Analysis of a Market for Tradable Credits, Policy Uncertainty Effects on Investment Decisions, and the Potential to Supply a Renewable Aviation Fuel Industry with an Experimental Industrial Oilseed

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Evan Lawrence Markel
May 2017

Copyright © 2017 by Evan L. Markel.
All rights reserved.

Dedication

To friends, family and loved ones who have supported me on this endeavor.

Acknowledgements

I would like to acknowledge my major advisor Burt English who has exercised great patience and understanding as he guided my doctoral education. I would also like to thank my doctoral committee members Dayton Lambert, Charles Sims, Hamparsum Bozdogan and Donald Hodges. The entire faculty and staff of the Agricultural and Resource Economics Department has contributed to my education and development. Jamey Menard has provided countless hours of mentoring and advice. Brad Wilson has provided support and tutoring in geospatial information systems and Chad Hellwinckel has provided support and tutoring in the Policy Analysis System. I would also like to acknowledge my fellow graduate students. Without their friendship and shoulders to lean on, this journey would have been insurmountable.

Abstract

This research is aligned with identifying barriers throughout the alternative jet-fuel supply chain. Prices are analyzed in the market for tradable credits known as renewable identification numbers (RINs). The RIN market is a key policy instrument used in the implementation of the renewable fuel standard (RFS). The program is highly complex and drivers of RIN price are not always clear. RIN prices also exhibit multiple regimes where the price of nested RINs converge. Therefore, a smooth transition autoregressive model is employed to examine drivers of RIN price and to identify drivers of price regime change. Through research in the RIN market and renewable fuel standard, a common theme of policy uncertainty is identified in the literature.

A two-variable real option model is utilized to examine the effect of policy uncertainty on the decision to invest in new production of second-generation biofuel. This represents the first attempt to isolate general market uncertainty from policy uncertainty in the biofuel producer's optimal investment decision. RFS policy uncertainty adds to the aggregate uncertainty faced by the biofuel producer and may be impeding the original intentions of the policy program.

Finally, an experimental biofuel feedstock and its potential to supply an alternative jet-fuel industry is considered. The experimental feedstock is known as pennycress, which produces an industrial oilseed. Using a partial equilibrium model of the agricultural sector, supply curves are simulated and its impacts on the agricultural sector are investigated.

Table of Contents

Introduction 1
Chapter I: Thresholds and Regime Change in the Market for Renewable Identification Numbers 8
Chapter I: Abstract 9
1.1. Introduction 10
1.2. Renewable Identification Numbers 13
1.2.1. RIN Pricing 15
1.3. Data 18
1.4. Nonlinearity in RIN price 20
1.4.1 Visually inspecting for nonlinearity 20
1.4.2 Testing for Nonlinearity 21
1.5. Thresholds in the RIN Market 23
1.5.1 Model Selection for D6/D4 Price Ratio 24
1.5.2 Estimation of SETAR and TAR models 26
1.6. Smooth Transition Autoregressive Model 29
1.6.1 Linearity testing against the LSTAR model 30
1.6.2 Testing the significance of the transition variable 30
1.6.3 Estimation of the STAR Model 31
1.6.4 STAR Model Diagnostics 33
1.6.5 An application to individual RIN prices 36
1.7. Conclusion 41
1. A. Appendix 43
1. A.1. Tables 43
1. A.2 Figures 50
Chapter II: Renewable Fuel Standard Uncertainty and Tax Incentive Effects in the Optimal Investment Decisions of Second-Generation Biofuel Producers 58
Chapter II: Abstract 59
2.1. Introduction 60
2.2. The Fundamentals of Biodiesel RIN Price 63
2.3. A Model of Optimal Entry Into and Exit From the Biofuel Market 67
2.3.1 Market Dynamics 68
2.3.2 The Firm's Decision to Enter 70
2.3.3. The Firm's Decision to Exit 71

2.4. An Application to a Second-generation Biofuel Investment 72
2.4.1 Estimating Stochastic Price Processes 73
2.4.2 Cost Parameters 77
2.4.3 Numerical Methods 79
2.5. Results 82
2.5.1 Results of the Single Stochastic Variable Case 84
2.5.2 Results of the Two-Stochastic Variable Case 84
2.5.3 Tax Incentives 88
2.6. Conclusion 92
2. A. Appendix 94
2. A.1 Tables 94
2. A.2 Figures 97
Chapter III: Potential for Pennycress to Support an Alternative Jet-Fuel Industry 105
Chapter III: Abstract 106
3.1. Introduction 107
3.2. Farm-Gate Costs 109
3.3. A Model of Optimal Decisions in the Production of Pennycress 111
3.3.1 Market Dynamics for Pennycress 113
3.3.2 The Farmer's Decision to Produce Pennycress 114
3.3.3 The Farmer's Decision to Exit 115
3.3.4 Parameterization of the model 116
3.3.5 Cost Parameters 118
3.3.6 The Price of Pennycress Which Triggers Entry and Exit 122
3.4. Geographic Suitability 123
3.5. Land Allocation and Supply Simulation 124
3.5.1 Results of Land Allocation and Supply Simulation 128
3.5.2 Regions of Pennycress Production 131
3.6. Plant-Gate Costs 132
3.7. Conclusion 135
3. A. Appendix 138
3. A.1 Tables 138
3. A.2 Figures 147

Conclusion 154
References 155
Vita 166

List of Tables

Table 1: RIN type by D-code and production pathway 43
Table 2: Optimal lag lengths of price ratio and individual prices 43
Table 3: Test of serial correlation in D6/D4 RIN price 43
Table 4: Test of heteroskedasticity in D6/D4 RIN price ratio 44
Table 5: Tsay and Keenan tests for nonlinearity in D6/D4 price ratio 44
Table 6: Linear model of D6/D4 RIN price ratio 44
Table 7: Test for most significant transition variable of D6/D4 price ratio 45
Table 8: D6/D4 Price ratio STAR estimation results 45
Table 9: Linear, TAR and STAR model comparison for D6/D4 RIN price ratio 46
Table 10: Test results of model diagnostics for RIN price ratio and individual RIN price 46
Table 11: Linear model of D4 RIN price 46
Table 12: Linear model of D6 RIN price 47
Table 13: Tsay and Keenan tests for nonlinearity of Individual RIN prices 47
Table 14: Tests of most significant transition variable for D4 RIN price 47
Table 15: Tests of most significant transition variable for D6 RIN price 48
Table 16: D4 STAR estimation results 48
Table 17: D4 linear and STAR model comparison 49
Table 18: D6 STAR estimation results 49
Table 19: D6 linear and STAR model comparison 49
Table 20: OLS results of D4 RIN price 94
Table 21: OLS results of conventional fuel price 94
Table 22: Parameter estimation results 94
Table 23: Cost parameters sourced from prior literature 95
Table 24: Comparison of entry and exit prices 95
Table 25: Comparison of production costs with tax incentives 95
Table 26: Comparison of entry and exit prices with current policy incentives 96
Table 27: Stochastic yield and price parameter estimations 138
Table 28: Best of fit distribution for historical soybean yield 138
Table 29: Probability of occurrence for yield reductions and expected cost 138
Table 30: Summary of pennycress cost of production parameters 139
Table 31: Entry and exit trigger prices for pennycress production 139
Table 32: Simulated pennycress production (figures in millions) 139
Table 33: Summary of plant-gate costs in the conversion of pennycress oil to green-jet 140
Table 34: Summary of plant-gate costs in the conversion of pennycress oil to green-diesel 141
Table 35: Summary of plant-gate costs in the conversion of pennycress oil to green-propane . 142
Table 36: Summary of plant-gate costs in the conversion of pennycress oil to green-naphtha.. 143
Table 37: Pennycress enterprise budget assuming 1600lbs per acre 144
Table 38: Changes from base case to alternate scenario with pennycress price of $0.15 per pound 145
Table 39: Changes from base case to alternate scenario with pennycress price of $0.50 per pound 146

