magnitude of this effect, and whether it is economically important, depends on the particular parameterization of the model.

My model predicts that larger refineries should be surrounded by younger orchards, *ceteris paribus*. This is a strong, testable empirical prediction.

**Proposition 3.5** *The change in optimal growing region size with respect to a change in biorefinery capacity is generally ambiguous, but if  $y''(n^*) < 0$ , then increased biorefinery capacity leads to increase growing region size, i.e.  $\frac{dL^*}{dQ} > 0$ .*

Figure 3.3 shows a numerical example of how the cost minimizing orchard age and planted land area change as the biorefinery processing capacity is increased. The example is calibrated to sugarcane mills in São Paulo state, mills ranging in capacity from 1 million to 36 million tons per year. The details of the calibration are provided in appendix A.2. Standard errors and sensitivity analysis are not presented for this example.

The shape of the expansion path corresponds to the results of the comparative statics of  $n^*$  and  $L^*$  as I increase  $\bar{Q}$ . As biorefinery capacity increases from 1 to 36 million tons the

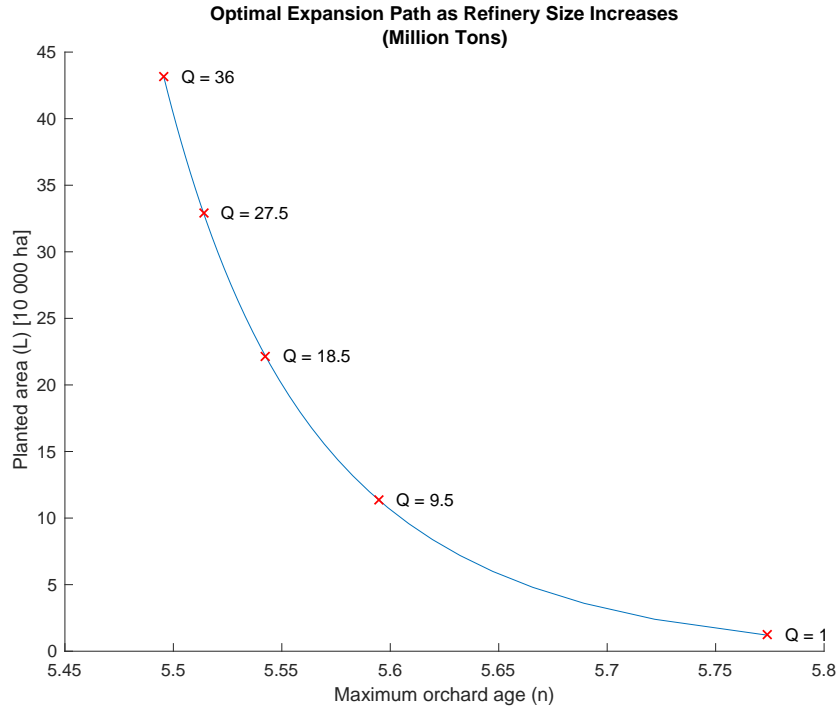


Figure 3.3: Cost minimizing orchard age and planted land area as biorefinery capacity is increased from 1 million tons to 36 million tons.  $n_{MY} = 5.31$ . Details of calibration are in appendix A.2

cost minimizing orchard age decreases from 5.77 to 5.50, corresponding to an increase in the  $n$ -orchard yield from 83.31 tons per hectare to 83.52 tons per hectare. The majority of the additional feedstock is supplied from additions to planted land area.

Qualitatively, from these results I expect to see more variation in cost minimizing orchard age among smaller refineries than larger refineries. The difference in cost minimizing orchard age between the smallest and the median biorefinery size is almost three times greater than the age difference between the median and the largest biorefinery. The intuition behind this effect comes from the peak in the  $n$ -orchard yield function. As the cost minimizing  $n$  approaches  $n_{MY}$  the additional yield from an age reduction declines, but the additional cost of more frequent replanting remains positive. Hence larger increases in  $\bar{Q}$  are necessary for the benefits of additional yield to justify the additional replanting costs.

Quantitatively, the magnitude of the yield change between largest and smallest biorefinery is small. I anticipate that it would be difficult to detect empirically this small effect among the noise introduced by exogenous year-to-year variation in yields (by weather, for example).



## Comparative statics of $n^*$ and $L^*$ with respect to other exogenous parameters ( $\bar{Q}$ held constant)

**Proposition 3.6** *Table 3.1 presents the signs of the comparative statics of  $n^*$  and  $L^*$  with respect to the listed exogenous parameters.*

$\mathbf{x}$	$\frac{dn^*}{dx}$	$\frac{dL^*}{dx}$
$C_f$	$(< 0)$	$(< 0)$
$C_n$	$(> 0)$	$(> 0)$
$C_D$	$(< 0)$	$(< 0)$
$\alpha$	$(< 0)$	$(< 0) \Leftrightarrow L > e^{-1/(\alpha-1)}$

Table 3.1: Signs of comparative statics of  $n^*$  and  $L^*$  with respect to four key parameters.

The parameter  $C_f$  represents the fixed cost of growing an orchard of any age — these are the age-structure independent costs, the costs of maintaining the land on which the orchard stands. An increase in  $C_f$  increases the marginal costs of land, so the optimal mixture moves away from land and toward yield through a reduction in the age.

The parameter  $C_n$  represents the age-structure dependent costs of the orchard. In the case of sugarcane this cost includes replanting costs, as well as the cost of maintaining plants in the years after they are planted.<sup>6</sup> An increase in these costs always increases the attractiveness of older plants, since older plants are relatively cheaper than new plants.

The parameter  $C_D$  represents the costs of delivering feedstock from the field to the biorefinery. It affects both costs of yield and of land, since both these variables determine the quantity of feedstock transported. However, since this analysis is focused on a cost-minimization problem with respect to a fixed biorefinery size, a change in delivery cost only affects the marginal cost of land, leaving the quantity of feedstock produced unchanged. Unsurprisingly then, an increase in the delivery cost increases the marginal cost of land, and like the effect of  $C_f$ , shifts the optimal mixture of land and yield towards yield and away from land.

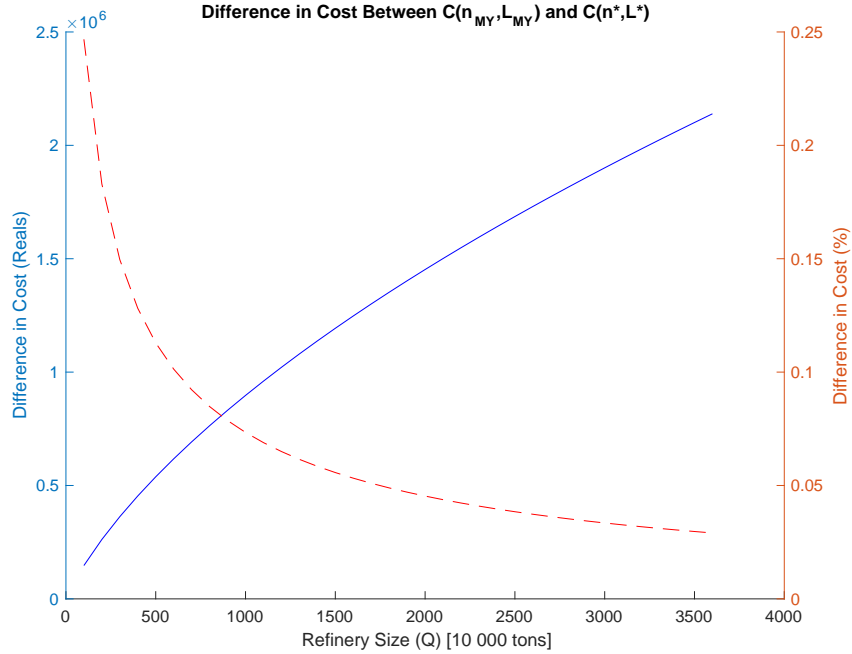


Figure 3.4: Details of calibration are in appendix A.2

### By how much is cost reduced when $n$ is endogenized?

Figure 3.4 shows the cost savings from allowing  $n$  to be chosen endogenously, compared to fixing  $n = n_{MY}$  and only allowing  $L$  to change as biorefinery size is scaled up. The solid blue line represents the absolute cost savings (in Reals). The absolute cost savings increase as the biorefinery size is scaled up range from around R\$150 000 to around R\$2 150 000 for the largest biorefinery. These cost savings are small, however, when compared to the total cost of production. The relative cost savings are highest for the smallest refineries, with the smallest biorefinery reducing its costs by around 0.25 percent, while the largest biorefinery only reduces its costs by 0.03 percent. The cost savings are relatively more elastic for smaller refineries since cost minimizing age is more elastic for smaller refineries (see figure 3.3).

### How robust is this result?

Given the small magnitude of the percentage cost reduction in the previous section, a natural question arises: how robust is this result? Does this particular parameter constellation lead to an unusually small cost reduction? Or, alternatively, is this about the best one can hope for from this model?

<sup>6</sup>The cost of maintaining a plant is much less than the replanting cost. The total cost of replanting and maintaining is dependent on the maximum orchard age, since  $\frac{1}{n}$  of the orchard is being replanted and  $\frac{n-1}{n}$  is being maintained each year. See appendix A.2 for more details.

The small result for the Brazilian example suggests that, for that case, using the maximum yield age is almost as good as using the cost minimizing age, further suggesting that the rule-of-thumb suggested by Tisdell and De Silva (1986) is a good one in this case. But are there other cases where this rule-of-thumb would be misleading?

To explore this question I randomly drew parameter constellations from a plausible parameter space. Keeping the biorefinery size fixed at 1 000 000 tons, I drew 7-vectors from a uniform distribution over the hyper-rectangle described in table 3.2. For each draw from the distribution I attempted to solve for the cost minimizing orchard age and planted area, discarding the results for which the cost minimization problem could not be solved, or were inconsistent with the theory, for example results with a negative cost difference ( $TC(n_{MY}) < TC(n^*)$ ).

I repeated this process 100 000 times. A histogram of resulting cost-reduction percentages is presented in figure 3.5. Since over 98 percent of parameter draws had cost reductions of less than 1 percent, I can conclude that the Tisdell and De Silva rule of thumb is a good one in most cases, assuming that the case is included in the parameter search space.

Parameter	Min Value	Max Value
a	0	10
b	0	10
c	-3	0
$C_f$	0	10 000
$C_n$	0	10 000
$C_D$	0	2
$\alpha$	1	3

Table 3.2: Support for cost minimization with random parameters

### 3.4 Discussion and Future Research

Although the simulation results were consistent with the predictions from the theoretical model, their magnitudes were small, with a maximum cost saving from endogenizing yield of only 0.25 percent. In this chapter I have not explored the reason for the small numerical result, leaving a thorough analysis for future research. I can, however, speculate on its causes.

The optimal orchard age is very close to the yield-maximizing orchard age in all of the simulations. From the comparative statics in table 3.1 I know that increases in age-structure independent costs, delivery costs, and delivery cost convexity all decrease the optimal orchard age. I can infer that in the Brazilian example these three forces dominate the age-structure dependent costs. If these three costs were lower, or the age-structure dependent costs were higher I would expect to see older orchard ages. For example, if the transportation network were improved, thereby lowering delivery costs, I would expect to see older orchard ages.

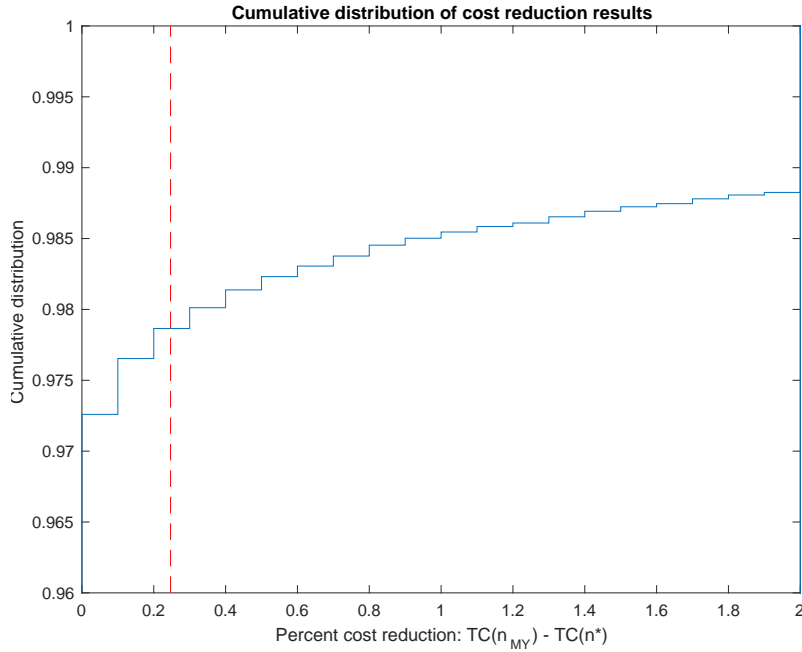


Figure 3.5: Truncated graph of cumulative percentage cost reductions from 100 000 draws from a uniform distribution with support described in table 3.2. The result from the simulation using the parameters calculated in appendix A.2 is shown with the dashed, red line. Note that the y-axis begins at 0.96, and approximately 1.3 percent of the results had a cost reduction of greater than 2 percent.

Likewise if replanting costs increased, or the costs for maintaining established plants decreases, I would expect to see older orchard ages.

Also, since optimal orchard ages are so close to the yield-maximizing age, it is likely that greater reductions in total cost can be achieved through reducing the age-structure independent costs of land (e.g. managerial effort, irrigation costs) or the delivery costs, through improving and straightening roads used for delivery, or improving the logistical management of the feedstock network.

The shape of the age-yield relationship will also affect the potential cost-savings from age-structure endogenization. I have conducted the numerical analysis with a given age-yield relationship, omitting sensitivity analysis of alternative age-yield specifications. The analytical structure of the Hoerl function allows me, in principle, to conduct comparative static experiments on the three Hoerl function parameters,  $a$ ,  $b$ , and  $c$ . Unfortunately, none of these parameters map directly to the intuitive targets of a comparative static exercise, such as the time to peak yield, or the relative ‘flatness’ of the age-yield relationship. Further, as discussed in section 3.2, the Hoerl function has some drawbacks modeling a perennial crop that quickly reaches peak yield. The comparison of alternative choices of the functional

form of the age-yield relationship and an analysis of the effect of the shape of the age-yield relationship on the optimal choices of orchard age and land use are left for future research.

The numerical results for São Paulo, while illustrative, require more analysis before firm policy conclusions can be drawn from them. In particular, my choice of the yield-maximizing age as the reference age may not reflect the age used by managers in the region. For example, in January 2014, the average age (not maximum age) of all sugarcane fields Brazil's South-Central growing region (including São Paulo) was 4.6 years, implying that there are many fields over 5 years old, and possibly reaching 6, 7, or 8 years old (UNICA, 2014). In the case of a manager whose cost parameters are similar to those in this study, but whose fields are this old, the cost savings from reducing the maximum orchard age to the optimal orchard age could be substantial.

A surprising result from the model is that smaller refineries have proportionally larger cost savings. One might think, *a priori*, that the costs savings would be larger for larger refineries because they require more feedstock, so lowering the cost of feedstock would be more beneficial. However, for large refineries, the optimal orchard age is so close to the yield-maximizing orchard age that endogenizing the age makes very little difference to the choice of yield, thus resulting in proportionally small cost reductions. It is precisely because larger refineries require large quantities, thereby incurring high marginal delivery costs due to the greater extent of the growing region, that they are least benefited by endogenizing yield.

Focusing on cost-minimization hides a potential channel for age-structure endogenization to affect the management of the biorefinery-feedstock operation. By keeping the biorefinery size fixed, a change in a parameter, say age-structure dependent costs, only causes a substitution between land and age-structure. In a more flexible model that also endogenized the biorefinery capacity choice, that same parameter change would have two effects: the substitution effect from the change in the relative cost of land and age; and the size effect, where the change in the parameter alters the optimal size of the biorefinery. These two effects may reinforce or offset each other. For example, a decrease in delivery costs may increase the optimal orchard age through the substitution effect, and also increase the optimal size of the biorefinery, also increasing the optimal orchard age. This example is speculative, based on the comparative statics in sections 3.3 and 3.3, but a full, simultaneous analysis of both the substitution and size effects would require a net present value maximization model with endogenous biorefinery capacity choice.

## 3.5 Conclusion

I have, to my knowledge, presented the first model of optimal orchard age when the output of the orchard is used as a feedstock for a biorefinery of a given size.

To account for non-convexities in the cost-minimization problem, I proved that, under certain assumptions on the age-yield relationship, the first order conditions of the model have a solution, and provided a sufficient condition for this solution to also solve the original cost-minimization problem. I provided analytical comparative statics of this solution with

respect to biorefinery size and cost parameters. I also provided a numerical simulation, calibrated to the sugarcane industry in São Paulo, Brazil, to illustrate the magnitude of the costs savings possible when orchard age is endogenized.

From the results I can conclude that it is indeed *possible* to reduce biorefinery feedstock costs by endogenizing the feedstock age-structure, but in the case of the São Paulo, the benefits are small relative to the reference case of the yield-maximizing age. I confirm that the rule-of-thumb suggested by Tisdell and De Silva (1986) and implied by Margarido and Santos (2012) is relevant for São Paulo.

The importance of age-structure endogenization to managers must be considered on a case-by-case, depending on whether the parameters of the particular problem generate a large cost saving from endogenization, or whether the yield-maximizing age is a good heuristic. Further, if the orchard is currently managed at some age other than the yield-maximizing age, this alternative age must be used as a reference, possibly generating still greater cost savings.

# Chapter 4

## The Economics of Perennial Orchards with Endogenous Age-Classes

### 4.1 Introduction

The economics of perennial crops encompasses a variety of interesting questions, including the following examples. When does a farmer choose to plant perennials? That is, when is the net present value (NPV) of growing a perennial crops superior to the NPV of growing his next best alternative? What could cause a farmer to switch from one perennial to another? For instance, farmers in California's Central Valley have been uprooting established raisin grape vineyards to replace them with almond trees (Philpott, 2015). What does the profit differential between these two crops need to be to justify the large capital expense of planting a new perennial crop, and the waiting period before the new trees start producing a crop? How does the farmer decide the optimal age for replacing his trees?

Other questions include: When does a farmer manage the orchard as a single tree (plant and remove all trees at the same time), which allows for the use of more established single tree replacement models Wessler (1997); Feinerman and Tsur (2014), or when does the farmer manage his trees as an age-structured orchard, having multiple age-classes active at once, planting, replacing, or uprooting different sections of his orchard each year? That is, when is a single tree model a good approximation of farmer behavior, and when is a richer, more complex, age-structured model necessary? And, given an initial orchard, how does a farmer get to an optimal allocation of perennials? That is, if he has a desired steady-state age-structure, how and when can he reach it from an arbitrary initial orchard?

Many perennial crops must be processed after harvesting before they can be sold to wholesalers and consumers. Tree nuts, for example, must be shelled before they can be eaten or used in food production. A particularly important example regarding the processing of perennial crops comes from the biofuels sector, where perennial feedstocks like sugarcane, *miscanthus*, and switchgrass are processed by large biorefineries into ethanol. A key problem facing a biofuels processor is the choice of biorefinery size. Variability in feedstock supply

is costly (Debnath et al., 2014), and the variability in feedstock supply caused by the age-structure of perennials is a new area of study (see chapter 2). For a vertically integrated biofuel grower/processor, what is the optimal biorefinery size, and what is the optimal management strategy for the perennial feedstock?

Finally, how do the answers to these questions change with parameters? For example, what if improvements in plant breeding led to a shorter establishment period, or a less steep yield decline after the yield peak?

In this chapter I develop an analytical framework for answering these questions, with a particular focus on the importance of age-structure. Although many of these questions have been asked, and answered, before, in the remainder of this introduction I show how my approach differs from previous work, and the advantages and need for this analytical approach.

Devadoss and Luckstead (2010) identify four key differences between annual and perennial crops:

- (1) perennial crop supply is vastly different from the annual crop supply because
- (2) trees are a long-term investment;
- (3) perennial crops have long gestation intervals between initial plantings and first harvest;
- (4) once trees start yielding, there is an extended period of productivity, and then gradual decline in production; and
- (5) after trees reach their final productivity decline, they are removed.