List of Figures

Figure 1: Core value of a RIN in ethanol (a) and biodiesel markets (b) 50
Figure 2: Nested structure of RINs (U.S. Environmental Protection Agency 2016) 50
Figure 3: Historical RIN prices 51
Figure 4: Coupled RIN prices 51
Figure 5: Decoupled RIN prices 52
Figure 6: First-differenced Ratio of D6/D4 lag plots 53
Figure 7: Historical Ratio of D6/D4 RIN prices 54
Figure 8: Kernel density plot to illustrate bi-modal distribution 54
Figure 9: First-differenced ratio D6/D4 RIN prices 54
Figure 10: Evaluation of the transition function for D6/D4 RIN price ratio 55
Figure 11: Fitted values compared to actual values of $\Delta(D6/D4)$ for the STAR model and the linear model 55
Figure 12: Evaluation of D4 Transition Function 55
Figure 13: Fitted values compared to actual values of $\Delta ln(D4)$ price for the STAR model and the linear model 56
Figure 14: Evaluation of D6 Transition Function 56
Figure 15: Fitted values compared to actual values of $\Delta ln(D6)$ price for the STAR model and the linear model 57
Figure 16: Core Value of the RIN 97
Figure 17: Outward demand shift causes decrease in RIN value 97
Figure 18: Effect of a blender tax credit on the core value of a RIN 98
Figure 19: Effect of a shift in the RFS mandate on the core value of a RIN 98
Figure 20: Historical biodiesel RIN prices and biodiesel blend margin 99
Figure 21: ACF and PACF of $\Delta lnRt$ 99
Figure 22: ACF and PACF of $\Delta lnPt$ 100
Figure 23: Historical Diesel Price 100
Figure 24: Two-variable entry and exit threshold curves 101
Figure 25: Two-variable entry and exit threshold curves with fuel price and RIN price variance set to zero 101
Figure 26: Threshold curve sensitivity holding conventional fuel price volatility constant 102
Figure 27: Threshold curve sensitivity holding RIN price volatility constant 102
Figure 28: Effect of changes in the firms per gallon capital cost on entry and exit thresholds .. 103
Figure 29: Effect of changes in the firm's operating cost on entry and exit thresholds 103
Figure 30: Entry and exit threshold curves while accounting for producer policy incentives ... 104
Figure 31: Distribution of Illinois soybean yields 147
Figure 32: Observed data overlaid by best of fit theoretical distribution 147
Figure 33: Entry and Exit Thresholds for all combinations of price and yield 148
Figure 34: Supply curves under the assumption of 1600 pounds per acre yield 148
Figure 35: Pennycress disruption effects on commodity production and price 149
Figure 36: Pennycress disruption effects on commodity net returns 149
Figure 37: Simulated top pennycress producing county 150
Figure 38: Regions that pose a risk to pennycress based on historical climate data 151

Figure 39: Regions and their expected yields based on climate risk regions 152
Figure 40: Simulated pennycress production in year 2039..... 153

Introduction

Research studying the use of agriculture to produce energy began to arise in the early 1970's. The impetus was through discoveries in technological processes while researching novel ways of converting coal to liquid fuel. Conversion of coal to a low-sulfur liquid fuel previously relied on the use of hydrogen under high pressures and temperatures, in the presence of a catalyst. The Bureau of Mines was interested in a novel system which did not rely on hydrogen. This lead Appell, Wender et al. (1969) to discover that low-rank coals could be converted to low-sulfur oils using carbon monoxide and water, a result even more effective than hydrogen reactions. Low-rank coals are those that are low in carbon content and have less energy content such as lignite and sub-bituminous coals. These coals are less mature in their formation. As time, heat and pressure increases, the energy content increases and so does the rank. So these low-rank coals are low in energy content in a similar way that organic matter is low in energy content.

To understand why the reaction of coal to carbon monoxide and water was so effective, carbon monoxide plus water was reacted with several other model compounds to produce oil and compared to their respective reactions with hydrogen. Two of those model substances tested were cellulose and lignin, both primary components of plant material. The reaction and subsequent production of oil of cellulose and lignin with carbon monoxide and water was more rapid and complete than the reaction with hydrogen. Later, Appell, Miller et al. (1970) converted urban refuse, wood wastes, cellulosic wastes and sewage sludge to a low-sulfur oil by heating the substances under pressure with carbon monoxide and steam. In addition to urban refuse, agricultural wasters, wood and sewage sludge Appell, Fu et al. (1971) further develop their conversion processes by producing oil from bovine manure and lignin. By this time the U.S.

Bureau of Mines had developed liquefaction, hydrogasification and pyrolysis processes to convert biomass into fuel (Steffgen 1974).

The importance of this new literature was heightened with the 1973 oil embargo which caused fuel shortages and rapidly rising prices. Dugas (1973), provided an early techno-economic analysis of the yields and feasibility of producing energy from crops, trees and algae using bacterial fermentation, yeast fermentation and pyrolysis conversion processes. Szego and Kemp (1973), Steffgen (1974), Grantham and Ellis (1974) and Bassham (1976) , discuss the potential advantages and disadvantages of dedicated energy plantations and specially grown forests for energy. Grantham and Ellis (1974), found that standing timber, mill and logging residues may be used as fuel inside the logging industry, but that fueling a large share of national energy demands would be infeasible due to the vast land requirements necessary.

As the research progressed and technological feasibility became established, the literature turned to the economic feasibility of biofuels. The early studies of economic feasibility ranged from regional studies to large comprehensive national analysis.

In reaction to the 1973 oil embargo, President Richard Nixon announced Project Independence, with the goal of achieving energy self-sufficiency by 1980 (USDOE 2016). The Project Independence Task force conducted a macroeconomic assessment of competing demands on the nation's production capacity through 1985, and of the methods by which investment spending by the private may be financed. The task force estimated that if an accelerated plan were followed biomass could provide up to eight percent of the nation's energy demands by year 2000 (Kelly and White 1974) and (Alich and Inman 1976).

Alich and Inman (1976), discuss the potential of biomass fired electricity generation power plants and provide a techno-economic analysis of direct biomass combustion, gasification/turbine generation and substitute natural gas production. The authors found costs for electricity generation via biomass combustion to be reasonable, but limited by transportation costs. Therefore power plants located near biomass supplies were found to be advantageous.

Oursbourn, Lacewell et al. (1978), evaluate the collection and transportation costs of delivering agricultural residues to centrally located conversion facilities in Texas. High costs of collection and transportation lead the authors to conclude that conversion costs would need to be between 32 and 36 percent the cost of natural gas for the use of crop residues to be feasible in replacing natural gas in Texas. However, the authors note several limitations to the study such as the exclusion of indirect costs, storage costs, crop yield variabilities and values of the crop residuals themselves.

In 1979 the Office of Technology Assessment released an extensive study of alcohol based fuels from agricultural products titled *Gasohol*. The report examines the technical, economic, environmental and social factors related to the use of gasohol, a blend of one part ethanol and nine parts gasoline. One of the key findings of this study is that 1-2 billion gallons of ethanol could be produced each year with minimal impact on food and feed prices. The study noted that by the 1990's ethanol might be competitively produced using cellulosic feedstocks such as crop residues and wood, having little impact on food prices. The study also found the choice of feedstock is the largest factor in determining the environmental impact of gasohol. Additionally, the expansion of ethanol production through the use of gasohol would result in intensive cultivation of new land to provide additional feedstocks. If corn is the chosen feedstock, the authors warn of significant increases in soil loss along with fertilizer and pesticide use. Other

feedstocks such as perennial grasses would have much less of an erosion effect. In addition, the authors of the study noted that investment in the production of ethanol would be limited by uncertainties in commodity prices, feedstock availability and conversion technology costs. The authors further noted that long term viability of an ethanol industry would be dependent on sustained market demand, and costs of converting cellulosic feedstocks. Large capital investments in cellulosic processes were found to be a major constraint in long term viability of the ethanol industry (USOAT 1979).

From 1978 to 1986 ethanol fuel from biomass production increased from about 4 to 5 million gallons per year to about 850 million gallons per year. This increase in production was mostly driven by the Energy Tax Act of 1978 which provided gasoline blended with 10 percent ethanol an exemption from the federal gasoline excise tax and energy investment tax credits for equipment that converted biomass into ethanol (Zerbe 1988). In a report by the U.S. Department of Agriculture, ethanol was found to be economically infeasible without massive government subsidies and the fuel ethanol industry would be terminated or sharply curtailed after 1992 when federal subsidies were scheduled to expire (USDA 1986).

Since 1977 the USDOE had been supporting the development of woody biomass through the Short-Rotation Woody Crops Program (SRWCP). The aim of this program is to reduce uncertainty and risks associated with short-rotation intensive crops (SRIC) and move the technology closer to commercialization through the private sector. Perlack, Ranney et al. (1986), provide a review of the SRIC technology noting that adoption by the private sector would depend on future cost reductions and the prices of fossil fuels. Cost reductions in the form of technological and genetic advances along with reductions in harvest costs were noted as the two key potential drivers for commercialization.

Zerbe (1988), noted that producing ethanol from non-grain feedstock was possible, but production and research to improve conversion technologies were deterred by low oil prices. Imported oil prices peaked in 1980 at $98.02 per barrel before falling to $30.30 per barrel in 1986, in February 2016 real dollars. Oil prices did not significantly increase until the year 2000 when imported oil prices surpassed the 1986 low of $30.30 by averaging $38.40 in the year 2000 (USDOE 2016). A low oil price throughout this period was a major deterrent to private investment into non-grain biomass fuel production. However, this period of low oil price was not the only deterrent as conversion of non-grain biomass feedstocks such as woody biomass were typically found to be economically infeasible. Strauss, Blankenhorn et al. (1988), found the conversion of woody biomass to ethanol was not economically feasible due to a prohibitively high cost of woody biomass supply.

The U.S. Department of Energy in 1983 established the Biofuels and Municipal Waste Technology Division (BMWT) to encourage the development of a biomass energy industry infrastructure, through the Regional Biomass Energy Program. The goal of the program was to identify and plan for regionally appropriate biomass technologies and to advance the production of biomass feedstocks and their conversion to fuels and energy by the private sector, including the conversion of municipal waste (Lailas 1989).

Although public interest in alternative fuels waned during the 1980s after oil prices fell in 1986, the ethanol industry continued to grow at a moderate rate. In 1992 about 1 billion gallons of ethanol were produced, the industry being supported through a mix of Federal and State incentives. The Clean Air Act of 1990 gave rise to a market for oxygenates, creating a new market for ethanol. The Energy Policy Act of 1992 put forth a series of regulations and incentive programs involving fleet vehicle standards, alternative fuels, energy conservation and tax credits.