From this list it is clear that, unlike annuals, a static model is inadequate for studying perennials. This sentiment is echoed by Just and Pope (2001, p. 706) in their review of agricultural production economics: “Specifically, lags and dynamic processes appear to be at the heart of understanding large-animal livestock and perennial crop production problems.”

Having satisfactory models of the production decisions of perennial crop growers gives us insight into an important and valuable sector of the world’s agricultural industry. Hence the study of perennial crops has been given considerable attention by agricultural economists, particularly because of their importance for the economies of tropical developing countries (Bellman and Hartley, 1985), but they also play an important role in the economies of developed countries.

In the US, the value of production of fruit and tree nuts was 30 billion USD, production mainly from perennial crops. Tree nut value is around \$10b, led by almonds, walnuts, and pistachios (USDA, 2016a). The fruit production value is led by grapes, apples, strawberries, and oranges (strawberries are grown as an annual crop in California (California Strawberry Commission, 2017), but can be grown as a perennial, a practice often done in colder climates. The other three crops are perennials). Perennial crops are also important to regional economies. For example, Apple production value in Washington state was greater than \$1 billion in 2007 (Devadoss and Luckstead, 2010).

Perennial crops are also a source of value through their conversion to biofuel. For example, Brazilian biofuel is a particularly valuable liquid fuel in the fight against climate change, with Crago et al. (2010) estimating that its life-cycle carbon emissions are about half those of corn



ethanol. Furthermore, ethanol is an important component of Brazil's economy, accounting for around 2.3 percent of its GDP in 2010 (Valdes, 2011).

Finally perennial crops can provide substantial environmental benefits, relative to their annual cousins, through increasing soil carbon, reducing erosion, and requiring fewer chemical inputs (Glover et al., 2010).

The questions posed at the start of this section are normative questions, relating to what a farmer or manager *should* do. The majority of analysis of perennial crops in agricultural economics has focused on the positive questions of how farmers actually manage their perennial crops. Due to their positive focus, the majority of work in agricultural economics on the econometrics of perennial crop production has been on econometric studies (Rausser, 1971; French and Matthews, 1971; French et al., 1985; Trivedi, 1987; Knapp and Konyar, 1991; Kalaitzandonakes and Shonkwiler, 1992; Devadoss and Luckstead, 2010; Brady and Marsh, 2013).

There has been far less normative work, much of which has been based on simulation and computation, with very little analytical, theoretical work. Bellman and Hartley (1985) present a comprehensive dynamic programming framework to numerically simulate the supply of a perennial crop. They do not actually perform the simulation. Knapp (1987) presents a model and calibrated simulation for calculating the dynamic equilibrium in a market for perennial crops. He applies his model to the alfalfa industry in California. And Franklin (2012) presents a thorough age-structured dynamic programming computational model to simulate wine grape production in South Australia.

There are few purely analytic studies of the management of perennials, and those that have been done mostly focus on the management of a single tree, or an even-aged orchard. Kalaitzandonakes and Shonkwiler provide a possible reason for this in their 1992 paper:

The theoretical underpinnings of perennial investment decisions have been elaborated by Bellman and Hartley, and Trivedi and need not be repeated here. These studies have demonstrated that due to the complex dynamics of perennial technologies and the heterogeneity of perennial capital, intertemporal profit maximization of the perennial firm admits analytical solutions only under rather restrictive technological assumptions. Such technological assumptions could not be supported on a priori grounds for the Florida grapefruit industry and hence the formal optimization problem of the firm is not explicitly considered in this study.

However, the complexity of perennial crop production is not a compelling argument against the development and analysis of stylized, analytic models. As in many other branches of economics, the development of analytic and empirical models are substitutes, not complements, creating a virtuous cycle in which new insights and knowledge are synthesized from their joint production. The analytic dimension of perennial crop production is far less developed than then empirical dimension, and it is not clear, *a priori*, that this imbalance is optimal. The value of the analytical approach to modeling perennials is that it can provide generalizable

intuition about the behavior of perennials, identifying that which can be said in general, and that which depends on the specifics of the particular crop and context.

Of the few analytic perennial crop management models, most are focused on the single tree management problem. The single-aged tree replacement problem is an example of the classic durable asset replacement problem (see Miranda and Fackler (2002, section 7.2.2) for a simple example, or Rust (1987) for a more complex one). Wessler (1997) provides a framework for calculating the NPV of an even-aged orchard from an arbitrary starting age. Feinerman and Tsur (2014) analyzes the impacts of stochastic interruptions to an even-aged orchard production cycle (in particular, reduced water supply due to drought).

The only purely theoretical analyses of age-structured orchards that I am aware of are the studies by Mitra et al. (1991) and Wan (1993). The framework for their theoretical analysis is built on the multi-sector growth model literature (Majumdar, 1987; Brock and Dechert, 2008). The multi-sector growth approach is an intertemporal extension of general equilibrium theory, and is thus “analytically organized around existence of equilibrium, the core and equilibria, the two welfare theorems, as well as the ‘anything goes’ theorem of Sonnenschein, Mantel and Debreu (SMD)” (Brock and Dechert, 2008, p. 4). Agricultural economics, in contrast, is focused more on empirical applications, and the behavior of farmers in response to changes in economic parameters, e.g. building supply response functions.

Mitra et al. (1991) analyze an infinite horizon model of perennial production with arbitrarily many age-classes, fixed total land, and a positive discount factor. They characterize the set of optimal stationary orchards for both linear and strictly concave benefit functions. Their main results regard the convergence of an optimal program to the optimal stationary orchard under strictly concave benefit and positive discounting. They find that, by assuming the age-yield relationship is non-decreasing over the tree’s lifespan, the optimal program never converges to the optimal stationary forest, rather to a cycle in a neighborhood around it. They prove a neighborhood turnpike result—namely, that as the discount factor approaches 1, the amplitude of the optimal cycle approaches zero. Wan (1993) focuses on the existence and uniqueness of steady-states in a two age-class tree farm, as well as the possibility of cyclical tree planting in the long-run. In his remark in section 3, he notes that his framework is “quite general” and can be applied to the orchard model of Mitra et al. (1991). He states, without proof, the optimal control rule for 2-age-class perennials, which excludes the possibility of the neighborhood turnpike result developed by Mitra et al.. He does not discuss the difference between his result and that of Mitra et al.

A distinct, but related, strand of literature with a more agricultural and natural resource economics style is the series of three papers on the optimal management of age-structured forestry written by Salo and Tahvonen between 2002 and 2004. The forestry model is a relative of the orchard model, where forestry is an instance of *point input, point output* capital, while perennial crops are an example of *point input, flow output* (Mitra et al., 1991). Salo and Tahvonen develop their results using a dynamic Lagrangian approach (Chow, 1993, 1997), rather than the capital theoretic approach previously used in theoretical analyses of the economics of age-structured forestry models (Mitra and Wan, 1985, 1986), or orchard models (Mitra et al., 1991; Wan, 1993). Salo and Tahvonen bridge the gap between the

capital theoretic approach, and the resource economics approach, beginning their analysis by focusing on the long-run equilibrium conditions of their age-structured forest, but then developing numerical transition paths, and supply curves, which have a far more applied focus.

In this chapter I develop an age-structured model of perennial crop planting decisions, using a similar model to Mitra et al. (1991) and Wan (1993), but analyzed using the dynamic Lagrangian approach of Salo and Tahvonen (2004).

There are four key differences between my model and the earlier orchard models. First, I assume that there are only two age-classes. This is an analytical convenience to help develop the modeling framework. In future research I intend to relax this assumption. Second, I allow a flexible total land quantity in each period, and assume a constant opportunity cost of land, which simplifies the analysis by imposing fewer constraints on the problem. Further, because I do not have the explicit land constraint, and allow the aging constraint to bind or not, I can generate a richer set of trajectories. Third, I use a finite time horizon with a salvage value after the final period. I conjecture that using the finite time horizon model with an appropriate choice of parameters and scrap value function can generate a trajectory identical to one generated by a Mitra et al. (1991) or Wan (1993) in their fixed land, infinite horizon models. If this is so, this analysis nests Wan's two-age-class variation of Mitra et al. (1991), but also allows for a richer variety of qualitative planting trajectories. And fourth, to facilitate my analysis I use a quadratic benefit function for much of the chapter, a benefit function that is more special than theirs. This is another assumption I wish to relax in future work.

Similarly to Mitra et al. (1991) and Wan (1993), I choose a discrete time framework. My choice of a discrete time framework for this analysis is twofold. First, the choice of a discrete time framework is realistic for the perennial crop setting. While timber may be harvested at any point during the year in principle (especially in regions without strong seasonal variation in weather), many perennial crops are harvested on a yearly cycle due to the biology of the crop. Second, it is a far more thoroughly studied case, and the methods of analysis are more elementary, allowing me to focus on the applications and intuition of the model, and creating scope for reaching a wider audience. Fabbri et al. (2015) study the related Mitra and Wan forestry model in continuous time, a development pioneered by these authors. To analyze this problem they developed a new class of vintage models, allowing them to use measure valued controls, which allow for all members of a continuous time age-class to be cut at once. For their forestry model, they find that the majority of canonical results carry over to the continuous time case, with the notable exception of cyclical optimal solutions.

To my knowledge, this is the first study of an age-structured, finite-horizon model of orchard management.

This chapter is organized as follows: Section 4.2 presents two variations of a two-period, two-age-class model. Section 4.2 introduces the most basic two-age-class perennial crop model, where a farmer plants a perennial that provides benefits for two years, but he cannot change his land allocation after the initial planting. Section 4.2 relaxes this restriction, allowing him to change his land allocation in the second period, illustrating the importance of the

*aging constraint.* Section 4.3 presents a general, two-age-class,  $T$ -period model, which is then analyzed at an arbitrary time period. The results of this analysis are used to develop a proposition regarding the planting of young trees in the final period, and to analyze an example 3-period trajectory. Section 4.4 concludes the chapter with discussion of the analysis and directions for future work.

## 4.2 Two-Period Models

### A two-period perennial model with no uprooting or planting in second period

This section illustrates the basic difference between perennial and annual crops, namely that perennials produce for multiple years, while annuals only produce for a single year.

Assume there are only two periods. Assume that a farmer owns land that he is considering using to grow a perennial crop. On this land he may grow either a perennial crop that lives for both periods, or he can rent the land to someone else and receive a constant rental rate<sup>1</sup>,  $p$ . In the first period he chooses how much land to allocate to the perennial, an allocation that cannot be adjusted in the second period. How much land will he optimally allocate to perennials?

Let  $x$  be the area of land allocated to perennials. The perennial lives for two periods. A unit of land with young perennials produces  $f_y$  units of output, and the same unit of land with mature perennials produces  $f_m$  units of output. At this point I make no assumptions about the relative magnitudes of  $f_y$  and  $f_m$ .

I assume that the farmer derives benefits with diminishing marginal returns from perennials<sup>2</sup>. His benefit from perennials is given by  $u(\cdot)$  with  $u'(\cdot) > 0$  and  $u''(\cdot) < 0$ . His total net benefit in a period is the benefit from perennial production, less the rental rate of the land, which is the opportunity cost for choosing to grow perennials. The farmer discounts the future using a discount factor  $\beta$  ( $0 < \beta \leq 1$ ).

The present value of his net benefit in each period, with  $x$  units of land allocated to perennials, is given by

**Period 1:**  $u(f_y x) - px$

**Period 2:**  $\beta(u(f_m x) - px)$

The farmer's total benefit from the two periods is thus

$$U(x) = u(f_y x) - px + \beta(u(f_m x) - px)$$

---

<sup>1</sup>An alternative justification for a constant rental rate is that he may use the land to grow an annual commodity crop for which he is a price taker

<sup>2</sup>A possible reason for this is that he has a degree of market power when selling the perennial crop.

The farmer maximizes the benefits from the plot of land by setting the derivative of the benefit function with respect to  $x$  to zero.

$$\begin{aligned}
 U'(x) &= u'(f_y x) f_y - p + \beta(u'(f_m x) f_m - p) = 0 \\
 &= \underbrace{u'(f_y x) f_y + \beta u'(f_m x) f_m}_{\text{Marginal Benefit of Perennials}} - \underbrace{p(1 + \beta)}_{\text{Opportunity Cost of Land}} = 0
 \end{aligned}$$

A marginal increase in the land allocated to perennials benefits the farmer by providing a marginal young tree this period and a marginal mature tree in the following period. This benefit is obtained at the opportunity cost of land,  $p$ , in both this period and the next.

The optimal land allocation is found by setting this derivative to zero and solving for  $x$ . Let  $g(x) = u'(f_y x) f_y - p + \beta(u'(f_m x) f_m - p)$ . By total differentiation the comparative static of  $x$  with respect to the productivity of young trees,  $f_y$ , is given by

$$\frac{dx}{df_y} = \frac{-\frac{\partial g}{\partial f_y}}{\frac{\partial g}{\partial x}} = \frac{u''(f_y x) f_y x + u'(f_y x)}{u''(f_y x) f_y^2 + \beta u''(f_m x) f_m^2}$$

For this benefit function, the denominator is always negative, hence the sign of the comparative static is identical to the sign of the numerator. Let  $Q_1$  be the quantity of perennial output produced in the first period. The numerator is positive when

$$\begin{aligned}
 u''(f_y x) f_y x + u'(f_y x) &> 0 \\
 u''(f_y x) f_y x &> -u'(f_y x) \\
 \frac{-u''(f_y x) f_y x}{u'(f_y x)} &< 1 \\
 \frac{-u''(Q_1) Q_1}{u'(Q_1)} &< 1 \\
 -\varepsilon_{u'(Q_1)} &< 1
 \end{aligned}$$

Where  $\varepsilon_{u'(Q)}$  is the elasticity of marginal benefit from perennials with respect to quantity of perennial output. Thus, if and only if marginal benefit from perennials is *inelastic* with respect to quantity of perennial output, an increase in the productivity of young trees will lead to an increase in the land allocated to perennials.

Similarly, the comparative static with respect to the productivity of mature trees implies that the optimal land allocation to perennials will increase with an increase in the productivity of mature trees if and only if the elasticity of marginal benefit with respect to the quantity of perennial output in the second period is inelastic.

An increase in the rental rate always reduces the land allocated to perennial crops. Finally, an increase in the discount factor (where the farmer becomes more patient) increases the optimal allocation of land to perennials if and only if the marginal benefit of mature trees is

$\alpha$	$\frac{dx^*}{d\alpha}$	Condition
$f_y$	$> 0$	$\Leftrightarrow -\varepsilon_{u'(Q_1^*)} < 1$
$f_m$	$> 0$	$\Leftrightarrow -\varepsilon_{u'(Q_2^*)} < 1$
$p$	$< 0$	—
$\beta$	$> 0$	$\Leftrightarrow u'(f_m x^*) f_m > p$

Table 4.1: Conditional signs of comparative statics of  $x^*$  with respect to key parameters.

greater than the rental rate. Since the farmer cannot separate his planting decision between young and mature trees, it is possible that the optimal land allocation may result in positive net marginal benefit from the marginal mature tree. When this is the case, and the farmer becomes more patient, he exploits this surplus benefit by increasing his planting of perennials.

This model explains the basic decision of allocating land to perennials, but does not yet make clear the importance of young trees as the precursors to mature trees. I now extend this model to allow the possibility of uprooting mature trees (but not replacing them with young trees) in the second period, which will clarify the additional value of young trees in allowing the possibility of future mature trees.

## A two-period perennial model with uprooting in the second period

In the previous model every young tree became a mature tree. Now I give the farmer the choice to decide how many young trees he allows to reach maturity, allowing all or only some fraction of the trees to reach maturity in the second period. This extension of the model shows how young trees provide two sources of value to the farmer: the immediate benefit of the production from a young tree, and the future benefit of allowing the farmer to obtain the production from a mature tree in the next period.

Since the farmer may uproot some mature trees, I must separately keep track of the land allocated to young and mature trees. Let  $x_y$  be the land allocated to young trees in period 1, and let  $x_m$  be the land allocated to mature trees in period 2. In this section I also normalize the yield of young trees to 1, so  $f_m$  represents the yield of a mature tree relative to a young tree. If  $f_m > 1$  mature trees are more productive than young trees, and vice versa.

Since the farmer cannot have more mature trees in the second period than he had in the first period, I introduce a constraint on the number of mature trees  $x_m \leq x_y$ , called the *aging constraint*<sup>3</sup>.

<sup>3</sup>This constraint is often called a a cross-vintage bound, e.g. Wan (1994) or Salo and Tahvonen (2004), but, in my opinion, aging constraint is more intuitive.

The objective of the farmer now becomes

$$\begin{aligned} \max_{x_y, x_m} \quad & u(x_y) - px_y + \beta(u(f_m x_m) - px_m) \\ \text{s.t.} \quad & x_m \leq x_y \end{aligned}$$

The Lagrangian function for this maximization problem is given by

$$\mathcal{L} = u(x_y) - px_y + \beta(u(f_m x_m) - px_m) + \lambda(x_y - x_m)$$

And the Karush-Kuhn-Tucker (KKT) conditions for this problem are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_y} &= u'(x_y) - p + \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial x_m} &= \beta(u'(f_m x_m) f_m - p) - \lambda = 0 \\ \lambda &\geq 0; \quad \lambda(x_y - x_m) = 0 \end{aligned}$$

In this model there are two regimes to consider: the aging constraint is binding ( $\lambda > 0$ ); or the aging constraint is non-binding ( $\lambda = 0$ ). Let  $\hat{x}_y$  and  $\hat{x}_m$  be the optimal land allocations to young and mature trees if there were no aging constraint, that is  $\hat{x}_y$  and  $\hat{x}_m$  solve  $u'(\hat{x}_y) = p$  and  $u'(f_m \hat{x}_m) f_m = p$ . A binding aging constraint implies that  $\hat{x}_m > \hat{x}_y$ , and a non-binding aging constraint implies that  $\hat{x}_m \leq \hat{x}_y$ .

To explore the conditions when the aging constraint binds I specialize the problem by assuming that marginal benefits from perennial production is an affine function of perennial production, given by  $u'(Q) = b - aQ$  (with  $a, b > 0$ ).

The KKT conditions become

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_y} &= (b - ax_y) - p + \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial x_m} &= \beta((b - af_m x_m) f_m - p) - \lambda = 0 \\ \lambda &\geq 0; \quad \lambda(x_y - x_m) = 0 \end{aligned}$$

The KKT conditions for linear marginal benefit are shown in figures 4.1 and 4.2. Figure 4.1 shows the case of  $f_m > 1$  and figure 4.2 shows the case of  $f_m < 1$ .

The marginal benefit curves for young and mature trees cross at the vertical height  $\frac{bf_m}{1+f_m}$ . The value of the opportunity cost of land,  $p$ , relative to this threshold determines whether the aging constraint binds. However, whether the set of  $p$  that correspond to a binding aging constraint is above or below the threshold depends on the relative productivity of mature trees.

Recall the definition of a binding aging constraint:  $\hat{x}_m > \hat{x}_y$ . For linear marginal benefit  $\hat{x}_y = \frac{b-p}{a}$  and  $\hat{x}_m = \frac{b-p/f_m}{af_m}$ . Substituting these expressions into the binding aging constraint and solving for  $p$  gives

$$\begin{cases} p > \frac{f_m b}{f_m + 1}, & \text{if } f_m > 1 \\ p < \frac{f_m b}{f_m + 1}, & \text{if } f_m < 1 \end{cases}$$

**Proposition 4.1** *Assuming an interior (non-zero) solution and  $f_m \geq 1$ , if  $p \geq \frac{bf_m}{1+f_m}$  then  $\lambda > 0$ .*

In figure 4.1a the marginal benefit of mature trees ( $MB_m$ ) curve is to the right of the marginal benefit of young trees ( $MB_y$ ) curve for values of  $p$  above the threshold. Hence the aging constraint binds above the threshold. In figure 4.2a the  $MB_m$  curve is to the right of the  $MB_y$  curve for values of  $p$  below the threshold. Hence the aging constraint binds below the threshold.

Both of these cases are possible empirically, as shown by the following two examples: sugarcane, and apples. This model has two periods, but they do not necessarily correspond to two years. They represent two periods of the perennial crop's lifespan. The productivity of young trees may be greater or less than the productivity of mature trees, depending on the shape of the age-yield relationship (see chapters 1 and 2), and its division into young and mature periods. The relative productivity depends particularly on the time to peak yield and the rate of yield decline thereafter.

Sugarcane is an example where peak yield is reached rapidly, and yield declines rapidly too. For Brazilian sugarcane, peak yield is reached at first harvest, 12–18 months after planting. By its 7th year, around the end of its economic lifespan, its yield has halved from the peak. Its average yield in the first half of its life is slightly higher than the second half of its life (71 ton/ha vs. 67 ton/ha) (Margarido and Santos, 2012). In this model, sugarcane would be represented by  $f_m = 0.95 (< 1)$ .

Apples are an example where the peak is reached more slowly, with a longer non-bearing period, and the yield declines more slowly after the peak, so the average yield is higher in the second half of the life of apples. Red delicious apples in south-west France have an economic lifespan of around 35 years (Haworth and Vincent, 1977). The average yield in the first 15 years is around 200 kg/ha, while in final 20 years, the average yield is around 400 kg/ha. In this model, apples would be represented by  $f_m = 2 (> 1)$ .

Therefore, in this model the aging constraint for sugarcane binds for lower land rental rates, while for apples the aging constraint would bind for higher land rental rates.

### 4.3 A $T$ -Period Model

I now present a model with an arbitrary, finite time horizon that I use for the analysis in the remainder of the chapter.

Let the farmer grow trees over a  $T$  year time horizon. In each period  $t$  there is land allocated to young and mature trees, denoted by  $x_{yt}$  and  $x_{mt}$ . The farmer inherits a quantity of young trees from before the first period,  $x_{y0}$ , potentially allowing for mature trees in the first period.



The objective function is

$$\begin{aligned}
V(x_{y0}) = \max_{\substack{x_{yt}, t=1, \dots, T \\ x_{mt}, t=1, \dots, T+1}} & \sum_{t=1}^T \left[ \beta^{t-1} (u(x_{yt} + f_m x_{mt}) - p x_{yt} - p x_{mt}) \right] + \beta^T S(x_{m, T+1}) \\
\text{s.t. } & x_{mt} \leq x_{y, t-1} \quad t = 1, \dots, T + 1 \\
& x_{y, t-1}, x_{mt} \geq 0 \quad t = 1, \dots, T + 1 \\
& \text{and } x_{y0} \text{ given}
\end{aligned}$$

The Lagrangian for this problem is

$$\begin{aligned}
\mathcal{L} = & \sum_{t=1}^T \left[ \beta^{t-1} (u(x_{yt} + f_m x_{mt}) - p x_{yt} - p x_{mt}) \right] + \beta^T S(x_{m, T+1}) \\
& + \sum_{t=1}^{T+1} \left[ \lambda_t (x_{y, t-1} - x_{mt}) \right] + \sum_{t=1}^T \left[ \eta_{yt} (-x_{y, t}) \right] + \sum_{t=1}^{T+1} \left[ \eta_{mt} (-x_{m, t}) \right]
\end{aligned}$$

There are  $5T + 3$  Karush-Kuhn-Tucker conditions for this Lagrangian:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_{yt}} &= \beta^{t-1} (u'(x_{yt} + f_m x_{mt}) - p) + \lambda_{t+1} - \eta_{yt} = 0 \quad t = 1, \dots, T \\
\frac{\partial \mathcal{L}}{\partial x_{mt}} &= \beta^{t-1} (u'(x_{yt} + f_m x_{mt}) f_m - p) - \lambda_t - \eta_{mt} = 0 \quad t = 1, \dots, T \\
\frac{\partial \mathcal{L}}{\partial x_{m, T+1}} &= \beta^T S'(x_{m, T+1}) - \lambda_{T+1} = 0 \\
\lambda_t &\geq 0; \lambda_t (x_{y, t-1} - x_{mt}) = 0 \quad t = 1, \dots, T + 1 \\
\eta_{yt} &\geq 0; \eta_{yt} (-x_{y, t}) = 0 \quad t = 1, \dots, T \\
\eta_{mt} &\geq 0; \eta_{mt} (-x_{m, t}) = 0 \quad t = 1, \dots, T + 1
\end{aligned}$$

Most notation in this model has been introduced in the proceeding sections. The new notation is defined as follows:  $x_{nt}$  is the area of land allocated to age-class  $n$  in period  $t$ ;  $\eta_{nt}$  is the multiplier on the non-negativity constraint for age-class  $n$  in period  $t$ ;  $V(x_{y0})$  represents the optimized value of the farm of it's productive lifespan ( $T$  years plus salvage value), given the farmer starts with  $x_{y0}$  units of land planted with young trees that he may let mature in the first period;  $S(x_{m, T+1})$  is the salvage value function.

The salvage value represents the value in period  $T + 1$  if the farmer has  $x_{m, T+1}$  mature trees. It could represent sale price of his farm if the buyer operated the farm as a perennial orchard indefinitely and capital markets were perfect; it could represent the value of the mature trees in period  $T + 1$ , if the farmer chose to harvest them before choosing to change the enterprise mix on the farm; it could also be zero if the farmer abandons the farm at the end of period  $T$ .

This is a large system of equations. However, analysis of this system can be simplified by examining an arbitrary period  $t$ . In the setup of this model, each period  $1, \dots, T$  is symmetrical, having two non-negativity constrained choice variables  $x_{yt}$  and  $x_{mt}$ , and two potentially active aging constraints, one constraint determined by the area of young trees bequeathed to this period by the previous period,  $x_{mt} \leq x_{y,t-1}$ , and the other constraint determined by the desired area of mature trees in the next period,  $x_{yt} \geq x_{m,t+1}$ . This means that I can consider the problem in two stages. In the first stage I determine the optimal values of the choice variables, subject to the Lagrange multipliers implied by the boundary conditions, and in the second stage I find the set of Lagrange multipliers that maximize the value of the objective function, conditional on the maximized values of the choice variables in every period. The remainder of this chapter focuses on step one, while step two is left for future work.

To begin this analysis I specialize the problem by assuming that the marginal benefit function is linear (similar to the the earlier sections 4.2). Using this assumption, I can enumerate all possible combinations of the complementary slackness conditions (there are  $2^4 (= 16)$  cases), allowing me to identify the complementary slackness conditions that are consistent with any given set of parameters (price, patience, productivity, and boundary conditions), and hence the qualitative nature of the trajectory in each case.

The derivatives of the Lagrangian with respect to the choice variables in period  $t$ , assuming linear marginal benefit of perennial crop production, are

$$(4.1) \quad \frac{\partial \mathcal{L}}{\partial x_{yt}} = \beta^{t-1}(b - a(x_{yt} + f_m x_{mt}) - p) + \lambda_{t+1} - \eta_{yt} = 0$$

$$(4.2) \quad \frac{\partial \mathcal{L}}{\partial x_{mt}} = \beta^{t-1}(b f_m - a f_m (x_{yt} + f_m x_{mt}) - p) - \lambda_t - \eta_{mt} = 0$$

And the 4 complementary slackness conditions are

$$\begin{aligned} \lambda_t &\geq 0; \lambda_t(x_{y,t-1} - x_{mt}) = 0 \\ \lambda_{t+1} &\geq 0; \lambda_{t+1}(x_{yt} - x_{m,t+1}) = 0 \\ \eta_{yt} &\geq 0; \eta_{yt}(-x_{yt}) = 0 \\ \eta_{mt} &\geq 0; \eta_{mt}(-x_{mt}) = 0 \end{aligned}$$

Making assumptions on whether each of the 4 Lagrangian multipliers is zero or positive, I can use the two derivatives of the Lagrangian with respect to the choice variables to find the optimal values of the choice variables, and determine whether, and under what circumstances, the case is feasible.

Maintaining the assumption that  $f_m > 1$ , I enumerate the 16 cases for the Lagrangian multipliers for the period  $t$  KKT conditions in lemma C.1, detailed in appendix A.1. The details of the lemma are summarized in tables 4.2 and 4.3. Table 4.2 indicates the types of

land allocation possible in each of the cases. Depending on the case, the land allocation may be interior, boundary, origin, or infeasible, each possibility being a subset of the possibility before it. Table 4.3 is more specific, showing which boundary a boundary solution will lie on, and whether the land allocation in each age class is constrained by a neighboring period or not. The statement of the lemma in appendix A.1 includes necessary restrictions on the parameter values.

	1. ( $\eta_{yt} = 0, \eta_{mt} = 0$ )	2. ( $\eta_{yt} = 0, \eta_{mt} > 0$ )
1. ( $\lambda_t = 0, \lambda_{t+1} = 0$ )	Infeasible	Boundary
2. ( $\lambda_t = 0, \lambda_{t+1} > 0$ )	Interior	Boundary
3. ( $\lambda_t > 0, \lambda_{t+1} = 0$ )	Interior	Infeasible
4. ( $\lambda_t > 0, \lambda_{t+1} > 0$ )	Interior	Infeasible
	3. ( $\eta_{yt} > 0, \eta_{mt} = 0$ )	4. ( $\eta_{yt} > 0, \eta_{mt} > 0$ )
1. ( $\lambda_t = 0, \lambda_{t+1} = 0$ )	Infeasible	Origin
2. ( $\lambda_t = 0, \lambda_{t+1} > 0$ )	Boundary	Origin
3. ( $\lambda_t > 0, \lambda_{t+1} = 0$ )	Boundary	infeasible
4. ( $\lambda_t > 0, \lambda_{t+1} > 0$ )	Boundary	Infeasible

Table 4.2: The most relaxed possibility for each case in lemma C.1. The order of possibilities is Interior  $\Rightarrow$  Boundary  $\Rightarrow$  Origin  $\Rightarrow$  Infeasible. Each element of this list is a superset of the elements to the right, but not vice versa. Row  $i$  and column  $j$  refers to case  $(i,j)$  in lemma C.1

	1. ( $\eta_{yt} = 0, \eta_{mt} = 0$ )	2. ( $\eta_{yt} = 0, \eta_{mt} > 0$ )
1. ( $\lambda_t = 0, \lambda_{t+1} = 0$ )	Infeasible	$(x_{yt} \geq 0, x_{mt} = 0)$
2. ( $\lambda_t = 0, \lambda_{t+1} > 0$ )	$(x_{yt} = x_{m,t+1} \geq 0, x_{mt} \geq 0)$	$(x_{yt} = x_{m,t+1} \geq 0, x_{mt} = 0)$
3. ( $\lambda_t > 0, \lambda_{t+1} = 0$ )	$(x_{yt} \geq 0, x_{mt} = x_{y,t-1} \geq 0)$	Infeasible
4. ( $\lambda_t > 0, \lambda_{t+1} > 0$ )	$(x_{yt} = x_{m,t+1} \geq 0, x_{mt} = x_{y,t-1} \geq 0)$	Infeasible
	3. ( $\eta_{yt} > 0, \eta_{mt} = 0$ )	4. ( $\eta_{yt} > 0, \eta_{mt} > 0$ )
1. ( $\lambda_t = 0, \lambda_{t+1} = 0$ )	Infeasible	$(x_{yt} = 0, x_{mt} = 0)$
2. ( $\lambda_t = 0, \lambda_{t+1} > 0$ )	$(x_{yt} = x_{m,t+1} = 0, x_{mt} \geq 0)$	$(x_{yt} = x_{m,t+1} = 0, x_{mt} = 0)$
3. ( $\lambda_t > 0, \lambda_{t+1} = 0$ )	$(x_{yt} = 0, x_{mt} = x_{y,t-1} \geq 0)$	Infeasible
4. ( $\lambda_t > 0, \lambda_{t+1} > 0$ )	$(x_{yt} = x_{m,t+1} = 0, x_{mt} = x_{y,t-1} \geq 0)$	Infeasible

Table 4.3: A table of possibilities, testing the consistency of the solution to the KKTs against the maintained assumptions on  $\lambda$ 's and  $\eta$ 's, showing the values the land allocation variables may take in each case. Row  $i$  and column  $j$  refers to case  $(i,j)$  in lemma C.1

## When will young trees be planted in final period?

Having enumerated the possible cases each period, and developed necessary conditions for each case, I am now in a position to put lemma C.1 to work. I begin by asking under what circumstances a farmer would choose to plant young trees in the final period of his planning horizon. One's *a priori* intuition might be that the farmer plants young trees in the final period only if there is a sufficiently high salvage value of mature trees in the final period. This intuition is partly correct, but there are also circumstances for which a salvage value after the final period is not necessary.

**Proposition 4.2 (Planting young trees in period  $T$ )** *Assuming the aging constraint between period  $T$  and  $T + 1$  does not bind ( $\lambda_{T+1} = 0$ ), young trees will be planted in the final period ( $x_{yT} > 0$ ) only if either*

- $p \leq b$  and  $\lambda_T = 0$ , or
- $p < b$ ,  $\lambda_T > 0$  and  $x_{y,T-1} \leq \frac{b-p}{af_m}$

*Assuming the aging constraint between period  $T$  and  $T + 1$  binds ( $\lambda_{T+1} > 0$ ), young trees will be planted in the final period ( $x_{yT} > 0$ ) only if  $x_{m,T+1} \leq \frac{bf_m-p}{af_m}$ .*

This proposition is proved in appendix A.1. It shows that the opportunity cost of land must be low for young trees to be planted in the final period. A low opportunity cost of land means that the farmer forgoes little by planting lower yielding young trees. If the opportunity cost of land is high, then the farmer prefers to plant more productive mature trees (requiring more planting of young trees in earlier period), or choosing not to plant at all.

## Examples of qualitative trajectory types

Figure 4.3 shows four example trajectories, calculated using the constrained optimization solver in the Mathematica software package (ArgMax). There are several qualitatively different trajectory paths. Figure 4.3a shows a 3-period problem with a high opportunity cost of land, resulting in boundary conditions in the first and last period, and an interior condition in period 2. Total land is non-monotonic, increasing then decreasing. Figure 4.3b shows a 3-period problem with a low opportunity cost of land, resulting in boundary conditions in the first period, and interior conditions in the middle and final periods. Total land is non-monotonic, increasing then decreasing, but the increase is small relative to figure 4.3a. Figure 4.3c shows a 4-period problem with a high opportunity cost of land, resulting in boundary conditions in every period, and a constant quantity of total land. Figure 4.3c is qualitatively identical<sup>4</sup> to the first 4 periods of Wan's solution of the two-age class version of the model of

---

<sup>4</sup>The solution to the two-age-class model presented by Wan assumes one unit of land. Further assuming  $f_m > 1$ , the long-run solution is  $x_{m,t} = 1 - x_{m,t-1}$  [see remark in section 3 on page 418 of his paper]. This describes the solution in figure 4.3c, except that total land in my case is fixed at 0.92 units, so  $x_{mt} = 0.92 - x_{m,t-1}$ .

Mitra et al. (1991). In contrast to Wan I obtain constant land as a result, rather than as a maintained assumption of the model. Figure 4.3d shows a 4-period problem with a low opportunity cost of land. It is similar to 4.3a since the first and last periods are boundary, the middle periods are interior and the total quantity of land is non-monotonic. However this occurs in a low price, whereas figure 4.3a had a high price.

These examples show that I can generate a richer variety of land allocation trajectories, relative to the fixed land, infinite horizon model of Wan (1993). These are just four example trajectories, and there are infinitely many more trajectories I could generate by changing the parameter values. However, I am interested in the qualitative similarities between the trajectories. Can these trajectories be classified into a finite set of categories? I can use lemma C.1 to classify the qualitatively different types.

## Analyzing an example trajectory

I use lemma C.1 to analyze figure 4.3b, determining the case that pertains to each period in the example. I use backwards induction, starting at the final period and working back toward the initial period, to infer the shadow values on the aging and non-negativity constraints from observations on land allocations.

The trajectory in this example (calculated using the ArgMax function in Mathematica and rounded to one decimal place) is  $\{x_{y1} = 1.5, x_{m1} = 0, x_{y2} = 0.4, x_{m2} = 1.2, x_{y3} = 0.7,$  and  $x_{m3} = 0.4\}$ . By assumption  $S'(x_{m4}) = 0$ , so  $\lambda_4 = 0$ .

In the third period there is positive planting of both young and mature trees. Further, the quantity of land allocated to mature trees in period 3 is equal to the quantity of land planted with young trees in period 2. From these observations I infer that  $\lambda_3 > 0$ ,  $\lambda_4 = 0$ ,  $\eta_{yt} = 0$ , and  $\eta_{mt} = 0$ . This corresponds to case 3.1. Using the formulae in lemma C.1 for  $x_{y3}$  and  $\lambda_3$  in case 3.1, I calculated  $x_{y3} = 0.7$  and  $\lambda_3 = 0.625$ . This is consistent with the data, and the assumptions for this case.

In the second period there is again positive planting of young and mature trees, however, the quantity of mature trees in period 2 is less than the quantity of young trees in period 1. From these observations I infer that  $\lambda_2 = 0$ ,  $\lambda_3 > 0$ ,  $\eta_{yt} = 0$ , and  $\eta_{mt} = 0$ . This corresponds to case 2.1. Using the formulae in lemma C.1 for  $x_{m2}$  and  $\lambda_3$  in case 2.1, I calculated  $x_{m2} = 1.2$  and  $\lambda_2 = 0.625$ . This is consistent with the data, and the assumptions for this case.

In the first period there is positive planting of young trees and zero planting of mature trees. The initial condition is  $x_{y0} = 0$ . Hence the zero planting of mature trees is either induced by the aging constraint between period 0 and 1 (case 3.1), or by the non-negativity constraint on matures in period 1 (case 1.2). Either of these cases could be ruled out if the values of the free variables were calculated and found to be inconsistent with the assumptions in that case. Assuming case 3.1 holds, using the formulae in lemma C.1 for  $x_{y1}$  and  $\lambda_1$ , I calculated  $x_{y1} = 1.5$  and  $\lambda_1 = 0.625$ . This is consistent with the data, and the assumptions for this case. Assuming case 1.2 holds, using the formulae in lemma C.1 for  $x_{y1}$  and  $\eta_{m1}$ , I calculated  $x_{y1} = 1.5$  and  $\eta_{m1} = 0.625$ . This is also consistent with the data, and the assumptions for this case.

The result for period 1 represents a limitation of this method of analysis, namely, that it cannot distinguish between these two cases. The two cases are mutually exclusive, but from this analysis I cannot tell whether the objective function would be improved by increasing or decreasing  $x_{m1}$ .

## 4.4 Discussion and Future Work

Although this prototype model can be used to answer many of the questions posed in the introduction to this chapter, development of explicit answers is left for future work. However, I can provide preliminary answers to the questions, showing how this model could be used to answer them.

In this model the farmer chooses to plant perennials when the NPV of doing so is positive ( $V(x_{y0}) > 0$ ). The model includes the opportunity cost of land,  $p$ , so a positive NPV means that the value of the benefits he receives from perennial crop production exceeds the foregone benefits from allocating his land to perennials.

The choice between keeping an established perennial crop and switching to an alternative crop can be modeled by increasing the opportunity cost of land. Assuming the farmer starts with an established orchard ( $x_{y0} > 0$ ), I can use this model to find the price at which he chooses to shut down production of perennials. This represents any increase in the opportunity cost of land, not just increases from an alternative perennial crop.

The choice between two different perennials could be modeled more explicitly by comparing the value functions between the existing crop with a positive initial condition, and an alternative crop with a zero initial condition. Say the farmer is currently growing perennial crop  $a$ , and is considering switching to perennial crop  $b$ . In this framework the farmer will switch to crop  $b$  if  $V_b(0) > V_a(x_{y0})$ . This assumes away the presence of explicit switching costs and costly irreversibility, which can be quite important for perennial crops (Song et al., 2011).