The Act established mandates for Federal fleet vehicles to run on alternative fuels such as electricity, ethanol and other biofuels. The Act also promoted adoption of alternative fuels among the fleet vehicles in the private sector and also established a program on the production and use of biodiesel (Shapouri and Duffield 1993). These two acts of legislations provided a significant boost and support to the biofuels industry and reignited the literature on the economics of producing crops as biomass energy feedstocks. During this period a number of studies focused on regional cost-benefit analysis such as traditional farm budgeting (Reese, Aradhyula et al. 1993, Shapouri and Duffield 1993, Epplin 1996). A continued evolution in the literature resulted in the incorporation of enterprise budgeting and partial equilibrium analysis of the agricultural sector, along with geographic information system-based modeling (English, Menard et al. 2000, Graham, English et al. 2000, Ugarte, Walsh et al. 2000, Ugarte and Ray 2000, Walsh, Daniel et al. 2003).

In 2005, the ethanol industry received a major increase in support. The Energy Policy Act of 2005 established the Renewable Fuel Standard (RFS) requiring the use of biofuels such as ethanol and biodiesel in the U.S. automotive fuel supply. The RFS mandated that ethanol production increase from the 4 billion gallon level of 2006 to 7.5 billion gallons per year by 2012 (Markel, Clark et al. 2014). The Act also created a renewable energy research budget of over $2.2 billion through year 2009.

The Energy Independence Security Act (EISA) of 2007 enacted the Renewable Fuel Standard Two (RFS2). The new RFS2, (RFS from here out) included blending mandates for diesel and gasoline, and created new categories of renewable fuel and GHG performance thresholds for each category. The volumes of renewable fuels to be blended into transportation fuel were expanded from 7.5 billion gallons per year by 2012 to 36 billion gallons per year by 2022. The

RFS is implemented by requiring refiners that produce gasoline or diesel fuel in the United States and anyone who imports gasoline or diesel fuel into the United States, to meet Renewable Volume Requirements (RVOs). Compliance with RVOs is shown by submitting a tradable credit for every gallon of renewable fuel blended with gasoline or diesel. Those credits are known as renewable identification numbers (RINs). RINs are obtained by blending renewable fuel or by purchasing them on a secondary market. Expanding the blend mandate and establishing tradable credits called RINs, created downstream market effects and a new established market. The RIN program is the subject of much debate and ongoing research. Understanding the behavior of RIN prices and how these prices effect firm-level investment decisions is critically important and thus the subject of a majority of this research.

This research is composed of three parts. In chapter one, historical RIN price behavior and drivers of RIN price are examined. Historical RIN prices have been volatile, exhibiting large upward and downward spikes. Determining how the volatility impacts investment decisions is the subject of chapter two. Finally, the research concludes with chapter three, a study of economic feasibility for an experimental feedstock and its impact on agricultural markets.

Chapter I: Thresholds and Regime Change in the Market for Renewable Identification Numbers

Chapter I: Abstract

This paper examines potential nonlinearities in RIN price and applies nonlinear time-series techniques to model historical RIN prices. The central hypothesis is that RIN prices are nonlinear and the behavior of RIN prices fall under at least two regimes. Current vintage RIN prices are used to capture immediate reflexes to price drivers. A smooth transition autoregressive model is applied to address nonlinearities and test for the most significant transition variable. This type of model is well suited to handling nonlinearity and regime changes, such as those, which occur with RFS revisions.

Throughout their history, RIN prices have exhibited periods of a coupling behavior where biomass-based diesel D4 RINs and conventional renewable D6 RINs converge in price. The coupling of the RIN prices suggests biodiesel to be the marginal fuel used in overcoming the blend wall. Thus over-complying with the biodiesel mandates. During these periods, the ratio of D6/D4 is near, or equal to one. In other periods, the value of the ratio falls towards zero, suggesting at least two regimes. Examining drivers of regime change in RIN prices, and their ratio, provides important insight into a complex policy program known as the RIN market.

1.1. Introduction

Tradable credits known as Renewable Identification Numbers (RINs) are a key policy instrument used in the implementation of the Renewable Fuel Standard (RFS). Producers and importers of gasoline and/or diesel (referred to obligated parties) are obligated to blend a mandated percentage of biofuel into their product supplied. Obligated parties show compliance with the law by submitting one RIN for every gallon of biofuel blended. Therefore, the price of the credit represents the marginal cost of complying with the law. Understanding the drivers of RIN price and their past behavior is therefore critically important to the sustainability of the RFS program and its effect on the Nation's energy industry. A central hypothesis in this study is that RIN prices exhibit nonlinear behavior governed by at least two price regimes. Furthermore, it is hypothesized that the regime switching behavior is driven by the implied-ethanol mandate's proximity to the ethanol blend wall.

A number of studies and reports have examined RIN price fundamentals, developing a theoretical framework of the core value in RIN price. Thompson, Meyer et al. (2009) discuss core RIN values, and the hierarchal nature of RIN values. To demonstrate the core value of RINs, Thompson, Meyer et al. (2010) use complementary slackness equations, and supply-and-use tables to simulate RIN price. McPhail, Westcott et al. (2011), provide a conceptual model of RIN prices and discuss the factors affecting RIN price. The program is administered by setting an overall mandate and three submandates, resulting in four RIN types representing the four fuel categories. Of these four RIN types, only two were produced in significant numbers in the early years of the program. Those two RINs include conventional renewable (i.e. ethanol) RINs and biomass-based biodiesel RINs. Due to the hierarchical nested structure of the mandates, biodiesel

RINs should always be worth at least as much as ethanol RINs[1]. Therefore, the relationship between these two RIN prices has been an important indicator of the current state of the RIN market, since the two began trading. Thompson, Meyer et al. (2011) provide an early discussion of this relationship. At the time of writing, historical RIN data were not publicly available. However, the authors note that when biodiesel RIN price is greater than ethanol RIN price, the overall mandate is less binding than the biodiesel mandate. When the overall mandate is more binding than the biodiesel mandate, obligated parties are indifferent between using ethanol or biodiesel RINs to show compliance, causing their prices to converge. Due to the lack of historical data, the authors simulate RIN prices using a partial-equilibrium model. Through forward looking simulation results, the authors find that ethanol RIN prices remain relatively low while advanced and biodiesel RIN prices remain relatively high. Indicating the advanced and biodiesel mandates were the most difficult to satisfy. However, results partly depend on an expansion of 85% ethanol-blended gasoline (E85) consumption, thus overcoming the blend wall. At the time of their study, biodiesel RINs traded well above the price of ethanol RINs. However, the prices did indeed converge shortly after regulators proposed a significant strengthening of the overall mandate in 2013. With the availability of historical data, researchers have begun analyzing observed RIN price behavior. Whistance and Thompson (2014) examine RIN price behavior to determine if prices follow the nested-hierarchy expectations. Lade, Lin et al. (2015), show that prices reflect current and expected marginal compliance costs, along with expected RFS mandates.[2] As discussed in Lade, Lin et al. (2015), the coupling of ethanol and biodiesel

[1]For simplicity, this category can be thought of as the ethanol category. However, readers should be aware that other fuels could satisfy the requirements of this category, such as the case with butanol.
[2](Lade, Lin et al. 2015) find RIN prices to be non-stationary. However, the authors find no evidence of cointegration among RIN types or energy price series.

RINs suggests that the industry expectation was for biodiesel to be the marginal fuel pushing the industry beyond the blend wall and thus overcomplying with the biodiesel mandates. Expansion of the E85 market has not occurred for a number of reasons, leaving obligated parties with two methods of overcoming the blend wall. Either they overcomply with the biodiesel mandate, or they reduce production and imports, thereby reducing their renewable volume obligations.

Market expectations about the ability to meet an overall mandate beyond the blend wall, and the aforementioned change in RIN price behavior, lead to potential nonlinearities in RIN price. Due to the nested structure of the submandates, biomass-based diesel RINs should always be worth at least as much as the ethanol RINs. Therefore, the ratio of ethanol RIN price over biomass-based diesel RIN price should never exceed one. When this ratio is equal to one, or near one, this indicates that biodiesel is the marginal fuel of compliance (Irwin 2016). Meaning that, the market cannot absorb any more production of ethanol, forcing obligated parties to overcomply with the biodiesel mandate in order to generate enough RINs to meet the overall mandate. As the value of the ratio decreases, market participants are signaling their expectations are that the blend wall has been overcome, or is no longer binding due to decreased demand in gasoline and diesel. The value of this ratio can thus be broken into two or more regimes. Supposing there are only two regimes, the blend wall may be binding in the regime where the ratio is close to one and non-binding when the ratio is close to zero. However, there could be three regimes. In a third regime, the blend wall may be binding, but circumvented by reductions in gasoline and diesel production. Alternatively, perhaps the blend wall has been circumvented by increased blends of biodiesel, generating a sufficient surplus of biodiesel RINs to meet the implied ethanol mandate.