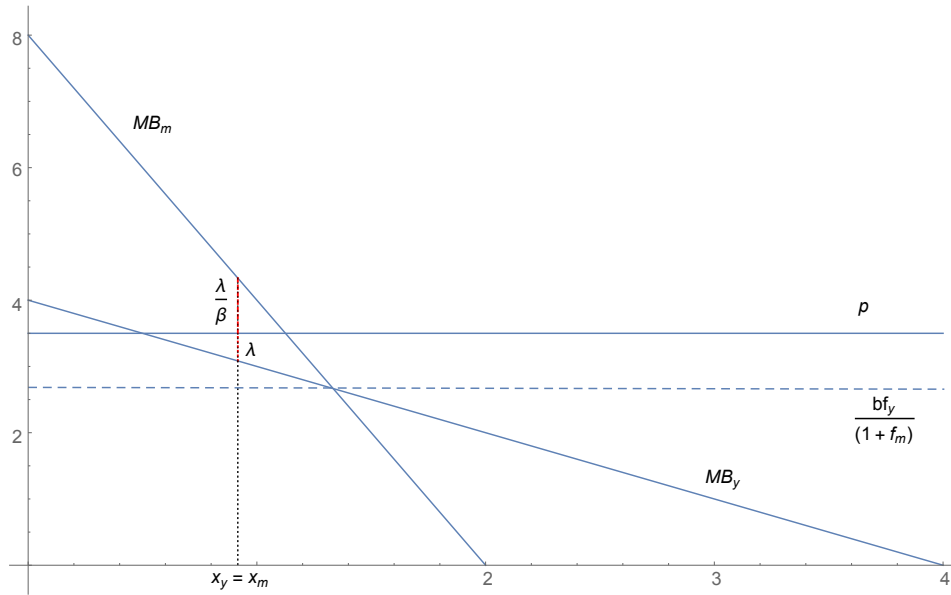
I can use lemma C.1 to determine when a farmer chooses to let a young tree age to become a mature tree by determining the conditions for which  $x_{mt} > 0$ . I demonstrated a method of analyzing a related question (when is  $x_{yT} > 0$ ?) in section 4.3.

Whether or not a farmer manages the orchard as if it were a single tree, or as an age-structured orchard, is determined by whether the optimal program has interior solutions in any period. If the solution to the farmer's problem consists entirely of boundary solutions—figure 4.3c, for example—the farmer is managing the orchard as if it were a single tree. Finding the conditions under which the solution to the farmer's problem consists entirely of boundary solutions will show the set of circumstances for which the single tree assumption, so common in the perennial crop economics literature, is justified.

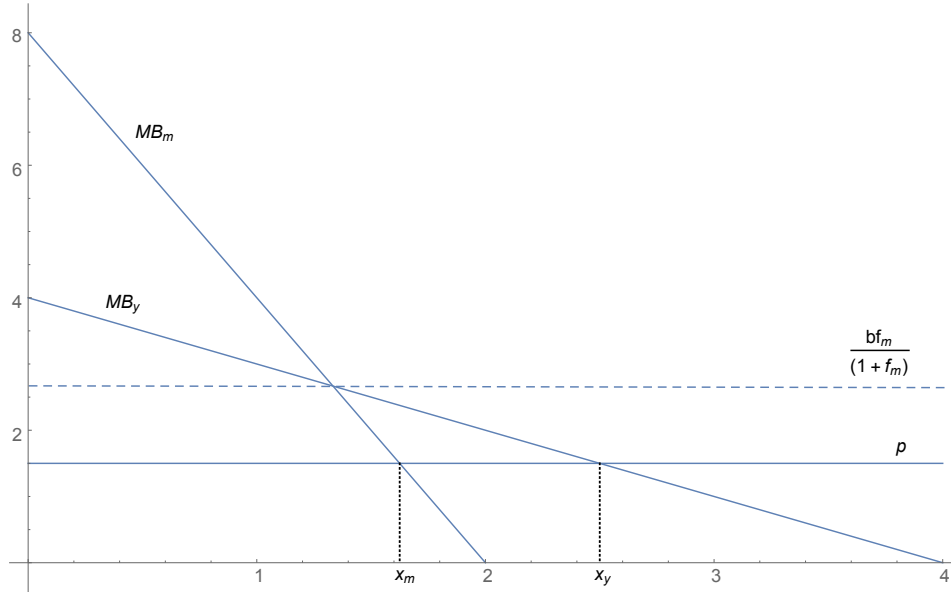
Processing of perennial crops can be incorporated into this model in at least two ways. The first alternative is to model the capacity constraint as a hard capacity constraint, where the model behaves as normal for quantities of perennial crop production less than the capacity constraint, but for production exceeding the capacity constraint the excess

production is either discarded, or provides benefits at a lower rate than the unconstrained case. A second alternative is to model the capacity constraint as a loss function centered around the biorefinery's cost-minimizing scale of production. In either case, optimal biorefinery capacity is determined as a two stage model, where the optimal planting sequence is determined conditional on the capacity given, and then the optimal capacity is chosen, conditional on the subsequent optimal land allocation of the farmer.

This chapter has presented a land-allocation model for perennial crops, with two possible age-classes. In addition to developing the issues just described, in future work I intend to expand the number of active age-classes to at least 3, allowing me to test the importance of the assumption of increasing yields used by Mitra et al. (1991) to develop their neighborhood turnpike result. I also intend to make a more explicit comparison between the infinite horizon models of Mitra et al. (1991) and Wan (1993), and this model, testing my conjecture that their models can be nested within mine by the correct choice of the salvage value function.



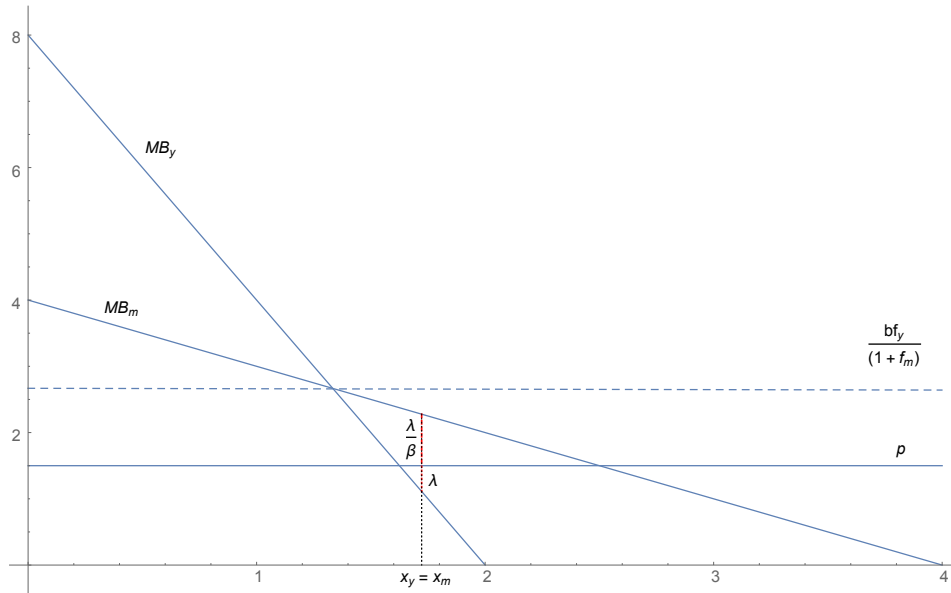
(a) An example of the aging constraint binding when  $f_m > 1$ . Here  $f_m = 2$ , and  $p = 3.5$ . Note that  $p$  is above the binding threshold  $\frac{bf_m}{1+f_m}$ .



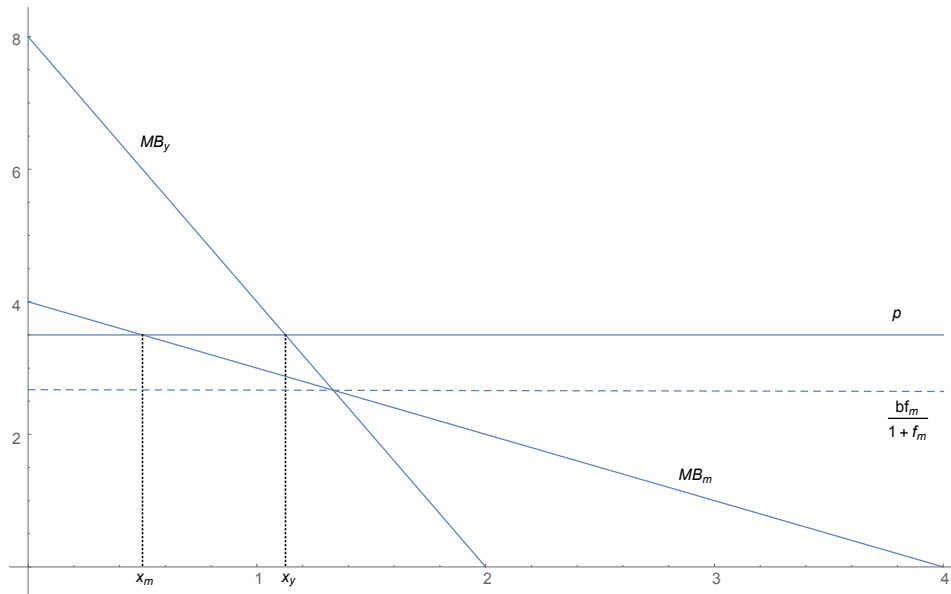
(b) An example of the aging constraint not binding when  $f_m > 1$ . Here  $f_m = 2$ , and  $p = 1.5$ . Note that  $p$  is below the binding threshold  $\frac{bf_m}{1+f_m}$ .

Figure 4.1: Binding and non-binding aging constraints with  $f_m > 1$ . The horizontal axis is the land allocated to perennials,  $x$ , and the vertical axis measures marginal benefit. Here I assume linear marginal benefit  $u(f_n x_n) = b - a(f_n x_n)$ , where  $a = 1$  and  $b = 4$ . Also  $\beta = 0.5$ .





(a) An example of the aging constraint binding when  $f_m < 1$ . Here  $f_m = \frac{1}{2}$ , and  $p = 1.5$ . Note that  $p$  is below the binding threshold  $\frac{bf_m}{1+f_m}$ .



(b) An example of the aging constraint not binding when  $f_m < 1$ . Here  $f_m = \frac{1}{2}$ , and  $p = 3.5$ . Note that  $p$  is above the binding threshold  $\frac{bf_m}{1+f_m}$ .

Figure 4.2: Binding and non-binding aging constraints with  $f_m < 1$ . The horizontal axis is the land allocated to perennials,  $x$ , and the vertical axis measures marginal benefit. Here I assume linear marginal benefit  $u(f_n x_n) = b - a(f_n x_n)$ , where  $a = 1$  and  $b = 4$ . Also  $\beta = 0.5$ .

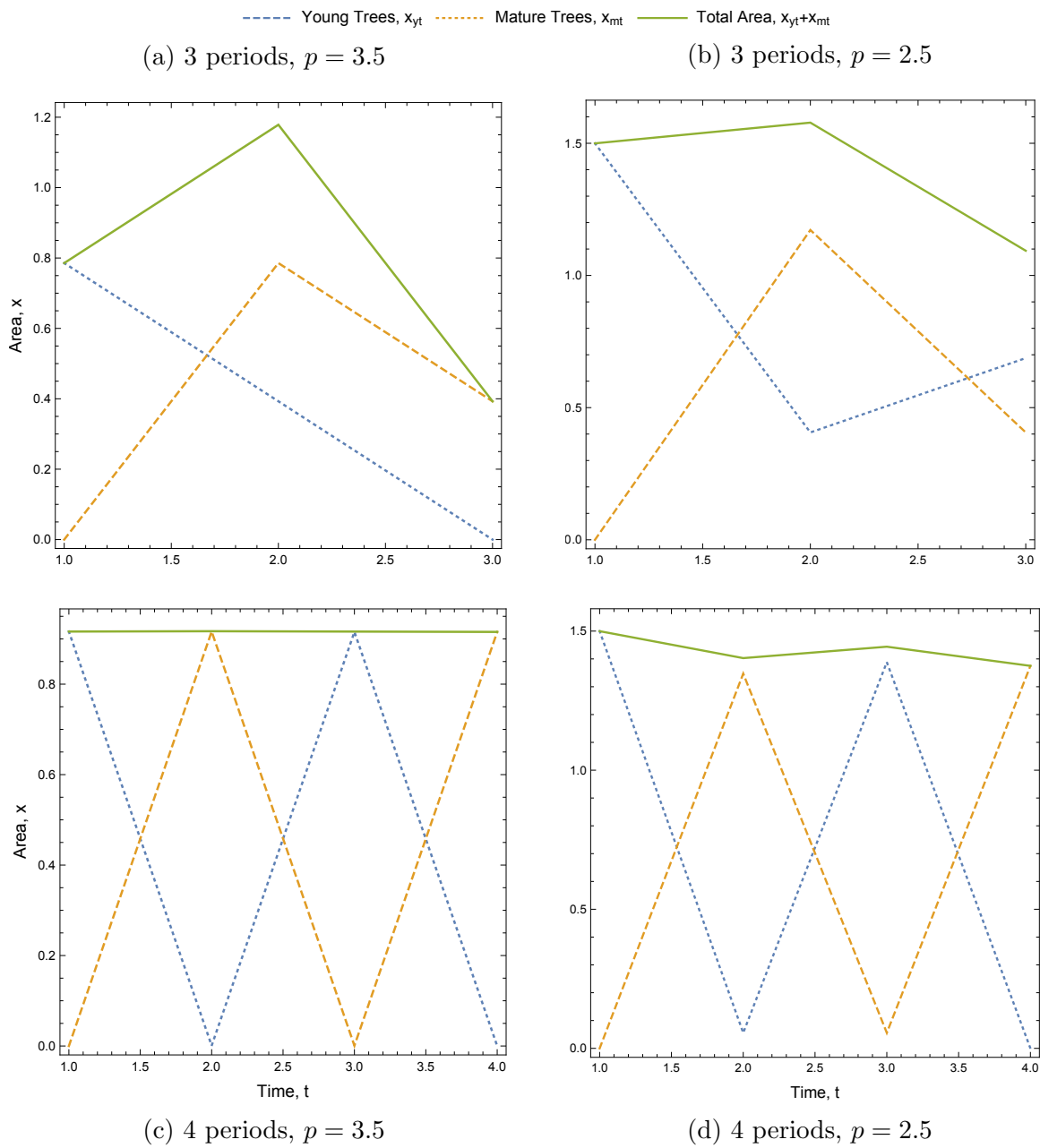


Figure 4.3: Four example trajectories generated by changing the time horizon and opportunity cost of land. Apart from price, the parameter values are the same as those used to generate figure 4.1, that is  $u(f_n x_n) = b - a(f_n x_n)$ , with  $a = 1$  and  $b = 4$ ,  $f_m = 2$ ,  $\beta = 0.5$ . Additionally,  $x_{y0} = 0$  and  $S'(x_{m,T+1}) = 0$ .

# Bibliography

- M. Alauddin and C. Tisdell. Decomposition Methods, Agricultural Productivity Growth and Technological Change: A Critique Supported by Bangladeshi Data. *Oxford Bulletin of Economics and Statistics*, 48(4):353–372, 1986.
- D. A. Ali, K. Deininger, and M. Duponchel. Credit Constraints and Agricultural Productivity: Evidence from Rural Rwanda. *Journal of Development Studies*, 50(5):649–665, 2014. ISSN 0022-0388. doi: 10.1080/00220388.2014.887687. URL <http://www.tandfonline.com/doi/abs/10.1080/00220388.2014.887687>.
- B. A. Babcock. Extensive and Intensive Agricultural Supply Response. *Annual Review of Resource Economics*, 7:333–348, 2015. ISSN 1941-1340. doi: 10.1146/annurev-resource-100913-012424. URL <http://dx.doi.org/10.1146/annurev-resource-100913-012424>.
- N. Barbosa. Financial regulation and the Brazilian response to the 2008-2009 financial crisis. *Revista de Economia Política*, 31(125):848–852, 2011. ISSN 0101-3157. doi: 10.1590/S0101-31572011000500013.
- P. J. Barry and L. J. Robison. Agricultural Finance: Credit, Credit Constraints, and Consequences. In B. Gardner and G. Rauser, editors, *Handbook of Agricultural Economics*, volume 1, pages 513–571. Elsevier Science B.V., 2001.
- R. E. Bellman and M. J. Hartley. The Tree-Crop Problem. Technical report, World Bank, 1985.
- E. M. Berry, Y. Arnoni, and M. Aviram. The Middle Eastern and biblical origins of the Mediterranean diet. *Public health nutrition*, 14(12A):2288–95, 2011. ISSN 1475-2727. doi: 10.1017/S1368980011002539. URL <http://www.ncbi.nlm.nih.gov/pubmed/22166186>.
- G. Bocquého and F. Jacquet. The adoption of switchgrass and miscanthus by farmers: Impact of liquidity constraints and risk preferences. *Energy Policy*, 38(5):2598–2607, 2010. ISSN 03014215. doi: 10.1016/j.enpol.2010.01.005.
- M. Bonomo, R. Brito, and B. Martins. Macroeconomic and Financial Consequences of the After Crisis Government-Driven Credit Expansion in Brazil. 2014.

- M. P. Brady and T. L. Marsh. Do Changes in Orchard Supply Occur at the Intensive or Extensive Margin of the Landowner? 2013.
- R. J. Brazee. The Faustmann face of optimal forest harvesting. In *The International Yearbook of Environmental and Resource Economics 2006/2007*, chapter 6, page 312. 2007.
- B. C. Briggeman, C. A. Towe, and M. J. Morehart. Credit constraints: Their existence, determinants, and implications for U.S. farm and nonfarm sole proprietorships. *American Journal of Agricultural Economics*, 91(1):275–289, 2009. ISSN 00029092. doi: 10.1111/j.1467-8276.2008.01173.x.
- W. A. Brock and W. D. Dechert. Growth Models, Multisector. In S. N. Durlauf and L. E. Blume, editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, second edition, 2008. ISBN 9789814407724. doi: 10.1142/9789814407724\_0018.
- D. Brough and R. Ewing. Global sugar mills retrench as depressed prices take toll. *Reuters*, page 2017, oct 2014. URL <http://uk.reuters.com/article/2014/10/08/sugar-mills-idUKL6NOS31SU20141008>.
- K. Burger and H. P. Smit. The Impact of Prices and Technology in the Replanting of Perennial Crops. In S. Bauer, R. Herrmann, and F. Kuhlmann, editors, *Märkte der Agrar- und Ernährungswirtschaft*, pages 263–271. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Münster-Hiltrup, 1997.
- California Strawberry Commission. About Strawberries, 2017. URL [http://www.californiastrawberries.com/about{}\\_strawberries](http://www.californiastrawberries.com/about{}_strawberries).
- O. Canuto. The Impact of the Financial Crisis on Developing Countries. Technical report, Institute of Development Studies, Sussex, 2008. URL <https://www.ids.ac.uk/files/dmfile/BRAZIL0tavianoCanuto.pdf>.
- O. Canuto. From Crisis Management to Long Term Growth Drivers: Perspectives from Brazil, 2010. URL [http://siteresources.worldbank.org/INTPREMNET/Resources/0C{}\\_Russia{}\\_April012010.pdf](http://siteresources.worldbank.org/INTPREMNET/Resources/0C{}_Russia{}_April012010.pdf).
- G. C. Chow. Optimal control without solving the Bellman equation. *Journal of Economic Dynamics and Control*, 17(4):621–630, 1993. ISSN 01651889. doi: 10.1016/0165-1889(93)90049-X. URL <http://linkinghub.elsevier.com/retrieve/pii/016518899390049X>.
- G. C. Chow. *Dynamic Economics*. Oxford University Press, New York, 1997.
- T. Cohen. Brazil sugar production costs rise. *Financial Times*, 2012.
- K. Cortes. Brazil May Boost Farm Loans 28%, Banco do Brasil Says. *The Brazilian-American Chamber of Commerce*, 2009. URL <http://brazilcham.com/articles/brazil-may-boost-farm-loans-28-banco-do-brasil-says>.

- C. L. Crago, M. Khanna, J. Barton, E. Giuliani, and W. Amaral. Competitiveness of Brazilian sugarcane ethanol compared to US corn ethanol. *Energy Policy*, 38(11):7404–7415, 2010. ISSN 03014215. doi: 10.1016/j.enpol.2010.08.016. URL <http://dx.doi.org/10.1016/j.enpol.2010.08.016>.
- E. Crooks and G. Meyer. Brazilian imports of US ethanol soar. *Financial Times*, may 2011.
- H. J. T. da Silva. Ethanol Production Costs in Brazil Fall Not Enough to Generate Profits. *World Ethanol & Biofuels Report*, 14(13):216–218, 2016.
- A. R. R. de Mendonça and S. Deos. Facing the Crisis: Brazilian Central Bank and Public Banking System as Minskyan “Big Banks”. 2013.
- A. De Meyer, D. Cattrysse, J. Rasinmäki, and J. Van Orshoven. Methods to optimise the design and management of biomass-for-bioenergy supply chains: A review. *Renewable and Sustainable Energy Reviews*, 31:657–670, 2014. ISSN 13640321. doi: 10.1016/j.rser.2013.12.036. URL <http://dx.doi.org/10.1016/j.rser.2013.12.036>.
- M. A. F. D. de Moraes and D. Zilberman. *Production of Ethanol from Sugarcane in Brazil From State Intervention to a Free Market*. Springer International Publishing, 2014. ISBN 9783319031392.
- D. Debnath, F. M. Epplin, and A. L. Stoecker. Managing Spatial and Temporal Switchgrass Biomass Yield Variability. *Bioenergy Research*, 7(3):946–957, 2014. ISSN 19391242. doi: 10.1007/s12155-014-9436-6.
- D. Debnath, F. M. Epplin, and A. L. Stoecker. Switchgrass procurement strategies for managing yield variability: Estimating the cost-efficient D (downtime cost) L (land to lease) frontier. *Biomass and Bioenergy*, 77:110–122, 2015. ISSN 1873-2909. doi: 10.1016/j.biombioe.2015.03.024. URL <http://dx.doi.org/10.1016/j.biombioe.2015.03.024>.
- S. Devadoss and J. Luckstead. An analysis of apple supply response. *International Journal of Production Economics*, 124(1):265–271, 2010. ISSN 09255273. doi: 10.1016/j.ijpe.2009.11.024. URL <http://dx.doi.org/10.1016/j.ijpe.2009.11.024>.
- F. Dong. Effects of credit constraints on household productivity in rural China. *Agricultural Finance Review*, 72(3):402–415, 2012. ISSN 0264-0473. doi: <http://dx.doi.org/10.1108/09564230910978511>.
- D. M. Drukker. Testing for serial correlation in linear panel-data models. *Stata Journal*, 3: 168–177, 2003.
- R. Ewing. Brazil ethanol returns to US as biofuel rules pave way. *Reuters*, sep 2013a.
- R. Ewing. Policy, Consolidation and Investment Trends in Brazilian Biofuels. *Reuters*, 2013b.

- R. Ewing. Brazil's rapid recovery in sugar output to slow. *Reuters*, oct 2013c.
- R. Ewing. Big crop, policy revive Brazil's sugar-ethanol mills. *Reuters*, feb 2013d.
- R. Ewing. Brazil's big cane mills cut costs, capex against losses. *Reuters*, 2014.
- G. Fabbri, S. Faggian, and G. Freni. On the Mitra-Wan Forest Management Problem in Continuous Time. *Journal of Economic Theory*, 157:1001–1040, 2015. ISSN 10957235. doi: 10.1016/j.jet.2015.03.004. URL <http://dx.doi.org/10.1016/j.jet.2015.03.004>.
- E. Feinerman and Y. Tsur. Perennial crops under stochastic water supply. *Agricultural Economics*, 45:1–10, 2014. ISSN 01695150. doi: 10.1111/agec.12120. URL <http://doi.wiley.com/10.1111/agec.12120>.
- F. F. Filho. Brazil's response: How did financial regulation and monetary policy influence recovery? *Brazilian Journal of Political Economy*, 31(125):880–888, 2011.
- B. Franklin. *A Dynamic Regional Model of Irrigated Perennial Crop Production*. Phd dissertation, UC Riverside, 2012.
- B. C. French and R. G. Bressler. The Lemon Cycle. *Journal of Farm Economics*, 44(4): 1021–1036, 1962.
- B. C. French and J. L. Matthews. A Supply Response Model for Perennial Crops. *American Journal of Agricultural Economics*, 53(3):478–490, 1971.
- B. C. French, G. A. King, and D. D. Minami. Planting and Removal Relationships for Perennial Crops: An Application to Cling Peaches. *American Journal of Agricultural Economics*, 67(2):215–223, 1985.
- Globo. Brasil quer ser 'Arábia Saudita' do biocombustível, diz Lula, mar 2007. URL <http://g1.globo.com/Noticias/Economia/0,,MUL14025-9356,00-BRASIL+QUER+SER+ARABIA+SAUDITA+DO+BIOCOMBUSTIVEL+DIZ+LULA.html>.
- J. D. Glover, J. P. Reganold, L. W. Bell, J. Borevitz, E. C. Brummer, E. S. Buckler, C. M. Cox, T. S. Cox, T. E. Crews, S. W. Culman, L. R. DeHaan, D. Eriksson, B. S. Gill, J. Holland, F. Hu, B. S. Hulke, A. M. H. Ibrahim, W. Jackson, S. S. Jones, S. C. Murray, A. H. Paterson, E. Ploschuk, E. J. Sacks, S. Snapp, D. Tao, D. L. Van Tassel, L. J. Wade, D. L. Wyse, and Y. Xu. Increased Food and Ecosystem Security via Perennial Grains. *Science*, 328(5986):1638–1639, jun 2010. URL <http://science.sciencemag.org/content/328/5986/1638.abstract>.
- N. Gotsch and M. Wohlgenant. A welfare analysis of biological technical change under different supply shift assumptions: the case of Cocoa in Malaysia. *Canadian Journal of Agricultural ...*, 49:87–104, 2001. ISSN 0008-3976. doi: 10.1111/j.1744-7976.2001.tb00292.x. URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1744-7976.2001.tb00292.x/abstract>.

- C. Guirkinger and S. R. Boucher. Credit constraints and productivity in Peruvian agriculture. *Agricultural Economics*, 39(3):295–308, 2008. ISSN 01695150. doi: 10.1111/j.1574-0862.2008.00334.x.
- D. T. Handley. Profitable Matted Row Strawberries – Getting Started. In *Proceeding of the New England Vegetable and Fruit Conference*, 2011. URL <http://www.newenglandvfc.org/2011{ }conference/proceedings2011.html>.
- J. M. Haworth and P. J. Vincent. Medium Term Forecasting of Orchard Fruit Production in the EEC: Methods and Analyses. Technical report, Eurostat, Brussels, 1977.
- G. James, editor. *Sugarcane*. Blackwell Science, 2004.
- B. M. Jenkins. A comment on the optimal sizing of a biomass utilization facility under constant and variable cost scaling. *Biomass and Bioenergy*, 13(1-2):1–9, 1997. ISSN 09619534. doi: 10.1016/S0961-9534(97)00085-8.
- R. E. Just and R. D. Pope. Agricultural Production. *Handbook of Agricultural Economics*, 1:629–741, 2001. ISSN 15740072. doi: 10.1016/S1574-0072(01)10015-0. URL <http://www.sciencedirect.com/science/article/pii/S1574007201100150>.
- N. G. Kalaitzandonakes and J. S. Shonkwiler. A State-Space Approach to Perennial Crop Supply Analysis. *American Journal of Agricultural Economics*, 74(2):343–352, 1992.
- E. W. Kemmerer. Agricultural Credit in the United States. *The American Economic Review*, 2(4):852–872, 1912.
- M. Khanna and C. L. Crago. Measuring Indirect Land Use Change with Biofuels: Implications for Policy. *Annual Review of Resource Economics*, 4:161–184, 2012. doi: 10.1146/annurev-resource-110811-114523.
- K. C. Knapp. Dynamic Equilibrium in Markets for Perennial Crops. *American Journal of Agricultural Economics*, 69(1):97–105, 1987.
- K. C. Knapp and K. Konyar. Perennial Crop Supply Response: A Kalman Filter Approach. *American Journal of Agricultural Economics*, 73(3):841–849, 1991.
- F. Mafakheri and F. Nasiri. Modeling of biomass-to-energy supply chain operations: Applications, challenges and research directions. *Energy Policy*, 67:116–126, 2014. ISSN 03014215. doi: 10.1016/j.enpol.2013.11.071. URL <http://dx.doi.org/10.1016/j.enpol.2013.11.071>.
- M. Majumdar. Multisector Growth Models. In J. Eatwell, M. Milgate, and P. Newman, editors, *Palgrave Dictionary of Economics*, pages 6878–6884. Palgrave Macmillan, first edition, 1987. ISBN 978-0-387-56892-8. doi: 10.1007/978-0-387-77358-2. URL <http://www.springerlink.com/index/10.1007/978-0-387-77358-2>.

- F. B. Margarido and F. Santos. Agricultural Planning. In F. Santos, A. Borém, and C. Caldas, editors, *Sugarcane Bioenergy, Sugar and Ethanol – Technology and Prospects*, chapter 1, pages 7–21. Ministry of Agriculture, Livestock and Food Supply, 2012.
- M. L. Mendonça, F. T. Pitta, and C. V. Xavier. The Sugarcane Industry and the global economic crisis. Technical Report June, Transnational Institute and the Network for Social Justice and Human Rights, 2013.
- R. Miao and M. Khanna. Costs of Meeting the Cellulosic Biofuel Mandate with Perennial Energy Crops: Implications for Policy. 2015. URL <https://ssrn.com/abstract=2618867>.
- M. J. Miranda and P. L. Fackler. *Applied Computational Economics and Finance*. MIT Press, 2002.
- T. Mitra and H. Y. Wan. Some Theoretical Results on the Economics of Forestry. *The Review of Economic Studies*, 52(2):263–282, 1985.
- T. Mitra and H. Y. Wan. On the Faustmann solution to the forest management problem. *Journal of Economic Theory*, 40(2):229–249, 1986. ISSN 00220531. doi: 10.1016/0022-0531(86)90073-6. URL <http://linkinghub.elsevier.com/retrieve/pii/0022053186900736>.
- T. Mitra, D. Ray, and R. Roy. The economics of orchards: an exercise in point-input, flow-output capital theory. *Journal of Economic Theory*, 53:12–50, 1991. URL <http://www.sciencedirect.com/science/article/pii/002205319190141P>.
- T. Molnar, P. Kahn, T. Ford, C. Funk, and C. Funk. Tree Crops, a Permanent Agriculture: Concepts from the Past for a Sustainable Future. *Resources*, 2(4):457–488, 2013. ISSN 2079-9276. doi: 10.3390/resources2040457. URL <http://www.mdpi.com/2079-9276/2/4/457/>
- M. Moreira. Stop and Go in Brazilian Ethanol Sector, 2015.
- M. H. Nguyen and R. G. H. Prince. A simple rule for bioenergy conversion plant size optimisation: Bioethanol from sugar cane and sweet sorghum. *Biomass and Bioenergy*, 10(5-6):361–365, 1996. ISSN 09619534. doi: 10.1016/0961-9534(96)00003-7.
- R. P. Overend. The average haul distance and transportation work factors for biomass delivered to a central plant. *Biomass*, 2(1):75–79, 1982. ISSN 01444565. doi: 10.1016/0144-4565(82)90008-7.
- T. Philpott. What America’s nutty demand for almonds is doing to California, 2015. URL <http://grist.org/food/what-americas-nutty-demand-for-almonds-is-doing-to-california/>.
- M. Pollan. *The Botany of Desire*. Random House, New York, 2001.



- G. C. Rausser. *A Dynamic Econometric Model of the California-Arizona Orange Industry*. Ph.d. thesis, University of California, 1971.
- A. A. Rentizelas, A. J. Tolis, and I. P. Tatsiopoulou. Logistics issues of biomass: The storage problem and the multi-biomass supply chain. *Renewable and Sustainable Energy Reviews*, 13(4):887–894, 2009. ISSN 13640321. doi: 10.1016/j.rser.2008.01.003.
- F. R. Riley. Olive Oil Production on Bronze Age Crete: Nutritional Properties, Processing Methods and Storage Life of Minoan Olive Oil. *Oxford Journal of Archaeology*, 21(1):63–75, 2002. ISSN 1468-0092. doi: 10.1111/1468-0092.00149.
- J. Rust. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5):999–1033, 1987.
- S. Salo and O. Tahvonen. Renewable Resources with Endogeneous Age Classes and Allocation of Land. *American Journal of Agricultural Economics*, 86(2):513–530, 2004.
- B. Sharma, R. G. Ingalls, C. L. Jones, and A. Khanchi. Biomass supply chain design and analysis: Basis, overview, modeling, challenges, and future. *Renewable and Sustainable Energy Reviews*, 24:608–627, 2013. ISSN 13640321. doi: 10.1016/j.rser.2013.03.049. URL <http://dx.doi.org/10.1016/j.rser.2013.03.049>.
- J. R. Smith. *Tree Crops: A Permanent Agriculture*. Harcourt Brace, New York, 1929.
- R. Sobreira and L. F. de Paula. The 2008 Financial Crisis and Banking Behavior in Brazil: The Role of the Prudential Regulation. *Journal of Innovation Economics and Management*, 2(6):77–93, 2010. doi: 10.3917/jie.006.0077.
- F. Song, J. Zhao, and S. M. Swinton. Switching to Perennial Energy Crops under Uncertainty and Costly Reversibility. *American Journal of Agricultural Economics*, 93(3):764–779, 2011. ISSN 00029092. doi: 10.1093/ajae/aar018.
- J. E. Stiglitz and A. Weiss. Credit Rationing in Markets with Imperfect Information. *The American Economic Review*, 71(3):393–410, 1981.
- K. Sydsæter, A. Strøm, and P. Berck. *Economists’ Mathematical Manual*. Springer, New York, NY, 3 edition, 2005.
- F. L. D. S. Teixeira. *Custo Médio Operacional da Lavoura de Cana-de-açúcar em Reais*, 2013.
- The Economist. The credit crunch reaches Brazil Inc. nov 2008.
- The Economist. Brazilian Brew, jan 2012. ISSN 00130613.
- C. A. Tisdell and N. T. M. H. De Silva. Supply-Maximising and Variation-Minimizing Replacement Cycles of Perennial Crops and Similar Assets: Theory Illustrated by Coconut Cultivation. *Agricultural Economics*, 37(2):243–251, 1986.

- P. K. Trivedi. A Framework for Studying the Supply Response of Perennial Crops. Technical report, The World Bank, 1986.
- P. K. Trivedi. On Understanding Investment Behavior in Perennial Crops Production. Technical Report October, The World Bank, 1987.
- UNICA. Final report of 2013 / 2014 harvest season South-Central region. Technical report, UNICA, 2014.
- UNICA. Sugarcane: One Plant, Many Solutions. Technical report, UNICA, São Paulo, 2015.
- UNICA. Virtual Mill, 2016. URL <http://www.unica.com.br/virtual-mill/>.
- USDA. Fruit and tree nuts lead the growth of horticultural production value, 2016a. URL <https://www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?chartId=76958>.
- USDA. Crops Account for the Largest Share of the Value of U.S. Agricultural Production, 2016b. URL <https://www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?chartId=58328>.
- C. Valdes. Brazil's Ethanol Industry: Looking Forward. Technical report, USDA-ERS, 2011.
- E. C. Venegas and V. W. Ruttan. An Analysis of Rice Production in the Philippines. *Economic Research Journal*, 11(1):159–78, 1964.
- P. Vossen. Olive Oil: History, Production, and Characteristics of the World's Classic Oils. *HortScience*, 42(5):1093–1100, 2007. ISSN 00185345.
- J. Wallace. Increasing agricultural water use efficiency to meet future food production. *Agriculture, Ecosystems & Environment*, 82:105–119, 2000. ISSN 01678809. doi: 10.1016/S0167-8809(00)00220-6. URL <http://linkinghub.elsevier.com/retrieve/pii/S0167880900002206>.
- H. Wan. Revisiting the Mitra-Wan Tree Farm. *International Economic Review*, 35(1):193–198, 1994.
- H. Y. Wan. A Note on Boundary Optimal Paths. In *General Equilibrium, Growth, and Trade II: The Legacy of Lionel McKenzie*, pages 411–426. 1993. ISBN 9780120846559. doi: 10.1016/B978-0-12-084655-9.50021-3.
- J. Wesseler. *Calculating the Economic Value of a Fruit Tree Orchard*. Deutsche Gesellschaft für technische Zusammenarbeit, GTZ, Eschborn, 1997.
- J. Wheatley. Brazil farmers lose debt battle. *Financial Times*, pages 13–15, dec 2008.

M. Wright and R. C. Brown. Establishing the optimal sizes of different kinds of biorefineries. *Biofuels, Bioproducts and Biorefining*, 1(3):191–200, 2007. ISSN 1932104X. doi: 10.1002/bbb.

# Appendix A

## Appendix to Chapter 2

### A.1 Formulae for Calculating the Yield Change following a Discrete Change in the Replant Rate that Changes the Number of Active Age-Classes

**The effect of a discrete increase in the renewal rate that reduces the number of active age-classes by one**

Unlike the marginal change case, a discrete change in the replant rate from  $R$  to  $R'$  can change the number of active age-classes.

Here I show the effect of an *increase* in the replant rate on the yield transition trajectory.

Say that at time  $t$  there are  $s + 1$  active age-classes (where  $s + 1 = \lceil \frac{1}{R} \rceil$ ), and that at time  $t + 1$  the number of active age-classes declines to  $s$ . What is the change in yield?

The yield at time  $t$  is

$$yield_t = f_0R' + \dots + f_tR' + f_{t+1}R + \dots + f_{s-1}R + f_s(1 - R(s - t) - R'(t))$$

and the yield at time  $t + 1$  is

$$yield_{t+1} = f_0R' + \dots + f_tR' + f_{t+1}R' + \dots + f_{s-1}(1 - R((s - 1) - (t + 1)) - R'(t + 1))$$

Notice how the oldest active age-class at  $t + 1$  is now  $s - 1$ , and that in the  $s - 1$  land allocation equation the  $R$  term is multiplied by  $(s - 1) - (t + 1)$ . This is because there are now  $s - 1$  other active age-classes.

The change in yield between  $t$  and  $t + 1$  is given by the difference between these two expressions

$$\Delta yield_t = f_{t+1}R' - f_{t+1}R + f_{s-1}(1 - R((s-1) - (t+1)) - R'(t+1)) - f_{s-1}R - f_s(1 - R(s-t) - R'(t))$$

which, after simplifying, becomes

$$\Delta yield_t = (f_{t+1} - f_{s-1})\Delta R + (f_{s-1} - f_s)(1 - R(s - t) - R'(t))$$

The first term in this expression is the ‘within age-class yield effect’ and the second term is the ‘between age-class yield effect’ which exists due to the change in the number of active age-classes. Notice that the ‘within yield effect’ is not exactly the same as the case when there was no change in the number of age-classes. The yield of the  $t + 1^{\text{th}}$  age-class is now being compared to the  $s - 1^{\text{th}}$  age-class, not the  $s^{\text{th}}$ .

## The effect of a discrete increase in the replant rate that reduces the number of active age-classes by $n$

The change in the replant rate must be big enough to change the number of active age-classes by  $n$  in one time step, otherwise the formula in section A.1 is sufficient with a redefinition of  $s$  each time step.

Say that at time  $t$  there are  $s + 1$  active age-classes, and that at time  $t + 1$  the number of active age-classes declines to  $s + 1 - n$ . What is the change in yield?

The yield at time  $t$  is

$$yield_t = f_0R' + \dots + f_tR' + f_{t+1}R + \dots + f_{s-1}R + f_s(1 - R(s - t) - R'(t))$$

and the yield at time  $t + 1$  is

$$yield_{t+1} = f_0R' + \dots + f_tR' + f_{t+1}R' + \dots + f_{s-n}(1 - R((s - n) - (t + 1)) - R'(t + 1))$$

The change in yield between  $t$  and  $t + 1$  is given by the difference between these two expressions

$$\begin{aligned} \Delta yield_t &= f_{t+1}R' - f_{t+1}R \\ &\quad + f_{s-n}(1 - R((s - n) - (t + 1)) - R'(t + 1)) - f_{s-n}R - f_{s-n+1}R - \dots \\ &\quad - f_{s-1}R - f_s(1 - R(s - t) - R'(t)) \end{aligned}$$

which, after simplifying, becomes

$$\Delta yield_t = (f_{t+1} - f_{s-n})\Delta R + (f_{s-n}(n - 1) - \sum_{i=1}^{n-1} f_{s-n+i})R + (f_{s-n} - f_s)(1 - R(s - t) - R'(t))$$

# Appendix B

## Appendix to Chapter 3

### A.1 Proofs of Propositions and Lemmas

**Lemma B.1** *The Hoerl function satisfies assumptions (3.1) - (3.4)*

**Proof of Lemma B.1.** Let  $f(x) = a x^b e^{cx}$

1.  $f(x)$  is continuous since it is the product of two continuous function,  $ax^b$  and  $e^{cx}$ , and the product of continuous functions is continuous.
2.  $f(0) = 0$  since  $f(0) = a 0^b e^0 = 0$
3.  $f(x)$  monotonically increases to a unique maximum and then monotonically declines to a minimum if and only if  $a > 0$ ,  $b > 0$ , and  $c < 0$ .