The ratio of ethanol RIN price over biodiesel RIN price is analyzed to examine potential nonlinearities and drivers of potential regime change. Following this, individual ethanol and

biodiesel RIN prices are examined for potential nonlinearity and regime change. Smooth transition autoregressive models are applied to the individual prices and to the ratio of prices. These types of models are well suited to handling nonlinearities and regime changes, such as those, which occur with RFS revisions. In this framework, it is also possible to test for the most significant transition variable in determining regime change. A candidate set of transition variables are considered, including a variable, which represents the implied weekly ethanol mandate. This variable is constructed by taking the product of the fractional ethanol mandate and weekly gasoline and diesel supplies.[3] It thus approximates the volume of ethanol blended on a weekly basis, excluding cellulosic ethanol and other advanced ethanol use (Good and Irwin 2015).

1.2. Renewable Identification Numbers

The RFS, established by the Energy Policy Act of 2005 (EPAct), requires refiners that produce gasoline or diesel fuel in the United States and anyone who imports gasoline or diesel fuel into the United States (U.S. Environmental Protection Agency 2014) to meet four different renewable volume requirements (RVOs): (i) total renewable fuel, (ii) advanced biofuel, (iii) biomass-based diesel, and (iv) advanced cellulosic biofuel.[4] The RVOs are based on the ratio of

[3] The RFS does not explicitly mandate ethanol. Therefore, the fractional ethanol mandate is implied as the annual percentage standard for total renewables minus the annual percentage standard of advanced biofuel (Irwin and Good 2015). The percentages are set by the U.S. Environmental Protection Agency and determined on an annual basis with revisions possible to occur year to year. The weekly product supplied for gasoline and diesel are provided by the U.S. Energy Information Association. Weekly product supplied is an approximation of weekly consumption as it measures the disappearance from primary sources such as refiners, blenders and distribution terminals (USDOE EIA 2016).

[4] For example, the RFS mandated that ethanol production increase from the 4 billion gallons of ethanol produced in 2006 to 7.5 billion gallons per year by 2012. Two years later the Energy Independence and Security Act of 2007 (EISA) amended the RFS to include (i) blending requirements for diesel as well as gasoline, (ii) an increase in the volume of renewable fuel required to be blended into transportation fuel, (iii) new categories of renewable fuel with separate minimum volume requirements for each category, and (iv) lifecycle greenhouse gas performance threshold

renewable fuels to all non-renewable gasoline and diesel fuels, and may be amended from year-to-year. For example, the EPA set the 2016 annual percentage standard for total renewable fuel at 9.63%, advanced biofuel at 1.88%, biomass-based diesel at 1.49% and cellulosic biofuel at 0.059%.[5] These annual percentage standards are based on mandated annual volume requirements and projected annual consumption of gasoline and diesel.

Obligated parties demonstrate compliance with RVOs using a tracking system developed by the EPA. In this system, every gallon qualifying as a renewable fuel produced or imported is assigned a RIN. The type of RIN generated depends on the type of renewable fuel being produced, feedstock and production process. To qualify as a renewable fuel under the RFS, the fuel's lifecycle GHG emissions must be less than the lifecycle GHG emissions of the 2005 baseline average gasoline or diesel fuel that it replaces. How much less, depends upon the category of biofuel. Table 1 summarizes RIN types, D-codes, and the primary fuel types that qualify under the corresponding categories.[6]

A qualifying feedstock is tracked through its processing at the biofuel plant and a RIN is attached to that batch of biofuel. The RIN must not be separated from the renewable fuel as it moves through the distribution system, and is transferred along with the renewable fuel as ownership changes. Once the renewable is blended into a conventional fuel, the RIN is separated and used

---

standards requiring the renewable fuel used to satisfy the RFS emit fewer greenhouse gases (GHG) than the petroleum-based fuel it replaces (Yacobucci 2012).

[5] These minimum requirements translate to a total renewable fuel percentage standard for 2013 of 9.74% (since the 16.55 billion gallons of renewable fuel represented 9.74% of all fuel volume), an advanced biofuel percentage standard of 1.52%, a biomass-based diesel percentage standard of 1.13% and a cellulosic biofuel percentage standard of 0.004% (U.S. Environmental Protection Agency 2013).

[6] U.S. Environmental Protection Agency (2016), provides a full listing of approved pathways and detailed fuel types.

for compliance by obligated parties, or it may be traded as a stand-alone commodity (McPhail et al 2011).

1.2.1. RIN Pricing

The fundamental value, also known as the core value of a RIN, is determined by the gap between the cost of supplying biofuel, and the price at which blenders are willing to pay for biofuel at mandated quantities. If there is no gap between the cost of supplying biofuel at the mandated quantity and the price blenders are willing to pay for RFS2-mandated quantities, then the market would be in equilibrium and there would be no need for the RIN markets. RIN credits subsidize the biofuel market by providing an incentive for the market to trade at quantities greater than the market equilibrium when the mandate is not in place. When the mandate is binding, RIN prices are positive. When the mandate is non-binding, RIN prices are zero. A non-binding mandate implies renewable fuel production levels in equilibrium are greater than RFS mandated requirements, rendering the mandate superfluous. Conversely, a binding mandate implies renewable fuel production levels in equilibrium are lower than RFS mandated requirements, necessitating the market for RINs. A binding mandate is necessary for a strong RIN market.

Figure 1 illustrates the core value of a RIN in ethanol markets (panel (a)) and biodiesel markets (panel (b)). Both figures follow conventions found throughout the literature, with the distinction between the two markets being made at the slope of the respective demand curves. Demand for biodiesel is horizontal and thus perfectly elastic for prices of biodiesel equal to conventional diesel. Biodiesel is not constrained by a blend wall, and has an energy content 93%-99% of conventional low-sulfur diesel. Higher blends of biodiesel have less potential to inflict damage on engines and current infrastructure. Blends of 20% biodiesel are common in transportation fuels. Higher blends or even 100% biodiesel transportation fuels are compatible with a large

portion of vehicles built since 1994 (AFDC 2016). Thus, biodiesel is often modeled as a near perfect substitute for conventional fuel, resulting in perfectly elastic demand for prices equal to conventional diesel. Ethanol is constrained by the blend wall and has an energy content 73-83% of gasoline. Higher blends of ethanol have damaging effects on vehicle engines. Many automobiles manufactured today cannot handle blends of ethanol exceeding ten percent. Some manufacturers offer vehicles, known as Flex-Fuel vehicles, which can handle up to 85% ethanol blends. However, demand for the flex-fuel variety has been relatively low. For these reasons, the ethanol market is often modeled as having a downward sloping demand curve as in panel (a) of Figure 1.

The nested nature of the submandates means that biofuels, which qualify for multiple RIN categories, have a higher value. Therefore, an advanced biofuel RIN will be worth at least as much as a conventional biofuel RIN because advanced biofuel RINs count towards both the advanced biofuel RVO and the ethanol RVOs. Similarly, biodiesel D4 RINs will be worth at least as much as D5 advanced biofuel RINs because they count towards the biodiesel RVO, the advanced biofuel RVO or the ethanol RVO (Figure 2). How much of a premium exists for D4 over D5 and D6 depends on the relative levels of advanced mandates versus conventional mandates and market equilibriums. However, due to this nested structure and aside from temporary trading anomalies, RIN prices satisfy

$$R_{D3}, R_{D4} \geq R_{D5} \geq R_{D6} \, .$$

Beyond the core value, RINs also have a time value, or a speculative component. RINs can be held for up to two years and applied towards 20% of future renewable volume obligations. Therefore, the observed price of a RIN can be represented as $R = core\ value + time\ value$ (Irwin 2014). As discussed above, the core value is represented by the gap between the biofuel

supply price and the conventional fuel price, plus the value of the blender's tax credit, for years when the credit is available. This gap, plus the blender's tax credit is known as the blend margin and is synonymous with the core value of the RIN.

Hence, the $core\ value = (P_b - P) + blender's\ tax\ credit$ and the observed price of a RIN (R) can be rewritten as

$$R = [(P_b - P) + blender's\ tax\ credit] + time\ value.$$ [7]

The core value is the largest component of observed RIN prices and a simple linear regression, with no constant, results in R^2 of 0.8660. This indicates that about 86% of the variation in observed RIN price is explained by the blend margin and blender tax credit.

Because RINs are bankable, an obligated party can accrue surplus RINs during periods of low compliance cost, to hedge against future costlier periods. This characteristic adds to the time-value in a RIN. However, market expectations also provide a time value to the RIN.

Non-obligated parties may participate in the RIN market by trading RINs much like any other commodity. If speculators anticipate an increase or decrease in RIN prices, they may buy or sell excess RINs, further increasing or decreasing RIN prices. An example of a non-obligated party is a producer or importer with an output of less than 75,000 barrels per day, or producers or importers of jet fuel and other fuels not falling under the umbrella of the RFS. The price of imports also affects RIN prices by causing a shift in the domestic supply of renewable fuels. Increases in crude oil prices also create stronger demand for alternative fuels, further

[7] In the years, 2011, 2013 and 2016 biodiesel blenders were eligible to receive a tax credit of $1.00 for every gallon of biofuel blended. Biodiesel RINs are multiplied by 1.5 to convert prices to ethanol equivalent values (Irwin 2014).For example, one gallon of biodiesel generates 1.5 D4 biodiesel RINs.

incentivizing blending with biofuels and thereby reducing RIN value. Increased costs of feedstock production reduce feedstock supply and increase RIN value, while decreases in feedstock production costs increase feedstock supply and negatively affect the value of RINs.

Blender tax credits incentivize blending with more biofuels, increasing biofuel demand and potentially causing a decrease in RIN price. Despite this, uncertainty about renewed availability of tax credits can counteract and lead to increased RIN price. In addition, market expectations about the future stringency of the policy can affect the time value of a RIN. In fact, uncertainty regarding RFS policy has been shown to have a significant effect on RIN price, a topic that will be revisited in the following chapter.

1.3. Data

RIN price data is provided by EcoEngineers, a provider of daily and historical RIN price.[8] Data ranges from January 6 2010 to April 15 2016 for D4, D5 and D6 RIN credits. To examine the movements of these three RIN credits over time and in relation to each other, it is useful to plot them together against time. Ethanol RINs began trading in April of 2008, biodiesel RINs began trading in September of 2009 and advanced biofuel RINs began trading in January of 2011 (AGMRC 2015). Data in this study begins in January of 2010 when biodiesel RINs and ethanol RINs were trading at similar prices. The two prices diverged with biodiesel RINs trading at a premium over ethanol RINs between 2010 and 2013. The three RIN prices began to converge following jumps in D6 RIN value, which began in early 2013. On January 2, the first trading day

[8]EcoEngineers provides comprehensive compliance management, market data, and facility planning and project development services for biofuel companies operating under US regulations. All price indices utilize volume-weighted averages either in the calculation of the index or as a component of the calculation of the Index. Weighted averages are utilized in an effort to minimize any trading anomalies or distress trading activity that might otherwise distort the data sample (EcoEngineers 2012).

of 2013, the price of a D6 ethanol RIN was just seven cents. Over the next 41 trading days, D6 RIN price jumped from seven cents to 74 cents per RIN.

By March 4 2013, the D4 premium was just five cents and on December 5 2013 ethanol RINs eliminate the biodiesel RIN premium. The D4 premium over D5 advanced biofuel RINs narrowed in October of 2012 when D4 is priced just eight cents over D5 on October 11 2013. Between the start of D5 trading in January of 2011 and October 10 2012, the average D4 premium was $0.57. This premium is to be expected from the nested nature of the RIN credits. The average D4 premium over D6, between January 6 2010 and March 4 2013, is $0.8758. The D4 premium over D5 from March 4 2013 to December 21 2015 averages $0.0306, while the D4 premium over D6 during the same period averages $0.1218. The convergence pattern is strongest between March 4 2012 and December 31 2014. The total spread is the sum of the individual spread between D4 and D6 and the spread between D4 and D5. During this period of strong convergence, the mean total spread is $0.1068.

The first trading day of 2015, January 2, is responsible for a significant departure from the strong convergence pattern. On this day, D4 RIN prices increase by more than 10 cents per RIN, signaling the beginning of a weakly convergent pattern. From this point, the convergence is much weaker with the total spread between the three RIN prices calculated to be a mean of $0.2435. The weakly convergent period is mostly attributable to diverging D6 ethanol RIN price, as D4 and D5 RINs continue to track each other in trading price.

Since D5 advanced biofuel RINs are lower and upper bound by ethanol and biodiesel RINs, they are excluded from the analysis. The focus turns to individual price of ethanol and biodiesel RINs, and the ratio of these two RIN prices. The ratio of D6/D4 RIN price is utilized to examine the

convergence of ethanol and biodiesel RIN price. If this ratio exhibits nonlinearity, it may signal important regime changing behavior in the underlying data generating process.

1.4. Nonlinearity in RIN price

1.4.1 Visually inspecting for nonlinearity

One method of identifying nonlinearity in a time series Y_t is to examine the joint distribution of Y_t and Y_{t-1} or Y_{t-s} where $t \neq s$. If the time series Y_t and Y_{t-s} are not jointly normal, there is evidence of nonlinearity. The Wold decomposition shows that the best linear predictor is the best one-step-ahead predictor if and only if innovation terms e_t satisfy the martingale condition. The martingale condition states that the conditional mean of e_t given e_{t-s} is identically equal to zero and this condition holds when e_t is a sequence of independent, identically distributed random variables with zero mean. Conversely, when the martingale condition fails, a nonlinear predictor will be the best one-step-ahead predictor. Given the Wold decomposition, the best linear predictor is approximated by finite ARIMA models. In linear ARIMA models, the errors are assumed independent and identically normally distributed. The normal error assumption implies the time series is also normally distributed and thus any two sets of time series are jointly normal. Therefore, nonlinearity can be identified by finding a non-normal joint distribution of Y_t and Y_{t-s} . To perform this task, a scatter diagram of Y_t against Y_{t-s} is plotted (Cryer and Chan 2008).

In this case Y_t represents the first-differenced ratio of D6 over D4 RIN price.[9] To aid in the visualization, a nonparametric nearest neighbor regression is fit to the data in each scatter plot.

[9]The same procedure is later carried out for the individual prices although the figure is not presented to reduce redundancy.

The nearest neighbor method fits a locally weighted polynomial regression. Each nearest neighbor regression is specified as a first-degree polynomial using 30% of the sample to calculate the bandwidth and local (Tricube) weighting. By weighting the observations, those points furthest from the local regression point are weighted less. Therefore the weighted regression minimizes the weighted sum of squared residuals providing a locally weighted scatterplot smoothing (Lowess) technique (Cleveland and Loader 1996). By fitting the nearest neighbor regression, the nonlinearity can be visually observed (Figure 6).

Figure 7 illustrates the historical price ratio in levels. From this figure, it appears the ratio is in at least two regimes, with a possible third regime. The transition between the first and second regime occurs abruptly, while the second transition appears to occur smoothly. Figure 8 illustrates the bi-modal distribution of the price ratio in levels. The presence of multiple modes is another visual indication of nonlinearity in the price ratio.

The price ratio in levels is nonstationary and integrated of order one. Therefore, the first-difference is applied to generate a stationary time series. By examining the stationary first-differenced ratio, one can see the possibility of a third regime occurring in 2014 and beyond (Figure 9).

1.4.2 Testing for Nonlinearity

Beyond visual inspection, empirical tests provide further evidence of nonlinearity. Two particular tests have gained in popularity over the years, the test for quadratic nonlinearity, (Tsay 1986), and a Tukey non-additive type test for nonlinearity (Keenan 1985).

Keenan's test for nonlinearity is based upon a general form of nonlinear stationary time series models that are known as Volterra expansions (Cryer and Chan 2008). Volterra expansions take the form

$$Y_t = \mu + \sum_{\mu=-\infty}^{\infty} \alpha_\mu \varepsilon_{t-\mu} + \sum_{\mu=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \alpha_{\mu v} \varepsilon_{t-\mu} \varepsilon_{t-v} + \sum_{\mu=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \sum_{w=-\infty}^{\infty} \alpha_{\mu v w} \varepsilon_{t-\mu} \varepsilon_{t-v} \varepsilon_{t-w} + \cdots$$

where $\{ \varepsilon_t, -\infty < t < \infty \}$ is a strictly stationary process, assumed to be independently and identically distributed with a zero mean. The right hand side terms in the expansion include an intercept, linear, quadratic and cubic terms. However, the test of linearity is equivalent to a test of no multiplicative terms. Therefore, the null hypothesis tests whether or not higher order expansion terms vanish.

Keenan (1985), provides a three step process to estimate the test for nonlinearity.

i. Regress Y_t on $\{1, Y_{t-1,} Y_{t-2}, \dots, Y_{t-p}\}$ where p is a predetermined lag order. Obtain the fitted values $\widehat{Y_t}$, predicted residuals $\hat{e}_t$ for $t = p + 1, \dots, n$ and calculate the sum of squared residuals ${\hat{e}_t}^2$.
ii. Regress ${\widehat{Y_t}}^2$ on $\{1, Y_{t-1,} Y_{t-2}, \dots, Y_{t-p}\}$ and obtain the predicted residuals $\hat{\epsilon}_t$ for $t = p + 1, \dots, n$.
iii. Regress $\hat{e}_t$ on $\widehat{\epsilon}_t$ with no intercept for $t = p + 1, \dots, n$. This allows the user to obtain $\hat{\eta} = \hat{\eta}_0 \sqrt{\sum_{t=p+1}^{n} \hat{\epsilon}_t^2}$ where $\hat{\eta}_0$ is the regression coefficient.

Keenan's test statistic is thus

$$\hat{F} = \frac{\hat{\eta}^2 (n - 2p - 2)}{{\hat{e}_t}^2 - \hat{\eta}^2}$$

Under the null hypothesis of linearity $\hat{F}$ approximately follows an F-distribution with degrees of freedom 1 and $n - 2p - 2$. Keenan's test is equivalent to testing that the coefficient $\hat{\eta} = 0$ in the regression model

$$Y_t = c + \beta_1 Y_{t-1} + \cdots + \beta_p Y_{t-p} + \exp\left\{ \eta \left(\sum_{j=1}^{p} \beta_j Y_{t-j} \right)^2 \right\} + \varepsilon_t.$$

If the coefficient η equals zero, then $\exp(0) = 1$ and the model simply becomes an AR (p).