If  $a = 0$  then the function is always zero and the condition cannot be obtained. If  $a < 0$ ,  $f(x) < 0$  for  $x > 0$ . Hence  $a > 0$ . If  $b = 0$ , then the  $f(x)$  behaves as an exponential function, which has no local extrema. If  $b < 0$  then  $f(0)$  is undefined, violating assumption (3.2). Hence  $b > 0$ . If  $c = 0$  the  $f(x)$  behaves as a polynomial  $ax^b$ , which for  $b > 0$  monotonically increases, but has no local maximum for  $x > 0$ . If  $c > 0$ , the function approaches infinity as  $x$  approaches infinity, violating assumption (3.4).

The derivative of  $f(x)$  is  $f'(x) = ax^b e^{cx} (\frac{b}{x} + c)$ .  $f(x)$  is initially increasing since there always exists some  $x$  sufficiently close to 0 such that  $\frac{b}{x} > -c$ . There exists a unique maximum of  $f(x)$  at  $x = \frac{-b}{c}$ . For  $x < \frac{-b}{c}$ , the derivative is always positive, so  $f(x)$  is monotonically increasing. For  $x > \frac{-b}{c}$ , the derivative is always negative, so  $f(x)$  is monotonically declining.

Hence I have shown  $a > 0$ ,  $b > 0$ ,  $c < 0$  is necessary and sufficient for  $f(x)$  to exhibit a monotonic increase to a unique maximum and a monotonic decrease thereafter.

4.  $\lim_{x \rightarrow \infty} x f(x) = \lim_{x \rightarrow \infty} a x^{b+1} e^{cx} = 0$  since  $e^{cx}$  approaches 0 more rapidly than  $x^{b+1}$  approaches infinity. This can be shown by repeated applications of L'Hôpital's rule.

■

**Lemma B.2** *The optimal maximum orchard age must be greater than or equal to the maximum yield age, i.e.  $n^* \geq n_{MY}$ .*

**Proof of Lemma B.2.**

From my assumptions on the age-yield relationship in section 3.2, for all yields less than the maximum yield, there are two orchard ages that generate that yield. That is, for all  $\bar{y} \in (0, y(n_{MY}))$ , there exist  $n_{\bar{y}}^- < n_{MY} < n_{\bar{y}}^+$ , such that  $y(n_{\bar{y}}^-) = y(n_{\bar{y}}^+) = \bar{y}$ .

On the graph of the isoquant, these two  $n$  values generate the same area,  $\bar{L}$ , since  $L = \frac{\bar{Q}}{y(n)}$  so  $\frac{\bar{Q}}{y(n_{\bar{y}}^-)} = \frac{\bar{Q}}{y(n_{\bar{y}}^+)} = \bar{L}$ .

Now compare the costs of these two  $n$  values.

$$\begin{aligned} C(n_{\bar{y}}^-, \bar{L}) - C(n_{\bar{y}}^+, \bar{L}) &= \left( C_f + \frac{C_n}{n_{\bar{y}}^-} \right) \bar{L} + C_D y(n_{\bar{y}}^-) \bar{L}^\alpha - \left( C_f + \frac{C_n}{n_{\bar{y}}^+} \right) \bar{L} - C_D y(n_{\bar{y}}^+) \bar{L}^\alpha \\ &= \left( C_f + \frac{C_n}{n_{\bar{y}}^-} \right) \bar{L} + C_D \bar{y} \bar{L}^\alpha - \left( C_f + \frac{C_n}{n_{\bar{y}}^+} \right) \bar{L} - C_D \bar{y} \bar{L}^\alpha \\ &= C_n \bar{L} \left( \frac{1}{n_{\bar{y}}^-} - \frac{1}{n_{\bar{y}}^+} \right) (> 0) \end{aligned}$$

Hence for any level of yield, the cost minimizing maximum orchard age is greater than or equal to the maximum yield age, i.e.  $n^* \geq n_{MY}$ . ■

**Lemma B.3** *The minimum of the isoquant is located at  $n_{MY}$ .*

**Proof of Lemma B.3.**

The isoquant is defined by  $y(n)L = \bar{Q}$ . This can be rewritten so that  $L$  is a function of  $n$ , i.e. for a particular level of feedstock production  $L = \frac{\bar{Q}}{y(n)}$ . The minimum of this function (i.e. the least quantity of land necessary to produce the desired quantity) occurs when the derivative of this function is set to zero.

$$\left. \frac{dL}{dn} \right|_{\text{isoquant}} = \frac{-\bar{Q}y'(n)}{[y(n)]^2} = 0 \Leftrightarrow y'(n) = 0$$

From the conditions imposed on the age-yield relationship in section 3.2 there is a unique maximum of the yield function located at  $n_{MY}$ . Hence the unique minimum of the isoquant function occurs at  $n_{MY}$ . ■

**Lemma B.4** *The minimum of the isocost curve is located at  $n < n_{MY}$ .*

**Proof of Lemma B.4.**

The isocost curve is defined by a level set of the cost function:  $C(n, L) = \bar{C}$ . I wish to locate

the set of local extrema of the isocost curve, where  $L$  is expressed as a function of  $n$ . This set is a subset of the critical points of  $\frac{dL}{dn}$ .

Totally differentiate the cost function:

$$\left[ \left( C_f + \frac{C_n}{n} \right) + \alpha C_D y(n) L^{\alpha-1} \right] dL + \left[ \frac{-C_n L}{n^2} + C_D y'(n) L^\alpha \right] dn = 0$$

Thus

$$\left. \frac{dL}{dn} \right|_{\text{isocost}} = \frac{C_n L/n^2 - C_D y'(n) L^\alpha}{\left( C_f + \frac{C_n}{n} \right) + \alpha C_D y(n) L^{\alpha-1}} = 0 \Leftrightarrow C_n L/n^2 = C_D y'(n) L^\alpha$$

since all the terms in the denominator are non-negative. The only term in this last equality that can change sign is  $y'(n)$ . All other terms are constrained to be non-negative. Hence the equality cannot be satisfied if  $y'(n) < 0$ , which occurs when  $n > n_{MY}$ . Also, if  $n = n_{MY}$  it must be that  $L = 0$  for the equality to be satisfied. If  $L = 0$  I have  $C(n_{MY}, 0) = 0$ , so for any positive level of cost  $(n_{MY}, 0)$  is not an element of the graph of the isocost function, and  $n_{MY}$  cannot be a critical point. Hence for any positive level of cost, any extrema of the isocost function must occur when  $n < n_{MY}$ . ■

**Lemma B.5** *The isocost curve has a positive slope for all  $n \geq n_{MY}$ .*

**Proof of Lemma B.5.**

This follows immediately from the proof of lemma B.4 since the expression for the slope of the isoquant curve is strictly positive for all  $n > n_{MY}$ . ■

**Proof of proposition 3.1.**

*The optimal  $n$  must be strictly greater than  $n_{MY}$ , i.e.  $n^* > n_{MY}$ .*

The isocost curve has a positive slope for all  $n \geq n_{MY}$  (lemma B.5). The isoquant curve has a zero slope at  $n_{MY}$  (lemma B.3). Hence the isocost and isoquant curves cannot be tangential at  $n_{MY}$ , so  $n^* \neq n_{MY}$ . Combining this with lemma B.2 gives me the result. ■

**Lemma B.6** *Assumptions (3.1)-(3.4) imply that  $0 < \lim_{n \rightarrow \infty} \int_0^n f(a) da < \infty$*

**Proof of Lemma B.6.**

I can split  $\lim_{n \rightarrow \infty} \int_0^n f(a) da$  in two by partitioning its domain:

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_0^n f(a) da &= \lim_{n \rightarrow \infty} \int_0^k f(a) da + \lim_{n \rightarrow \infty} \int_k^n f(a) da \\ &= \int_0^k f(a) da + \lim_{n \rightarrow \infty} \int_k^n f(a) da \end{aligned}$$

Now consider  $\int_0^k f(a) da$ . The age-yield relationship is bounded below by 0 by construction ( $f(a)$  represents a physical quantity). Assumptions (3.1)-(3.3) imply that  $f(a)$  is bounded



above. Hence  $f(a)$  is bounded on the domain  $[0, k]$  for all  $k \in \mathbb{R}_{>0}$ . Thus  $0 \leq \int_0^k f(a) da < \infty$  since this is the integral of a bounded positive function on a finite domain.

I must consider two possibilities when analyzing  $\lim_{n \rightarrow \infty} \int_k^n f(a) da$ : either  $f(a) > 0$  for all  $a \in \mathbb{R}_{\geq 0}$ , or there exists some  $\hat{k} \in \mathbb{R}_{\geq 0}$  such that for all  $a > \hat{k}$   $f(a) = 0$ . In the first case, I must establish that  $\lim_{n \rightarrow \infty} f(a)$  approaches zero fast enough that  $\lim_{n \rightarrow \infty} \int_k^n f(a) da$  is not infinite. Assumption (3.4) implies that there exist  $k \in \mathbb{R}_{\geq 0}$  and  $p > 1$  such that for all  $a > k$   $f(a) < \frac{1}{a^p}$  (if such  $k$  and  $p$  did not exist,  $\lim_{n \rightarrow \infty} \int_k^n f(a) da$  would either be strictly positive, or infinite). Thus

$$\lim_{n \rightarrow \infty} \int_k^n f(a) da < \lim_{n \rightarrow \infty} \int_k^n \frac{1}{a^p} da < \infty$$

since integrals of the form  $\int_k^\infty \frac{1}{x^p} dx$  are convergent if and only if  $p > 1$ . In the second case,  $\lim_{n \rightarrow \infty} \int_{\hat{k}}^n f(a) da = 0$ , and  $\lim_{n \rightarrow \infty} \int_0^n f(a) da = \int_0^{\hat{k}} f(a) da$ . Thus  $0 \leq \lim_{n \rightarrow \infty} \int_k^n f(a) da < \infty$

Assumption 3.3 implies that  $f(a)$  is strictly positive on some subset of  $\mathbb{R}_{\geq 0}$  with non-empty interior. Hence  $\lim_{n \rightarrow \infty} \int_0^n f(a) da > 0$

Therefore  $0 < \lim_{n \rightarrow \infty} \int_0^n f(a) da < \infty$  ■

### Proof of proposition 3.2.

Given assumptions (3.1)-(3.4), a solution,  $n^*$ , to the cost minimization problem exists such that  $n^* \in (n_{MY}, \infty)$  and  $\varepsilon_{y(n^*)} > -1$ , where  $\varepsilon_{y(n)}$  is the elasticity of the yield function with respect to age.

**Sketch of the proof:** I have already demonstrated that  $n^*$  must be greater than  $n_{MY}$ . At  $n_{MY}$  the slope of the isocost curve is strictly positive and the slope of the isoquant curve is zero. I show that as  $n$  approaches infinity, the slope of the isocost curve approaches zero, while the slope of the isoquant curve approaches a positive value. By continuity the slope functions must cross at least once, and hence there must exist at least one point where the isocost and isoquant curves are tangent to each other.

I begin by showing that the slope of the isocost curve approaches zero as  $n$  approaches infinity. The slope of the isocost function when  $L$  is written as a function of  $n$  (as derived in lemma B.3)

$$\left. \frac{dL}{dn} \right|_{\text{isocost}} = \frac{C_n L(n)/n^2 - C_D y'(n) L(n)^\alpha}{(C_f + \frac{C_n}{n}) + \alpha C_D y(n) L(n)^{\alpha-1}}$$

To take the limit of this expression as  $n$  approaches infinity, I need to know how  $L(n)$  on the isocost function behaves as  $n$  approaches infinity. The isocost function is defined as

$$C(n, L) = (C_f + \frac{C_n}{n})L + C_D y(n) L^\alpha = \bar{C}$$

This implicitly defines  $L$  as a function of  $n$ .

$$C(n) = (C_f + \frac{C_n}{n})L(n) + C_D y(n) L(n)^\alpha = \bar{C}$$

Now I take the limit of this expression as  $n \rightarrow \infty$  and solve for the unknown value  $L_\infty$ .

$$\begin{aligned} \lim_{n \rightarrow \infty} (C_f + \frac{C_n}{n})L(n) + C_D y(n) L(n)^\alpha &= \bar{C} \\ \Rightarrow C_f L_\infty &= \bar{C} \\ \Rightarrow L_\infty &= \frac{\bar{C}}{C_f} \quad \text{A constant} \end{aligned}$$

Returning to the derivative of the isocost function

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{dL}{dn} \Big|_{\text{isocost}} &= \lim_{n \rightarrow \infty} \frac{C_n L(n)/n^2 - C_D y'(n) L^\alpha}{(C_f + \frac{C_n}{n}) + \alpha C_D y(n) L(n)^{\alpha-1}} \\ &= \frac{0 - 0}{C_f + 0 + 0} \quad \text{Since } y(n) \text{ and } y'(n) \text{ both approach 0,} \\ &= 0 \quad \text{and } L(n) \text{ approaches a constant as } n \rightarrow \infty \end{aligned}$$

Now I show that under a certain condition the slope of the isoquant function approaches a positive constant as  $n \rightarrow \infty$ . The isoquant function is given by  $y(n) L = \bar{Q}$  and can be rewritten as

$$\begin{aligned} L &= \frac{\bar{Q}}{\frac{1}{n} \int_0^n f(a) da} \\ &= \frac{\bar{Q} n}{\int_0^n f(a) da} \end{aligned}$$

The slope of the isoquant function is given by

$$\frac{dL}{dn} \Big|_{\text{isoquant}} = \frac{\bar{Q} (\int_0^n f(a) da - n f(n))}{[\int_0^n f(a) da]^2}$$

The limit of the slope as  $n$  approaches infinity is

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{dL}{dn} \Big|_{\text{isoquant}} &= \lim_{n \rightarrow \infty} \frac{\bar{Q} \left( \int_0^n f(a) da - n f(n) \right)}{\left[ \int_0^n f(a) da \right]^2} \\
&= \bar{Q} \frac{\lim_{n \rightarrow \infty} \left( \int_0^n f(a) da - n f(n) \right)}{\lim_{n \rightarrow \infty} \left[ \int_0^n f(a) da \right]^2} && \text{since } \lim_{n \rightarrow \infty} \int_0^n f(a) da > 0 \\
&&& \text{(Lemma B.6)} \\
&= \bar{Q} \frac{\overbrace{\lim_{n \rightarrow \infty} \int_0^n f(a) da}^+ - \lim_{n \rightarrow \infty} n f(n)}{\underbrace{\lim_{n \rightarrow \infty} \left[ \int_0^n f(a) da \right]^2}_+} && \text{since} \\
&&& \lim_{n \rightarrow \infty} \int_0^n f(a) da \\
&&& > n f(n) = 0 \\
&&& \text{and} \\
&&& 0 < \lim_{n \rightarrow \infty} \int_0^n f(a) da < \infty \\
&&& \text{(Lemma B.6)} \\
\implies 0 < \lim_{n \rightarrow \infty} \frac{dL}{dn} \Big|_{\text{isoquant}} < \infty
\end{aligned}$$

Now define a function that returns the difference in the slopes of the isocost and isoquant functions,  $h(n) = \frac{dL}{dn} \Big|_{\text{isocost}} - \frac{dL}{dn} \Big|_{\text{isoquant}}$ . Since both constituent functions are continuous on the interval  $(n_{MY}, \infty)$ ,  $h(n)$  is also continuous on this interval. At the maximum yield age  $h(n_{MY}) > 0$  (from lemmas B.3 and B.5) and, as I have just shown, when  $n$  approaches infinity the limit of  $h(n)$  is strictly less than zero. Hence by the intermediate value theorem, there must exist some  $n^* \in (n_{MY}, \infty)$  such that  $h(n) = 0$ , and the isocost and isoquant curves are tangent to one another.

Additionally I prove the elasticity condition on the yield function at the optimum. Instead of treating the isocost and isoquant curves separately, I can consider the joint problem, where the quantity constraint is substituted into the cost function, written as a function of  $n$  only.

$$C(n, L(n)) = \left( C_f + \frac{C_n}{n} \right) \frac{\bar{Q}}{y(n)} + C_D y(n) \left( \frac{\bar{Q}}{y(n)} \right)^\alpha$$

The first order condition for a minimum with respect to  $n$  is

$$\frac{dC}{dn} = g(n) = \frac{-C_f y'(n)}{[y(n)]^2} - \frac{C_n (y(n) + n y'(n))}{[n y(n)]^2} + \frac{(1 - \alpha) C_D y'(n) \bar{Q}^{\alpha-1}}{[y(n)]^\alpha} = 0$$

Observe that the first and second terms in this expression are positive since the marginal yield must be negative at the optimum. If  $y(n) + n y'(n) \leq 0$ , then the second term is non-negative,

and the first order condition cannot be satisfied. Rewriting this condition I get

$$\begin{aligned} y(n) + n y'(n) &\leq 0 \\ n y'(n) &\leq -y(n) \\ \frac{n y'(n)}{y(n)} &\leq -1 \\ \varepsilon_{y(n)} &\leq -1 \end{aligned}$$

Hence at an optimum I must have the necessary condition that  $\varepsilon_{y(n)} > -1$ . ■

### Proof of proposition 3.3.

*The condition*

$n^3 \left( y'(n)^2 ((\alpha - 3) \alpha C_D L^{\alpha-1} + 2\lambda) - y(n) y''(n) (\lambda - C_D L^{\alpha-1}) \right) + 2 C_n (n y'(n) + y(n)) > 0$   
when evaluated at  $(n^*, L^*)$  is a sufficient condition for  $(n^*, L^*)$  to solve the cost-minimization problem.

The determinant of the bordered Hessian of the Lagrangian is given by

$$D(n, L) = \begin{vmatrix} 0 & \frac{\partial \mathcal{L}}{\partial n} & \frac{\partial \mathcal{L}}{\partial L} \\ \frac{\partial \mathcal{L}}{\partial n} & \frac{\partial^2 \mathcal{L}}{\partial n^2} & \frac{\partial^2 \mathcal{L}}{\partial n \partial L} \\ \frac{\partial \mathcal{L}}{\partial L} & \frac{\partial^2 \mathcal{L}}{\partial L \partial n} & \frac{\partial^2 \mathcal{L}}{\partial L^2} \end{vmatrix}$$

From Sydsæter et al. (2005) if  $D(n^*, L^*) < 0$  then  $(n^*, L^*)$  solves the local minimization problem. For the perennial crop field-biorefinery supply chain cost minimization problem, the determinant of the Hessian of the Lagrangian evaluates to

$$\begin{aligned} D(n, L) &= \begin{vmatrix} 0 & Ly'(n) & y(n) \\ Ly'(n) & \frac{2LC_n}{n^3} - y''(n) (\lambda L - C_D L^\alpha) & y'(n) (\alpha C_D L^{\alpha-1} - \lambda) - \frac{C_n}{n^2} \\ y(n) & y'(n) (\alpha C_D L^{\alpha-1} - \lambda) - \frac{C_n}{n^2} & (\alpha - 1) \alpha C_D L^{\alpha-2} y(n) \end{vmatrix} \\ &= - \frac{y(n)L}{n^3} \left( n^3 (y'(n))^2 ((\alpha - 3) \alpha C_D L^{\alpha-1} + 2\lambda) - y(n) y''(n) (\lambda - C_D L^{\alpha-1}) \right) \\ &\quad + 2 C_n (n y'(n) + y(n)) \end{aligned} \tag{B.1}$$

Since  $\frac{y(n)L}{n^3}$  is positive, the condition that  $D(n^*, L^*) < 0$  simplifies to

$$\begin{aligned} n^3 \left( y'(n)^2 ((\alpha - 3) \alpha C_D L^{\alpha-1} + 2\lambda) - y(n) y''(n) (\lambda - C_D L^{\alpha-1}) \right) \\ + 2 C_n (n y'(n) + y(n)) > 0 \end{aligned}$$

■

**Proof of proposition 3.4.**

As biorefinery size increases, the optimal orchard age decreases, i.e.  $\frac{dn^*}{d\bar{Q}} < 0$ .