Tsay (1986), extended Keenan's test improving the power of the test to account for more general nonlinear terms. These more general terms are accounted for by

replacing $\exp\left\{\eta\left(\sum_{j=1}^{p}\beta_j Y_{t-j}\right)^2\right\}$ with

$$\exp\{\delta_{1,1}Y_{t-1}^2 + \delta_{1,2}Y_{t-1}Y_{t-2} + \cdots \delta_{1,p}Y_{t-1}Y_{t-p} + \cdots$$
$$+\delta_{2,2}Y_{t-2}^2 + \delta_{2,3}Y_{t-2}Y_{t-3} + \cdots \delta_{2,p}Y_{t-2}Y_{t-p} + \cdots$$
$$\delta_{p-1,p-1}Y_{t-p+1}^2 + \delta_{p-1,p}Y_{t-p+1}Y_{t-p} + \delta_{p,p}Y_{t-p}^2\}\,.$$

The null hypothesis in Tsay's test is that all coefficient terms $\delta_{i,j}$ are equal to zero. The test statistic follows an F-distribution and thus an F-test that all $\delta_{i,j}'s$ equal zero, tests the null hypothesis that the true process is linear. The null hypothesis of linearity, in the fist-differenced D6/D4 price ratio of lag-order two, is rejected with p-values of 0.0126 and 0.0811 in the Tsay and Keenan tests respectively.[10] In the following section, additional tests of linearity are carried out against a specific alternative.

1.5. Thresholds in the RIN Market

As D6 and D4 RIN prices converge, the price ratio nears unity and prices exhibit a coupling patter. A central hypothesis in this paper is that the coupling patterns observed in the RIN market are brought about by some proximity to the ethanol blend wall. In other words, there is some unobservable threshold, that when crossed, brings about a change in the behavior of the RIN market. Therefore, a key variable in this paper is the fractional ethanol mandate multiplied by weekly product supplied of gasoline and diesel. This is referred to as weekly ethanol mandate. Weekly product supplied is an approximation of weekly consumption because it measures the

[10]Similarly, the null the hypothesis of linearity in the individual D4 and D6 RIN price is rejected under both the Tsay and Keenan. The table of results is presented in the chapter appendix.

disappearance of gasoline and diesel from primary sources such as refiners, blenders and distribution terminals (USDOE EIA 2016). The ethanol blend wall is generally thought to be ten percent of gasoline consumption. The weekly ethanol mandate approximates the volume of ethanol blended on a weekly basis, excluding cellulosic ethanol and other advanced ethanol use. Where this observed data lies in relation to some unobserved threshold is presumed to trigger regime changes in the ratio of D6/D4 RIN price and in the individual D6 RIN price.

Prices for D4 biodiesel RINs are driven primarily by biodiesel blend margins and other market forces unrelated to the ethanol blend wall (Irwin 2014). Therefore, the primary threshold variable is not expected to be the weekly ethanol mandates. Rather, it is hypothesized that regime changes in D4 RIN prices are triggered by endogenous factors to the D4 RIN market. Therefore, the key threshold variable is expected to be lagged values of D4 RIN price.

1.5.1 Model Selection for D6/D4 Price Ratio

The first step in building a nonlinear model is to specify a linear model to form a starting point for further analysis. The proposed linear model for the D6/D4 RIN price ratio is one of lagged dependent variables, and two exogenous variables including, ethanol blend margins, and biodiesel blend margins. Ethanol blend margins are defined as the difference between ethanol price and gasoline price, and applicable blender tax credits for the periods when such tax incentives were in effect. Therefore, the blend margins are composed of three explanatory variables including ethanol price, gasoline price and the value of blender tax credits. Similarly, the biodiesel blend margins are composed of three explanatory variables to include the biodiesel price, conventional diesel price and biodiesel-blender tax credits. The blend margins approximate the core value of RINs as depicted in Figure 1. When blending mandates are pushed beyond the blend wall, ethanol RINs should be in high demand and in fact, there may be a deficit

of available D6 RINs to fulfill the ethanol mandate. If that is the case, the deficit may be made up by submitting D4 RINs, as discussed earlier. In this state of the world, D6 RIN prices ought to be driven to their ceiling, which is defined by the floor of D5 and D4 RIN prices. Recall that since D5 RIN prices are bounded by D6 and D4 RINs, they are excluded from the analysis.

First, the optimal lag length of the price ratio is determined using an iterative process where the optimal lag is chosen to minimize a suite of information criteria such as Akaike's Information Criteria (AIC) and Bozdogan's Information Complexity Criteria (ICOMP). Based on the results, the optimal lag length for the D6/D4 RIN price ratio is determined to be lag-order one. This step is then carried out on the individual prices and a lag-order of one is found to minimize the AIC and ICOMP for both of the individual price series (Table 2).

Next, the linear AR (1) model is estimated via ordinary least squares for the price ratio and the individual prices. From the linear model, the first lag of the dependent variable is highly significant and the biodiesel blend margin are highly significant. Ethanol blend margins are not found to be statistically significant but remain in the model as this result may change in a nonlinear specification under different regimes.

Before proceeding with any model selection the residuals are examined for possible serial correlation and unspecified heteroskedasticity. The presence of heteroskedasticity is likely to have a significant effect on inference and thus model selection results. A LM-type test is carried out by regressing the squared residuals on explanatory variables to identify any possible heteroskedasticity dependent upon the model's regressors. However, the presence of serial correlation will invalidate any test of heteroskedasticity so a Breusch-Godfrey test for serial correlation is performed first. The null hypothesis is of no serial correlation in residuals up to a specified lag order. In this case, tests are carried out for second-order serial correlation.

Results in Table 3 indicate a failure to reject the null hypothesis of no second-order serial correlation in the price ratio. Therefore, a LM-type test of heteroskedasticity is carried out on the residuals of the linear model.

Table 4 reports the test results of the LM-type test of heteroskedasticity. Results strongly indicate the null hypothesis of homoscedasticity cannot be rejected. However when performing a test of autoregressive conditional heteroskedasticity, the null hypothesis is rejected with confidence. Therefore, Newey-West corrected standard errors are specified.

The proposed model for the RIN price ratio is a smooth transition autoregressive model with exogenous variables. The proposed transition variable is the weekly ethanol mandate as defined earlier. Log transforming the weekly ethanol mandate ensures stationarity. This model will be tested against a competing STAR model, where the transition variable is the endogenous lagged dependent variable. Furthermore, this model is compared to a threshold autoregressive (TAR) and self-exciting threshold autoregressive (SETAR) which is estimated in the next section.

1.5.2 Estimation of SETAR and TAR models

The two-regime threshold autoregressive model is defined as

$$Y_t = \begin{cases} X'\delta_1 + \beta_{1,1}Y_{t-1} + \cdots + \beta_{p_1,1}Y_{t-p_1} + \sigma_1\varepsilon_t, & if\ Z_{t-d} \leq r \\ X'\delta_2 + \beta_{1,2}Y_{t-1} + \cdots + \beta_{p_2,1}Y_{t-p_2} + \sigma_2\varepsilon_t, & if\ Z_{t-d} > r \end{cases}.$$

The autoregressive lag orders of the two sub models are not necessarily identical and are thus labeled as p_1 and p_2. Exogenous variables make up the X vector and the coefficient $\delta_j$varies among regimes. Extending the two-regime model to any m regimes is done by further partitioning the thresholds so that $r_0 < r_1 < r_2 < \cdots r_{m-1} < r_m$. The location of Z_{t-d} in relation to the thresholds will determine the sub models. For example, a four-regime threshold autoregressive model is represented as the piecewise regression of

$$Y_t = \begin{cases} X'\delta_4 + \beta_{1,4}Y_{t-1} + \cdots + \beta_{p_4,4}Y_{t-p_4} + \sigma_4\varepsilon_t, & r_3 \leq Z_{t-d} \\ X'\delta_3 + \beta_{1,3}Y_{t-1} + \cdots + \beta_{p_3,3}Y_{t-p_3} + \sigma_3\varepsilon_t, & r_2 \leq Z_{t-d} < r_3 \\ X'\delta_2 + \beta_{1,2}Y_{t-1} + \cdots + \beta_{p_2,2}Y_{t-p_2} + \sigma_2\varepsilon_t, & r_1 \leq Z_{t-d} < r_2 \\ X'\delta_1 + \beta_{1,1}Y_{t-1} + \cdots + \beta_{p_1,1}Y_{t-p_1} + \sigma_1\varepsilon_t, & Z_{t-d} < r_1 \end{cases}.$$

An indicator function $I_j(Z_{t-d}, r_j)$ where $j = 1,2,3, \ldots m$, takes the value of 1 if the expression is true and zero if false, is used to combine the piecewise regression. Therefore, the above piecewise regression of four regimes can be combined into a single nonlinear regression of

$$\begin{aligned} Y_t = \{&(X'\delta_1 + \beta_{p_1,1}Y_{t-p_1} + \sigma_1\varepsilon_t)I_1(Z_{t-d}, r_1) + (X'\delta_2 + \beta_{p_2,2}Y_{t-p_2} + \sigma_2\varepsilon_t)I_2(Z_{t-d}, r_1, r_2) \\ &+ (X'\delta_3 + \beta_{p_3,3}Y_{t-p_3} + \sigma_3\varepsilon_t)I_3(Z_{t-d}, r_2, r_3) + (X'\delta_4 + \beta_{p_4,4}Y_{t-p_4} \\ &+ \sigma_4\varepsilon_t)I_4(Z_{t-d}, r_3)\}. \end{aligned}$$

The identity of the specification is determined by the threshold variable Z_{t-d}. If Z_{t-d} is the d-th lagged dependent variable then the model is a self-exciting threshold autoregressive (SETAR) model. If Z_{t-d} is some other exogenous variable then the model becomes the threshold autoregressive (TAR) model.

The problem is therefore to estimate the coefficients $\delta_j, \beta_{p_j,j}$ and the threshold values r_j. To do so nonlinear least squares is performed to minimize the sum of squares objective function. The least square estimator $\hat{\theta} = \left(\hat{\delta}_j, \hat{\beta}_{p_j,j}, \hat{r}_j\right)$ solves the minimization problem

$$\begin{aligned} \hat{\theta} = argmin \sum_{t=1}^{n} \{&Y_t - \left(X'\delta_1 + \beta_{p_1,1}Y_{t-p_1}\right)I_1(Z_{t-d}, r_1) - \cdots - (X'\delta_4 \\ &+ \beta_{p_4,4}Y_{t-p_4})I_4(Z_{t-d}, r_3)\}^2 . \end{aligned}$$

This model will be tested against the competing SETAR model, where the threshold variable is the endogenous lagged dependent variable. First, the best fitting threshold delay parameter is chosen for the exogenous transition variable $ln(ethmandate)_{t-d}$. The delay parameter is allowed to vary from one to five while specifying the model and choosing the delay parameter, which minimizes the sum of, squared residual (SSR). While setting the maximum number of

possible regimes to four, the delay parameter, which minimizes the SSR, is $ln(ethmandate)_{t-2}$ with a SSR of 0.328225.

In the SETAR model, the threshold variable is the endogenous lagged dependent variable. Before testing the TAR model against the SETAR model, the best fitting delay parameter for lagged values of D6 are found by minimizing the SSR of the SETAR model. The delay parameter is allowed to vary from one to five, while specifying the model and choosing the delay parameter, which minimizes the SSR. Again, the maximum number of regimes is set to four and the models are iteratively estimated, capturing the SSR for each specification. In this case, the best fitting threshold variable for the SETAR model is found to be $dratio_{t-4}$ with a SSR of 0.347404.

Next, the number of thresholds for each model is determined. Visual inspection of Figure 6 indicates there are possibly three regimes in the price ratio series, but not likely to exceed three. However it is possible there are less than three regimes and to determine this, a sequential estimation of the number of thresholds and the associated threshold values is performed (Bai and Perron 1998).

To do so, each model is specified with the established delay parameters. The model, which minimizes the sum of squared residuals, is deemed the best fitting model for the RIN price ratio. Results indicate the TAR outperforms the SETAR with respective SSRs of 0.355751 and 0.379480. Therefore, the exogenous threshold variable of weekly ethanol mandate outperforms endogenous lagged values of the dependent variable and the TAR model is deemed the most appropriate.

The model of choice for D6/D4 RIN price ratio is

$$Y_t = \{(X'\delta_1 + \beta_{p_1,1}Y_{t-1} + \varepsilon_t)I_1(Z_{t-d} < r_1) + (X'\delta_2 + \beta_{p_2,2}Y_{t-1} + \varepsilon_t)I_2(r_1 \leq Z_{t-d})\}$$

$$X = \text{vector of exogenous variables}$$

$$Z_{t-d} = \text{logarithm of weekly ethanol mandates}$$

$$r_1 = \text{threshold values where}$$

$$I_m = \text{indicator variable equal to one if the argument is true, where } m = 1{:}\,2$$

$$Y_{t-1} = \text{lagged dependent variable}$$

The model is estimated assuming a lag order of one in each regime. The vector of exogenous variables includes ethanol blend margins, and biodiesel blend margins, which are both first differenced to ensure stationarity. Next, a smooth transition autoregressive model is applied to the RIN price ratio and fit will be compared to the TAR model.

1.6. Smooth Transition Autoregressive Model

The smooth transition autoregressive (STAR) model is represented by the equation

$$y_t = \boldsymbol{\alpha}'\boldsymbol{x}_t + \boldsymbol{\theta}'\boldsymbol{x}_t G(z_t;\gamma,c) + u_t$$

where, $\boldsymbol{x}_t = (1, y_{t-1}, \dots y_{t-p};\ x_{1t}, \dots, x_{kt})'$, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots \alpha_m)'$, $\boldsymbol{\theta} = (\theta_0, \boldsymbol{\theta}_1') = (\theta_0, \theta_1, \dots \theta_m)$ and $u_t \sim i.i.d(0, \sigma^2)$. The function $G(z_t;\gamma,c)$, identifies transition thresholds, z_t is a transition variable, γ and c are slope and location parameters respectively. The transition function is a continuous function between zero and one, assumed to take a logistic form of $G(z_t;\gamma,c) = [1 + \exp\{-\gamma(z_t - c)\}]^{-1}$ so that the model is also known as the logistic STAR or LSTAR. As z_t increases, the logistic function changes monotonically from zero to one. The STAR model with a logistic transition function is a regime-switching model, where the transition from one regime to the next can be smooth or abrupt, and the regime that occurs at time t is determined by z_t. The size of the slope parameter γ determines the speed at which the transition occurs. Large values indicate the regime change is quick and the STAR model nests the TAR model.

1.6.1 Linearity testing against the LSTAR model

Given a specific nonlinear alternative to the linear model, Lagrange Multiplier tests can be calculated which are optimal in terms of power against that nonlinearity (Granger and Terasvirta 1993). This test has power against both LSTAR and ESTAR models. The following steps can be taken to calculate these LM type tests.

1. Regress y_t on x_t and compute the residuals $\hat{u}_t = y_t - a'x_t$ and $SSR_0 = \sum \hat{u}_t^2$
2. If the transition variable is known to be z_{td} compute the auxiliary regression of $\hat{u}_t$=$\beta'\boldsymbol{x}_t + \sum_{j=1}^{p}(\varphi_{dj}z_{td}x_{tj} + \psi_{dj}z_{td}^2x_{tj} + \kappa_{dj}z_{td}^3x_{tj}) + v_t$ and obtain the SSR
3. Calculate the statistic $LM = \frac{T(SSR_0 - SSR)}{SSR_0}$

The null hypothesis is thus $\varphi_{dj} = \psi_{dj} = \kappa_{dj} = 0$ and the LM statistic has an asymptotic $\chi^2(3p)$ where p is the number of parameters. Performing the above test on D6/D4 RIN price ratio, results in LM-test statistics of 58.257 with a p-value of 0.0000, rejecting the null hypothesis of linearity against the alternative LSTAR nonlinear model.

1.6.2 Testing the significance of the transition variable

The auxiliary regression used in testing for linearity may be used to select the best transition variable and written more succinctly as

$$\hat{u}_t = \beta_0'x_t + \beta_1'x_tz_{td} + \beta_2'x_tz_{td}^2 + \beta_3'x_tz_{td}^3 + \eta_t \,. \qquad (1)$$

When the correct transition variable is selected, the auxiliary regression is indeed the appropriately specified auxiliary regression against the true nonlinear alternative. An incorrect transition variable would render it misspecified (Granger and Terasvirta 1993). If linearity is rejected for several candidate transition variables, then the transition variable with the smallest p-value or the largest test statistic is selected. In this procedure, the same candidate transition

variables used in TAR and SETAR models are considered here for the STAR models. Namely, this includes the logarithm of the weekly ethanol mandate (ln (ethmandate)), and lagged values of the dependent variable.

1.6.3 Estimation of the STAR Model

Estimation is performed with non-linear least squares using a Gauss-Newton optimization procedure with Marquardt iteration steps. The log-likelihood function is constructed under the assumption of normality and is thus represented as

$$\ln(l) = -\frac{1}{2}\ln(2\pi) - \ln(\sigma) - \frac{1}{2}\left(\frac{y_t - \boldsymbol{\alpha}'\boldsymbol{x}_t - \boldsymbol{\theta}'\boldsymbol{x}_t G(z_t;\gamma,c)}{\sigma}\right)^2$$

In the LSTAR model, the function $G(z_t;\gamma,c)$ is normalized in the exponential function by the standard deviation of the transition variable $\hat{\sigma}(z_t)$. Doing so rescales the slope parameter of the transition function γ since its value could be much larger than other parameters. To be clear the transition function takes the form

$$G(s_t;\gamma,c) = \ [1 + \exp\{-\gamma(z_t - c)/\hat{\sigma}(z_t)\}]^{-1}.$$

Based on Table 7, the LSTAR is estimated for the D6/D4 RIN price ratio while specifying the second lag of weekly ethanol obligations. Starting values for the optimization procedure are chosen randomly using a normal random number generator, which generates a vector of starting values using a random draw from a normal distribution with a mean of zero and variance of one. The model is specified with the same explanatory variables as the linear and TAR models so that a direct comparison of goodness of fit can be made.

Newey-West Heteroskedastic and autocorrelation corrected (HAC) standard errors are estimated to correct for any possible unspecified heteroskedasticity. Although heteroskedasticity was not

found under LM-type tests, the null hypothesis of homoscedasticity is rejected against an autoregressive conditional heteroskedasticity alternative. While ARCH does not invalidate least squares inference, there could be a loss of efficiency, which may lead to spurious rejection of a null hypothesis in later model diagnostics. Therefore, HAC standard errors are estimated in an attempt to correct for any loss of efficiency.