Totally differentiating  $g(n, \bar{Q})$  (The derivative of the cost function when the constraint is used to eliminate  $L$  — derived in the proof of proposition 3.2) gives me an expression for the desired comparative static

$$\frac{dn^*}{d\bar{Q}} = \frac{-g_{\bar{Q}}}{g_n}$$

At an optimum the second order condition for a minimum must hold, so  $g_n$  must be positive. Hence

$$\text{sign} \left( \frac{dn^*}{d\bar{Q}} \right) = -\text{sign}(g_{\bar{Q}})$$

Differentiating  $g(n, \bar{Q})$  with respect to  $\bar{Q}$ , and evaluating at the optimum yields

$$g_{\bar{Q}} = -(1 - \alpha)^2 \underbrace{[y(n^*)]^{-\alpha}}_{+} \underbrace{y'(n^*)}_{-} \underbrace{\bar{Q}^{\alpha-2}}_{+} \quad (> 0)$$

Hence

$$\frac{dn^*}{d\bar{Q}} < 0$$

■

**Proof of proposition 3.5.**

The change in optimal growing region size with respect to a change in biorefinery capacity is generally ambiguous, but if  $y''(n^*) < 0$ , then increased biorefinery capacity leads to increase growing region size, i.e.  $\frac{dL^*}{d\bar{Q}} > 0$ .

To analyze this comparative static of the constrained cost minimization problem using the substitution method I need to define the inverse yield function,  $g(y) = n$  ( $g^{-1}(n) = y(n)$ ). Since the yield function is not surjective, I can only define and analyze the inverse yield on a subset of the domain. Fortunately, as shown by proposition 3.1, the optimal  $n$  is found in the subset  $n > n_{MY}$ . On this subset the yield function is bijective, and I am guaranteed the existence of  $g(y)$ .

Using the constraint on biorefinery capacity ( $y(n)L = \bar{Q} \Rightarrow y(n) = \frac{\bar{Q}}{L}$  and  $n = g(\frac{\bar{Q}}{L})$ ) I can rewrite the cost function as a function of growing region only.

$$C(n(L), L) = \left( C_f + \frac{C_n}{g\left(\frac{\bar{Q}}{L}\right)} \right) L + C_D \bar{Q} L^{\alpha-1}$$

The first order condition with respect to a minimum is

$$\frac{dC}{dL} = C_f + \frac{C_n}{g\left(\frac{\bar{Q}}{L}\right)} + \frac{\bar{Q} C_n g'\left(\frac{\bar{Q}}{L}\right)}{L \left[g\left(\frac{\bar{Q}}{L}\right)\right]^2} + (\alpha - 1)C_D \bar{Q} L^{\alpha-2} = 0$$

Cross multiply by  $L \left[g\left(\frac{\bar{Q}}{L}\right)\right]^2$

$$h(L) = C_f L \left[g\left(\frac{\bar{Q}}{L}\right)\right]^2 + C_n L g\left(\frac{\bar{Q}}{L}\right) + C_n \bar{Q} g'\left(\frac{\bar{Q}}{L}\right) + (\alpha - 1)C_D \bar{Q} \left[g\left(\frac{\bar{Q}}{L}\right)\right]^2 L^{\alpha-1} = 0$$

Totally differentiating  $h(n, \bar{Q})$  gives me an expression for the desired comparative static

$$\frac{dL^*}{d\bar{Q}} = \frac{-h_{\bar{Q}}}{h_L}$$

At an optimum the second order condition for a minimum must hold, so  $g_n$  must be positive. Hence

$$\text{sign}\left(\frac{dL^*}{d\bar{Q}}\right) = -\text{sign}(h_{\bar{Q}})$$

(B.2)

$$h_{\bar{Q}} = (\alpha - 1)C_D \left[g\left(\frac{\bar{Q}}{L}\right)\right]^2 L^{\alpha-1} \quad (>)$$

(B.3)

$$+ 2(\alpha - 1)C_D \bar{Q} g\left(\frac{\bar{Q}}{L}\right) g'\left(\frac{\bar{Q}}{L}\right) L^{\alpha-2} \quad (<)$$

(B.4)

$$+ 2C_f g\left(\frac{\bar{Q}}{L}\right) g'\left(\frac{\bar{Q}}{L}\right) \quad (<)$$

(B.5)

$$+ 2C_n g'\left(\frac{\bar{Q}}{L}\right) \quad (<)$$

(B.6)

$$+ \frac{C_n \bar{Q} g''\left(\frac{\bar{Q}}{L}\right)}{L} \quad (\text{Ambiguous})$$

If  $g''\left(\frac{\bar{Q}}{L}\right) < 0$  at  $L^*$ , the term B.6 in  $h_{\bar{Q}}$  is negative.

**Aside: Rewriting this condition in terms of  $y(n^*)$**

This condition on the second derivative of the inverse yield function is not particularly intuitive. I can rewrite this condition in terms of  $y(n^*)$  which makes it much easier to understand. To do this I must rewrite this condition on the second derivative of an inverse function in terms of the original function. The relationship between the second derivative of a function and its inverse is

$$(f^{-1})''(f(x)) = \frac{-f''(x)}{[f'(x)]^3}$$

For the inverse yield function this becomes

$$g''\left(\frac{\bar{Q}}{L}\right) = g''(y(n^*)) = \frac{-y''(n^*)}{[y'(n^*)]^3}$$

So

$$g''\left(\frac{\bar{Q}}{L}\right) < 0 \Leftrightarrow \frac{-y''(n^*)}{[y'(n^*)]^3} < 0$$

Since  $y'(n^*) < 0$  and the cubing operation preserves sign, this inequality is satisfied if and only if  $y''(n^*) < 0$ .

**Returning to the proof**

Given that  $y''(n^*) < 0$ , I now show that term (B.2) plus term (B.3) is negative.

$$(B.2) + (B.3) = (\alpha - 1)C_D \left[ g\left(\frac{\bar{Q}}{L}\right) \right]^2 L^{\alpha-1} + 2(\alpha - 1)C_D \bar{Q} g\left(\frac{\bar{Q}}{L}\right) g'\left(\frac{\bar{Q}}{L}\right) L^{\alpha-2}$$

Extract common factors

$$(B.2) + (B.3) = \underbrace{(\alpha - 1)C_D g\left(\frac{\bar{Q}}{L}\right) L^{\alpha-1}}_{>0} \left[ g\left(\frac{\bar{Q}}{L}\right) + 2\bar{Q} g'\left(\frac{\bar{Q}}{L}\right) L^{-1} \right]$$

Therefore

$$\text{sign}((B.2) + (B.3)) = \text{sign} \left( g\left(\frac{\bar{Q}}{L}\right) + 2\bar{Q} g'\left(\frac{\bar{Q}}{L}\right) L^{-1} \right)$$

Substitute the definition of  $\bar{Q} = y(n) L$

$$\begin{aligned} & g\left(\frac{y(n) L}{L}\right) + 2y(n) L g'\left(\frac{y(n) L}{L}\right) L^{-1} \\ &= g(y(n)) + 2y(n) g'(y(n)) \\ &= n + \frac{2y(n)}{y'(n)} \end{aligned} \quad \text{since } g(\cdot) \text{ is inverse of } y(\cdot)$$

Recall  $y'(n) = \frac{f(n)-y(n)}{n}$ , so

$$\begin{aligned} n + \frac{2y(n)}{y'(n)} &= n + \frac{2n y(n)}{f(n) - y(n)} \\ &= n \left( 1 + \frac{2y(n)}{f(n) - y(n)} \right) \\ &= n \left( 1 - \frac{2y(n)}{y(n) - f(n)} \right) \end{aligned}$$

For  $n > n_{MY}$ ,  $y(n) > f(n) \geq 0$ , hence  $\frac{2y(n)}{y(n)-f(n)} > 1$ , so (B.2) – (B.3)  $< 0$ ,  $h_{\bar{Q}} < 0$ , and  $\frac{dL^*}{d\bar{Q}} > 0$ . ■

### Proof of proposition 3.6.

See table on page 53

As explained in the proofs for propositions 3.4 and 3.5, the sign of the comparative static of  $n^*$  and  $L^*$  with respect to any exogenous variable  $x$  can be found by analyzing the sign of the relevant derivative of the first order condition, i.e.

$$\text{sign} \left( \frac{dn^*}{dx} \right) = -\text{sign}(g_x) \quad \text{and} \quad \text{sign} \left( \frac{dL^*}{dx} \right) = -\text{sign}(h_x)$$

I now present and sign the expressions of  $g_x$  for the parameters of interest.

$$g_{C_f} = -\frac{\overbrace{\bar{Q} y'(n^*)}}{y(n^*)^2} \quad (> 0)$$

$$g_{C_n^*} = \underbrace{\frac{-\bar{Q}}{n^* y(n^*)}}_{-} \underbrace{\left[ \frac{1}{n^*} + \frac{y'(n^*)}{y(n^*)} \right]}_{+} (< 0) \quad \text{Since } \varepsilon_{y(n^*)} > -1 \Rightarrow \frac{1}{n^*} + \frac{y'(n^*)}{y(n^*)} > 0 \quad (\text{Prop 3.2})$$

$$g_{C_D} = \underbrace{(1 - \alpha)}_{-} \underbrace{\bar{Q}^\alpha y(n^*)^{-\alpha}}_{-} \underbrace{y'(n^*)}_{-} \quad (> 0)$$

$$\begin{aligned} g_\alpha &= -C_D \bar{Q}^\alpha y(n^*)^{-\alpha} y'(n^*) ((\alpha - 1)(\ln(\bar{Q}) - \ln(y(n^*))) + 1) \\ &= -\underbrace{C_D \bar{Q}^\alpha y(n^*)^{-\alpha}}_{+} \underbrace{y'(n^*)}_{-} \underbrace{((\alpha - 1) \ln(L) + 1)}_{+} \quad (> 0) \end{aligned}$$

I now present and sign the expressions of  $h_x$  for the parameters of interest.

$$h_{C_f} = L(n^*)^2 \quad (> 0)$$



$$h_{C_n} = Ln^* + \frac{\bar{Q}}{y'(n^*)} \quad (< 0) \quad \text{Since } \varepsilon_{y(n^*)} > -1 \Rightarrow Ln^* + \frac{\bar{Q}}{y'(n^*)} < 0 \quad (\text{Prop 3.2})$$

$$h_{C_D} = (\alpha - 1)(n^*)^2 \bar{Q} L^{\alpha-1} \quad (> 0)$$

$$h_\alpha = (n^*)^2 \bar{Q} C_D L^{\alpha-1} (1 + (\alpha - 1) \ln(L)) \quad (> 0) \Leftrightarrow L > e^{-1/(\alpha-1)}$$

■

## A.2 Calibration of the Cost-Minimization Problem

The cost minimization problem has 7 parameters for which meaningful values must be found (assuming the choice of the Hoerl function for the age-yield relationship). From average yield function with a Hoerl age-yield relationship

$$y(n) = \frac{1}{n} \int_0^n ax^b e^{cx} dx$$

I have parameters  $a$ ,  $b$ , and  $c$ . From the simplified cost minimization problem I have  $C_f$ , the non-age-dependent per land unit growing costs,  $C_n$ , the age-dependent replanting costs,  $C_D$ , the average delivery cost per unit of feedstock, and  $\alpha$ , the shape parameter of the growing region.

Parameter	Value
a	34.13
b	3.74
c	-0.98
$C_f$	2259.67
$C_n$	1569.69
$C_D$	0.2649
$\alpha$	1.5

Table A.1: Parameter values used for cost minimization simulation

### Hoerl Parameters: $a$ , $b$ , $c$

To estimate parameters  $a$ ,  $b$ , and  $c$ , I fit the Hoerl function to age-yield data obtained from Margarido and Santos (2012). The econometric advantage of the Hoerl function is that its logarithm is linear in parameters (Haworth and Vincent, 1977).

Since having zero yield in the first year leads to a very bad fit of the Hoerl function, I adjusted the data from 0 yield in the first year, to a yield of 60 ton/ha in the first year, which is the average of the first and second year.

The fitted age-yield relationship is shown in figure 3.1. The parameter values obtained are  $a = 34.1342$ ,  $b = 3.74$ , and  $c = -0.98$ .

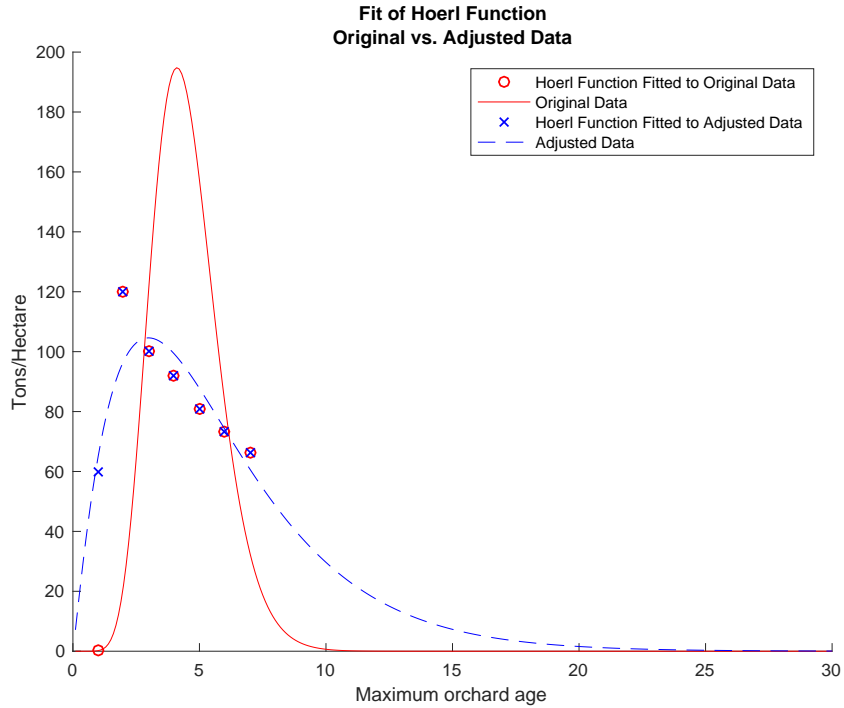


Figure A.1: The Hoerl function provides a poor fit for the observed sugarcane age-yield relationship. Fitting the Hoerl function to a modified dataset (increasing the observation for one-year-old canes) greatly improves the fit. I use the Hoerl function fitted to the adjusted data for the simulations in this study.

## Growing area shape parameter, $\alpha$

I assumed a circular growing area, which implies a value of  $\alpha = 1.5$ , as described above in the section on delivery costs.

## Cost parameters: $C_f$ , $C_n$ , $C_D$

I derived the feedstock cost parameters,  $C_f$  and  $C_n$ , from Teixeira (2013), and the delivery cost parameter,  $C_D$ , from Crago et al. (2010).

Teixeira (2013) presents an example operating budget for a 5-cut (6-age-class) sugarcane operation in São Paulo state.<sup>1</sup> Costs are divided into five categories, delivery costs, and four that account for farm gate feedstock costs: preparing the soil, planting, harvest, and maintenance of the ratoon. The total farm gate feedstock costs for a 6 hectare operation is

<sup>1</sup>Teixeira (2013) assumes that 80 percent of the cane is harvested burned, and 20 percent is harvested raw.

given by

$$\begin{aligned} \text{Total Farm Gate} \\ \text{Feedstock Costs} &= \text{Soil Preparation} + \text{Planting} \\ &+ 5 \times \text{Harvest} + 4 \times \text{Ratoon maintenance} \end{aligned}$$

Since the total cost is given for 5 hectares, the total cost per hectare is

$$\begin{aligned} \text{Total Farm Gate} \\ \text{Feedstock Costs} \\ \text{(Per Hectare)} &= \frac{1}{6} \times \text{Soil Preparation} + \frac{1}{6} \times \text{Planting} \\ &+ \frac{5}{6} \times \text{Harvest} + \frac{4}{6} \times \text{Ratoon maintenance} \end{aligned}$$

Assuming that these cost parameters are constant with respect to the number of age-classes I can write the total farm gate feedstock per hectare as a function of the age structure

$$\begin{aligned} \text{Farm gate} \\ \text{feedstock costs} (n) &= \frac{1}{n} \times \text{Soil Preparation} + \frac{1}{n} \times \text{Planting} \\ &+ \frac{n-1}{n} \times \text{Harvest} + \frac{n-2}{n} \times \text{Ratoon maintenance} \end{aligned}$$

Substituting Teixeira's numbers (in Reals) from the example budget, the cost function becomes

$$\text{Farm gate} \\ \text{feedstock costs} (n) = \frac{656.07}{n} + \frac{4159.83}{n} + \frac{n-1}{n} \times 1273.13 + \frac{n-2}{n} \times 986.54$$

Which on rearranging becomes

$$\text{Farm gate} \\ \text{feedstock costs} (n) = 2259.67 + \frac{1569.69}{n}$$

Hence for the simulations I use a baseline of  $C_f = 2259.67$  and  $C_n = 1569.69$ .

While Teixeira (2013) does include estimates of delivery costs, he does not include the biorefinery size that this example farm is feeding. I therefore turn to Crago et al. (2010) to derive the delivery cost parameter.

The total delivery cost from a growing region is given by

$$\begin{aligned} \text{Total} \\ \text{Delivery Costs} &= \text{Average Cost Per Ton Kilometer} \\ &\times \text{Quantity Transported} \\ &\times \text{Average Delivery Distance} \end{aligned}$$

Let  $\delta$  represent the average delivery cost per ton kilometer (i.e. the average cost to transport one ton of feedstock one kilometer). Crago et al. (2010) report an average transport cost of R\$6.7 to transport a ton of feedstock from the farm gate to the mill. The average delivery distance in this study was 22 kilometers so in this case  $\delta = 0.3045$ .

The average mill size in Crago et al. (2010) is 4.8 million tons. Given my assumption that the growing region produces the exact quantity required to feed the mill, this implied that the average quantity of feedstock transported was 4.8 million tons.

When calculating the average delivery distance, I must make a distinction between the area of land planted with sugarcane,  $L$ , and the area of the growing region,  $A$ . Although I am assuming that the growing region is circular, it is not necessarily the case that all the land is planted with sugarcane. In fact, relaxing the link between planted area and growing region area is necessary to correctly calibrate the model to the data in Crago et al. (2010).

Let  $d$  be the average density of sugarcane fields in the growing region, and  $A$  be the area of the growing region. Hence

$$L = d \times A$$

The average delivery distance is given by the expression

$$r_{av} = \frac{2}{3}r_{max} = \frac{2}{3}\sqrt{\frac{A}{\pi}}$$

Since the average delivery distance,  $r_{av}$ , from Crago et al. (2010) is 22km, the size of the growing region is  $A = 342119$  ha.

I calculate the density parameter from

$$\text{Total Quantity} = \text{Yield} \times \text{Density} \times \text{Growing Region Area}$$

Crago et al. (2010) reports an average yield of 75 tons per hectare. So I calculate the density as

$$4800000 = 75 \times d \times 342119 \Rightarrow d = 0.187$$

Hence the expression for the total delivery cost becomes

$$\begin{aligned} \text{Total Delivery Costs} &= \delta \times Q \times r_{av} \\ &= \delta \times Q \times \frac{2}{3}\sqrt{\frac{A}{\pi}} \\ &= \delta \times Q \times \frac{2}{3}\sqrt{\frac{L}{d \times \pi}} \\ &= \frac{2\delta}{3}\sqrt{\frac{1}{d \times \pi}} \times Q \times \sqrt{L} \\ &= \frac{2\delta}{3}\sqrt{\frac{1}{d \times \pi}} \times y(n)L\sqrt{L} \\ &= \frac{2\delta}{3}\sqrt{\frac{1}{d \times \pi}} \times y(n)L^{1.5} \\ &= C_D \times y(n)L^{1.5} \end{aligned}$$

For the  $d$  and  $\delta$  derived from Crago et al. (2010),  $C_D = 0.2649$ .

# Appendix C

## Appendix to Chapter 4

### A.1 Lemmas and Proofs

#### Lemma C.1

**Lemma C.1** This lemma enumerates the possible cases for the four Lagrange multipliers associated with the KKT conditions in period  $t$ . In each case I identify the free and pre-determined variables, present the equations determining the free variables, and present the conditions on the parameters required for the solutions to the determining equations to be consistent with the assumptions for each case.

The assumptions common across all cases are:

- $u'(x_{yt} + f_m x_{mt}) = b - a(x_{yt} + f_m x_{mt})$
- $a, b, p > 0$
- $f_m > 1$
- $0 < \beta < 1$
- $x_{y0} \geq 0$
- $V'(x_{m,T+1}) \geq 0$

**Case 1.1:**  $\lambda_t = 0, \lambda_{t+1} = 0, \eta_{yt} = 0, \eta_{mt} = 0$

Free Variables:  $x_{yt}, x_{mt}$

Pre-determined Variables:  $\lambda_t = 0, \lambda_{t+1} = 0, \eta_{yt} = 0, \eta_{mt} = 0$

Substituting the predetermined variables into equations (4.1) and (4.2), I get

$$\begin{aligned}x_{yt} + f_m x_{mt} &= \frac{b - p}{a} \\x_{yt} + f_m x_{mt} &= \frac{b - p}{f_m a}\end{aligned}$$

There is no solution to these two equations for  $f_m \neq 1$ . Hence this case is infeasible.

**Case 1.2:**  $\lambda_t = 0, \lambda_{t+1} = 0, \eta_{yt} = 0, \eta_{mt} > 0$

Free Variables:  $x_{yt}, \eta_{mt}$

Pre-determined Variables:  $\lambda_t = 0, \lambda_{t+1} = 0, \eta_{yt} = 0, x_{mt} = 0$

Substituting the predetermined variables into equations (4.1) and (4.2), I get

$$\begin{aligned}x_{yt} &= \frac{b - p}{a} \\ \eta_{mt} &= (f_m - 1)p\beta^{t-1}\end{aligned}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if  $p \leq b$ .

**Case 1.3:**  $\lambda_t = 0, \lambda_{t+1} = 0, \eta_{yt} > 0, \eta_{mt} = 0$

Free Variables:  $x_{mt}, \eta_{yt}$

Pre-determined Variables:  $\lambda_t = 0, \lambda_{t+1} = 0, \eta_{mt} = 0, x_{yt} = 0$

Substituting the predetermined variables into equations (4.1) and (4.2) and solving for the free variables, I get

$$\begin{aligned}x_{mt} &= \frac{bf_m - p}{af_m^2} \\ \eta_{yt} &= \left(\frac{1}{f_m} - 1\right)p\beta^{t-1}\end{aligned}$$

This never consistent with the baseline assumptions and the assumptions on the multipliers for this case for any allowable parameter values since  $1/f_m - 1 < 0$ . Hence this case is infeasible.

**Case 1.4:**  $\lambda_t = 0, \lambda_{t+1} = 0, \eta_{yt} > 0, \eta_{mt} > 0$

Free Variables:  $\eta_{yt}, \eta_{mt}$

Pre-determined Variables:  $\lambda_t = 0, \lambda_{t+1} = 0, x_{yt} = 0, x_{mt} = 0$

Substituting the predetermined variables into equations (4.1) and (4.2) and solving for the free variables, I get

$$\begin{aligned}\eta_{yt} &= (b - p)\beta^{t-1} \\ \eta_{mt} &= (bf_m - p)\beta^{t-1}\end{aligned}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if  $p \leq b$ .

**Case 2.1:**  $\lambda_t = 0, \lambda_{t+1} > 0, \eta_{yt} = 0, \eta_{mt} = 0$

Free Variables:  $x_{mt}, \lambda_{t+1}$

Pre-determined Variables:  $\lambda_t = 0, \eta_{yt} = 0, \eta_{mt} = 0, x_{yt} = x_{m,t+1}$

Substituting the predetermined variables into equations (4.1) and (4.2) and solving for the free variables, I get

$$\begin{aligned}x_{mt} &= \frac{bf_m - p - af_mx_{m,t+1}}{af_m^2} \\ \lambda_{t+1} &= \frac{(f_m - 1)p\beta^{t-1}}{f_m}\end{aligned}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if

- $p \leq bf_m$  and  $x_{m,t+1} \leq \frac{bf_m - p}{af_m}$ , or
- $p = bf_m$  and  $x_{m,t+1} = 0$

**Case 2.2:**  $\lambda_t = 0, \lambda_{t+1} > 0, \eta_{yt} = 0, \eta_{mt} > 0$

Free Variables:  $\lambda_{t+1}, \eta_{mt}$

Pre-determined Variables:  $\lambda_t = 0, \eta_{yt} = 0, x_{yt} = x_{m,t+1}, x_{mt} = 0$

Substituting the predetermined variables into equations (4.1) and (4.2) and solving for the free variables, I get

$$\begin{aligned}\lambda_{t+1} &= (p + ax_{m,t+1} - b)\beta^{t-1} \\ \eta_{mt} &= (bf_m - p - af_mx_{m,t+1})\beta^{t-1}\end{aligned}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if

- $p \leq b$  and  $\frac{b-p}{a} < x_{m,t+1} \leq \frac{bf_m - p}{af_m}$ , or
- $b < p < bf_m$  and  $x_{m,t+1} < \frac{bf_m - p}{af_m}$

**Case 2.3:**  $\lambda_t = 0$ ,  $\lambda_{t+1} > 0$ ,  $\eta_{yt} > 0$ ,  $\eta_{mt} = 0$

Free Variables:  $x_{mt}$ ,  $\lambda_{t+1}$ ,  $\eta_{mt}$

Pre-determined Variables:  $\lambda_t = 0$ ,  $\eta_{yt} = 0$ ,  $x_{yt} = x_{m,t+1} = 0$

This case is only feasible when  $x_{m,t+1} = 0$ , in which case there is a redundant constraint.  $\eta_{yt} > 0$  implies  $x_{yt} = 0$ , hence if  $x_{m,t+1} \neq 0$  the assumptions are violated, and this case is inconsistent. If  $x_{m,t+1} = 0$ , then equations (4.1) and (4.2) are underdetermined, a continuum of possibilities for  $\eta_{yt}$  and  $\lambda_{t+1}$ . Assuming  $x_{m,t+1} = 0$  and solving the equations yields

$$x_{mt} = \frac{bf_m - p}{af_m^2}$$

$$\eta_{yt} = \left( \frac{1}{f_m} - 1 \right) p\beta^{t-1} + \lambda_{t+1}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if  $x_{m,t+1} = 0$  and

- $p \leq bf_m$  and  $1 < f_m < \frac{b\beta^{t-1} + \lambda_{t+1}}{b\beta^{t-1}}$ , or
- $p < bf_m$  and  $f_m = \frac{b\beta^{t-1} + \lambda_{t+1}}{b\beta^{t-1}}$ , or
- $p < \frac{f_m\lambda_{t+1}}{(f_m-1)\beta^{t-1}}$  and  $f_m > \frac{b\beta^{t-1} + \lambda_{t+1}}{b\beta^{t-1}}$

**Case 2.4:**  $\lambda_t = 0$ ,  $\lambda_{t+1} > 0$ ,  $\eta_{yt} > 0$ ,  $\eta_{mt} > 0$

Free Variables:  $\lambda_{t+1}$ ,  $\eta_{yt}$ ,  $\eta_{mt}$

Pre-determined Variables:  $\lambda_t = 0$ ,  $x_{yt} = x_{m,t+1} = 0$ ,  $x_{mt} = 0$

Like case 2.3, this case is only feasible when  $x_{m,t+1} = 0$ , in which case there is a redundant constraint.  $\eta_{yt} > 0$  implies  $x_{yt} = 0$ , hence if  $x_{m,t+1} \neq 0$  the assumptions are violated, and this case is inconsistent. If  $x_{m,t+1} = 0$ , then equations (4.1) and (4.2) are underdetermined, a continuum of possibilities for  $\eta_{yt}$  and  $\lambda_{t+1}$ . Assuming  $x_{m,t+1} = 0$  and solving the equations yields

$$\eta_{yt} = (b - p)\beta^{t-1} + \lambda_{t+1}$$

$$\eta_{mt} = (bf_m - p)\beta^{t-1}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if  $x_{m,t+1} = 0$  and

- $p < b$ , or
- $p < bf_m$  and  $f_m < \frac{b\beta^{t-1} + \lambda_{t+1}}{b\beta^{t-1}}$ , or
- $p < \frac{b\beta^{t-1} + \lambda_{t+1}}{b\beta^{t-1}}$  and  $f_m \geq \frac{b\beta^{t-1} + \lambda_{t+1}}{b\beta^{t-1}}$



**Case 3.1:**  $\lambda_t > 0, \lambda_{t+1} = 0, \eta_{yt} = 0, \eta_{mt} = 0$

Free Variables:  $x_{yt}, \lambda_t$

Pre-determined Variables:  $\lambda_{t+1} = 0, \eta_{yt} = 0, \eta_{mt} = 0, x_{mt} = x_{y,t-1}$

Substituting the predetermined variables into equations (4.1) and (4.2) and solving for the free variables, I get

$$x_{yt} = \frac{b - p - af_m x_{y,t-1}}{a}$$

$$\lambda_t = (f_m - 1)p\beta^{t-1}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if

- $p \leq b$  and  $x_{y,t-1} \leq \frac{b-p}{af_m}$ , or
- $p = bf_m$  and  $x_{y,t-1} = 0$

**Case 3.2:**  $\lambda_t > 0, \lambda_{t+1} = 0, \eta_{yt} = 0, \eta_{mt} > 0$

Free Variables:  $x_{yt}, \lambda_t, \eta_{mt}$

Pre-determined Variables:  $\lambda_{t+1} = 0, \eta_{yt} = 0, \eta_{mt} = 0, x_{mt} = x_{y,t-1} = 0$

This case is always inconsistent. The aging constraint implies  $x_{mt} \leq x_{y,t-1}$  and the non-negativity constraint implies that  $x_{mt} \geq 0$ . Assuming both constraints bind, as in this case, the direction in which the objective function improves is opposite for each of these constraints. Hence if one constraint binds, the other must not bind. Therefore, positive values for both  $\lambda_t$  and  $\eta_{mt}$  are never possible.

**Case 3.3:**  $\lambda_t > 0, \lambda_{t+1} = 0, \eta_{yt} > 0, \eta_{mt} = 0$

Free Variables:  $\lambda_t, \eta_{yt}$

Pre-determined Variables:  $\lambda_{t+1} = 0, \eta_{mt} = 0, x_{mt} = x_{y,t-1}, x_{yt} = 0$

Substituting the predetermined variables into equations (4.1) and (4.2) and solving for the free variables, I get

$$\lambda_t = (bf_m - p - af_m^2 x_{y,t-1})\beta^{t-1}$$

$$\eta_{yt} = (b - p - af_m x_{y,t-1})\beta^{t-1}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if

- $p \leq b$  and  $x_{y,t-1} \leq \frac{b-p}{af_m}$ , or
- $p = bf_m$  and  $x_{y,t-1} = 0$

**Case 3.4:**  $\lambda_t > 0, \lambda_{t+1} = 0, \eta_{yt} > 0, \eta_{mt} > 0$

Free Variables:  $\lambda_t, \eta_{yt}, \eta_{mt}$

Pre-determined Variables:  $\lambda_{t+1} = 0, x_{yt} = 0, x_{mt} = x_{y,t-1} = 0$

This case is always inconsistent, for the same reason as case 3.2. The aging constraint implies  $x_{mt} \leq x_{y,t-1}$  and the non-negativity constraint implies that  $x_{mt} \geq 0$ . Assuming both constraints bind, as in this case, the direction in which the objective function improves is opposite for each of these constraints. Hence if one constraint binds, the other must not bind. Therefore, positive values for both  $\lambda_t$  and  $\eta_{mt}$  are never possible.

**Case 4.1:**  $\lambda_t > 0, \lambda_{t+1} > 0, \eta_{yt} = 0, \eta_{mt} = 0$

Free Variables:  $\lambda_t, \lambda_{t+1}$

Pre-determined Variables:  $\eta_{yt} = 0, \eta_{mt} = 0, x_{yt} = x_{m,t+1}, x_{mt} = x_{y,t-1}$

Substituting the predetermined variables into equations (4.1) and (4.2) and solving for the free variables, I get

$$\begin{aligned}\lambda_t &= (bf_m - p - af_m(x_{m,t+1} + f_mx_{y,t-1}))\beta^{t-1} \\ \lambda_{t+1} &= (p + a(x_{m,t+1} + f_mx_{y,t-1}) - b)\beta^{t-1}\end{aligned}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if

- $p \leq b$  and  $(x_{y,t-1} \leq \frac{b-p}{af_m}$  and  $\frac{b-p-af_mx_{y,t-1}}{a} < x_{m,t+1} < \frac{bf_m-p-af_m^2x_{y,t-1}}{af_m})$  or  $(\frac{b-p}{af_m} < x_{y,t-1} < \frac{bf_m-p}{af_m^2}$  and  $x_{m,t+1} < \frac{bf_m-p-af_m^2x_{y,t-1}}{af_m})$ , or
- $p = b$  and  $(x_{y,t-1} = 0$  and  $0 < x_{m,t+1} < \frac{bf_m-p}{af_m})$  or  $(x_{y,t-1} < \frac{bf_m-p}{af_m^2}$  and  $x_{m,t+1} < \frac{bf_m-p-af_m^2x_{y,t-1}}{af_m})$ , or
- $b < p < bf_m$  and  $x_{y,t-1} < \frac{bf_m-p}{af_m^2}$  and  $x_{m,t+1} < \frac{bf_m-p-af_m^2x_{y,t-1}}{af_m}$ .

**Case 4.2:**  $\lambda_t > 0, \lambda_{t+1} > 0, \eta_{yt} = 0, \eta_{mt} > 0$

Free Variables:  $\lambda_t, \lambda_{t+1}, \eta_{mt}$

Pre-determined Variables:  $\eta_{yt} = 0, x_{yt} = x_{m,t+1}, x_{mt} = x_{y,t-1} = 0$

As for cases 3.2 and 3.4, this case is always inconsistent. The aging constraint implies  $x_{mt} \leq x_{y,t-1}$  and the non-negativity constraint implies that  $x_{mt} \geq 0$ . Assuming both constraints bind, as in this case, the direction in which the objective function improves is opposite for each of these constraints. Hence if one constraint binds, the other must not bind. Therefore, positive values for both  $\lambda_t$  and  $\eta_{mt}$  are never possible.

**Case 4.3:**  $\lambda_t > 0$ ,  $\lambda_{t+1} > 0$ ,  $\eta_{yt} > 0$ ,  $\eta_{mt} = 0$

Free Variables:  $\lambda_t$ ,  $\lambda_{t+1}$ ,  $\eta_{yt}$

Pre-determined Variables:  $\eta_{mt} = 0$ ,  $x_{yt} = x_{m,t+1} = 0$ ,  $x_{mt} = x_{y,t-1}$

Substituting the predetermined variables into equations (4.1) and (4.2) and solving for the free variables, I get

$$\begin{aligned}\lambda_t &= (bf_m - p - af_m(x_{m,t+1} + f_m x_{y,t-1}))\beta^{t-1} \\ \lambda_{t+1} &= (p + a(x_{m,t+1} + f_m x_{y,t-1}) - b)\beta^{t-1}\end{aligned}$$

This is consistent with the baseline assumptions and the assumptions on the multipliers for this case if

- $1 < f_m \leq \frac{b\beta^{t-1} + \lambda_{t+1}}{b\beta^{t-1}}$  and  $p \leq bf_m$  and  $x_{y,t-1} \leq \frac{bf_m - p}{af_m^2}$  and  $x_{m,t+1} < \frac{bf_m - p - af_m^2 x_{y,t-1}}{af_m}$ , or
- $f_m > \frac{b\beta^{t-1} + \lambda_{t+1}}{b\beta^{t-1}}$  and
  - $p \leq \frac{f_m \lambda_{t+1}}{(f_m - 1)\beta^{t-1}}$  and  $x_{y,t-1} < \frac{bf_m - p}{af_m^2}$  and  $x_{m,t+1} < \frac{bf_m - p - af_m^2 x_{y,t-1}}{af_m}$ , or
  - $\frac{f_m \lambda_{t+1}}{(f_m - 1)\beta^{t-1}} < p < \frac{b\beta^{t-1} + \lambda_{t+1}}{\beta^{t-1}}$  and  $x_{y,t-1} < \frac{(b-p)\beta^{t-1} + \lambda_{t+1}}{af_m \beta^{t-1}}$  and  $x_{m,t+1} < \frac{(b-p - af_m x_{y,t-1})\beta^{t-1} + \lambda_{t+1}}{a\beta^{t-1}}$

**Case 4.4:**  $\lambda_t > 0$ ,  $\lambda_{t+1} > 0$ ,  $\eta_{yt} > 0$ ,  $\eta_{mt} > 0$

Free Variables:  $\lambda_t$ ,  $\lambda_{t+1}$ ,  $\eta_{yt}$ ,  $\eta_{mt}$

Pre-determined Variables:  $x_{yt} = x_{m,t+1} = 0$ ,  $x_{mt} = x_{y,t-1} = 0$

As for cases 3.2, 3.4, and 4.2, this case is always inconsistent. The aging constraint implies  $x_{mt} \leq x_{y,t-1}$  and the non-negativity constraint implies that  $x_{mt} \geq 0$ . Assuming both constraints bind, as in this case, the direction in which the objective function improves is opposite for each of these constraints. Hence if one constraint binds, the other must not bind. Therefore, positive values for both  $\lambda_t$  and  $\eta_{mt}$  are never possible.

## Proof of proposition 4.2

**Proof of proposition 4.2.** Planting young trees in the final period ( $x_{yT} > 0$ ) implies that  $\eta_{yT} = 0$ . There are 5 cases where this is possible. In these cases  $x_{yT}$  is either a free variable (non-binding aging constraint), or it is determined by the aging constraint and the terminal condition. The two cases with  $x_{yT}$  as a free variable are 1.2 and 3.1. The three cases where  $x_{yT}$  is determined by the terminal condition are 2.1, 2.2, and 4.1.

If the aging constraint between periods  $T$  and  $T + 1$  does not bind ( $\lambda_{T+1} = 0$ ) either case 1.2 or 3.1 pertains. In case 1.2 the multipliers are  $\lambda_t = 0$ ,  $\lambda_{t+1} = 0$ ,  $\eta_{yt} = 0$ ,  $\eta_{mt} > 0$ . By Lemma C.1 the necessary condition for this case is  $p \leq b$ . In case 3.1 the multipliers are

$\lambda_t > 0$ ,  $\lambda_{t+1} = 0$ ,  $\eta_{yt} = 0$ ,  $\eta_{mt} = 0$ . By Lemma C.1 the necessary conditions for this case are  $p < b$  and  $x_{y,t-1} \leq \frac{b-p}{af_m}$ .

If the aging constraint between periods  $T$  and  $T + 1$  binds ( $\lambda_{T+1} > 0$ ) there are three possible cases, 2.1, 2.2, and 4.1. In case 2.1 the multipliers are  $\lambda_t = 0$ ,  $\lambda_{t+1} > 0$ ,  $\eta_{yt} = 0$ ,  $\eta_{mt} = 0$ . By Lemma C.1 the necessary conditions for this case are ( $p \leq bf_m$  and  $x_{m,T+1} \leq \frac{bf_m-p}{af_m}$ ) or ( $p = bf_m$  and  $x_{m,T+1} = 0$ ). This second condition is not consistent with the assumption that  $x_{yT} > 0$  and may be discarded. In case 2.2 the multipliers are  $\lambda_t = 0$ ,  $\lambda_{t+1} > 0$ ,  $\eta_{yt} = 0$ ,  $\eta_{mt} > 0$ . By Lemma C.1 the necessary conditions for this case are ( $p \leq b$  and  $\frac{b-p}{a} < x_{m,T+1} \leq \frac{bf_m-p}{af_m}$ ) or ( $b < p < bf_m$  and  $x_{m,T+1} < \frac{bf_m-p}{af_m}$ ). In case 4.1 the multipliers are  $\lambda_t > 0$ ,  $\lambda_{t+1} > 0$ ,  $\eta_{yt} = 0$ ,  $\eta_{mt} = 0$ . By Lemma C.1 there are several possible necessary conditions for this case. In each possibility  $x_{m,T+1} < \frac{bf_m-p-af_m^2x_{y,t-1}}{af_m}$  which implies  $x_{m,T+1} \leq \frac{bf_m-p}{af_m}$ .

Hence assuming the aging constraint binds between periods  $T$  and  $T + 1$ ,  $x_{m,T+1} \leq \frac{bf_m-p}{af_m}$  is a necessary condition for  $x_{yT} > 0$ , ■