Table 8 reports the estimation results of the STAR model applied to the D6/D4 price ratio. A relatively large value for the smoothness parameter $\hat{\gamma}$ indicates the regime switch may be a quick transition rather than a smooth one. However, large values of $\hat{\gamma}$ are not of significant concern. As $\hat{\gamma} \rightarrow \infty$ the transition function $G(s_t; \gamma, c)$ becomes steep, which means the transition is fast. The threshold variable $\hat{c}$ is estimated to be 0.0854. To evaluate this threshold value, it is necessary to ensure it falls within the observed distribution of the transition variable $\ln(ethmandate)_{t-2}$. The minimum and maximum of the observed transition variable is -0.1421 and 0.1141 respectively, so the estimated threshold value falls towards the upper limits of the distribution, but well within the upper bound.

The linear model dominates when the transition function is equal to zero. During this regime, only the lagged dependent variable of $dratio_{t-1}$ is statistically significant. Ethanol blend margins and biodiesel blend margins are not significant in this regime. When the transition function takes on values greater than zero, the ethanol blend margin is significant at the 5% level and the biodiesel blend margin is highly significant at the 1% level. This indicates that a regime change, driven by the volume of weekly ethanol blended, results in a state of the world where blend margins are significant in explaining the value of the D6/D4 price ratio. In this state, the magnitude of the respective coefficients are partly determined by the value of the transition function. The sign of the coefficients in ethanol and biodiesel blend margins are in line with

expectations. As the ethanol blend margin increases, the core value of D6 ethanol RINs is expected to increase. Being in the numerator of the ratio, this naturally increases the value of the ratio. As biodiesel blend margins increase, the value of D4 RINs are expected to increase, thereby decreasing the value of the ratio.

Figure 10 indicates the transition function for the RIN price ratio is largely clustered around zero. However, it also indicates that the transition function takes on values between zero and one, illustrating the smoothness of the transition between regimes. The transition function can take on any value between zero and one. Therefore, the STAR model can be thought of as a continuum of regimes, represented by as many regimes, as unique values taken on by the transition function. Hence, the TAR model is nested and a special case of the STAR model. To compare the overall goodness of fit between the linear, TAR and STAR models, the SSR, R-squared and AIC are evaluated.

Table 9, indicates that both the STAR and TAR models provide a superior fit over the linear model. Based on the criteria it would appear that STAR models provide a better fit over TAR models for the D6/D4 price ratio. As noted, the TAR model is nested within the STAR model and thus it is not surprising that the overall goodness of fit is similar in magnitude.

1.6.4 STAR Model Diagnostics

In order to test the adequacy of the model, two tests are carried out. First, a test of remaining nonlinearity is conducted and second, a test of parameter constancy. These two tests aid in determining if the model is adequate in describing the behavior of the price ratio. Furthermore, the residuals of the model are inspected and tested for remaining autocorrelation. Eitrheim and Teräsvirta (1996), propose a LM-type test for remaining autocorrelation by estimating the auxiliary regression

$$\hat{u}_t = \pi \frac{\partial F(y_t)}{\partial \psi} + \sum_{i=1}^{q} \hat{u}_{t-i} + \nu_t \tag{1}$$

where $\frac{\partial F(y_t)}{\partial \psi}$ is the partial derivative of $\{\boldsymbol{\alpha}'\boldsymbol{x}_t + \boldsymbol{\theta}'\boldsymbol{x}_t G(z_t;\gamma,c)\}$ with respect to all parameters from the STAR specification. The LM-test statistic can be calculated as TR^2 with an asymptotic chi-squared distribution with degrees of freedom equal to the number of restrictions, which in this case is equal to q. The test of second-order serial correlation is carried out, resulting in a test-statistic of 6.948486 and a p-value of .030985. Given that the p-value is less than five percent, a test for third-order serial correlation is carried out. This test results in a LM-test statistic of 7.314276 and a resulting p-value of 0.065258 Therefore, results indicate there is no remaining autocorrelation in the residuals up to the third lag.

The test for remaining nonlinearity is performed in a similar manner. In this test, the lagged residuals from (1) are replaced by interacting the explanatory variables with the transition variable. Therefore, the auxiliary regression takes the form of

$$\hat{u}_t = \pi \frac{\partial F(y_t)}{\partial \psi} + \beta_1 x_t z_{t-d} + \beta_2 x_t {z^2}_{t-d} + \beta_3 x_t {z^3}_{t-d} + \nu_t \,. \tag{2}$$

The LM-test statistic is calculated so that $LM = TR^2$ with an asymptotic chi-squared distribution and degrees of freedom equal to $3p$ (Eitrheim and Teräsvirta 1996). Using this method, the LM-test statistic is calculated to be 22.09, which results in a p-value of 0.9667, indicating a failure to reject the null hypothesis of no remaining nonlinearity.

Finally, a test of parameter constancy is performed where the null hypothesis is that parameters do not vary over time. This test closely resembles the test for nonlinearity with the exception that z_{t-d} in equation (2) is replaced with a time-trend variable t (Dijk, Teräsvirta et al. 2002). Therefore, the auxiliary regression in the test for parameter constancy becomes

$$\hat{u}_t = \pi \frac{\partial F(y_t)}{\partial \psi} + \beta_1 x_t t + \beta_2 x_t t^2 + \beta_3 x_t t^3 + \nu_t \,. \tag{3}$$

The null hypothesis is that $\beta_1 = \beta_2 = \beta_3 = 0$ and the LM-test statistic can be calculated as TR^2 with an asymptotic chi-squared distribution and $3p$ degrees of freedom. Carrying out this test on the D6/D4 RIN price ratio, results in a test-statistic of 16.38 and a resulting p-value of 0.9979, strongly suggesting a failure to reject the null hypothesis that parameters do not vary over time.

Comparing in-sample fit and out-of-sample forecasts to actual values also indicates the STAR model is a suitable fit for the D6/D4 price ratio. Figure 11 compares the fitted values of the STAR model to actual values and provides a comparison against the linear model. Despite the improvement in goodness of fit for in-sample estimation, the STAR model performs poorly in out-of-sample forecasts, compared to the linear model. Out-of-sample forecasts result in a mean squared error of 0.0028 and 0.0010, for the STAR model and linear model respectively. The linear model also minimizes the mean absolute error with an MAE score of 0.0339, compared to an MAE of 0.0498 for the STAR model. However, this is perhaps not at all surprising. It is well documented in forecasting literature that linear models often outperform nonlinear models during out-of-sample forecasts (De Gooijer and Kumar 1992, Tiao and Tsay 1994, Brooks 1997, Clements and Hendry 1998, Stock and Watson 1998, Franses and Van Dijk 2000, Teräsvirta 2006). The linear model might provide a better forecast for a number of reasons, even though the nonlinear model more accurately describes the true nature of the data. For example, it is very possible that nonlinearity does not show up in the forecast period, despite being truly present in the estimation period. Alternatively, perhaps, a regime switch has occurred near the in-sample and out-of-sample boundary. Hence, out-of-sample forecasts provide a useful comparison against competing linear models, but should not necessarily be the benchmark to determine the most

appropriate fit. That determination is best made through visual inspection, testing for general nonlinearity, testing for specific nonlinearity, evaluation of nonlinear parameters, in-sample goodness of fit, and tests of misspecification. All of which have done here, and indicate the true data generating process behind the RIN price ratio is nonlinear.

1.6.5 An application to individual RIN prices

Applying a STAR model to the individual prices is used to determine if individual prices are nonlinear and this allows for the testing of the most significant transition variable in each price. The same processes of visual inspection, testing for general nonlinearity, testing for specific nonlinearity, evaluation of nonlinear parameters, in-sample goodness of fit, and tests of misspecification are applied to individual prices.

In each case, the optimal AR (p) process is chosen by scoring the information criteria, including the AIC score and the consistent ICOMP. In both cases of D6 and D4 RIN price, the optimal autoregressive process is found to be of lag-order one (Table 2). Therefore, a linear model is fitted to both RIN prices with right-hand side variables including a lagged dependent variable, the respective blend margins and a current period cross-price variable. To be clear, the explanatory variables for D6 RIN price are the first lag of D6 price, the weekly ethanol blend margins and the current period price of D4 RINs. The explanatory variables for D4 RIN price are the first lag of D4 price, the weekly biodiesel blend margins and the current period price of D6 RINs. Recall, the blend margins are composed of three explanatory variables including biofuel price, conventional fuel price and the value of blender tax credits. The blend margins thus represent the theoretical core value of a RIN's price. Results of the linear models are presented in Table 11 and Table 12.

Next, the two individual RIN prices are tested for linearity against the specific LSTAR alternative. Testing for nonlinearity and the most significant transition variable is done through specifying the auxiliary regression in equation (1). Performing the linearity test on D6 and D4, results in LM-test statistics of 46.044 and 43.231 for p-values of 0.0000 in both cases, rejecting the null hypothesis of linearity against the alternative LSTAR nonlinear model.

Several candidate transition variables are proposed and the procedure is carried out iteratively. The candidate variable which is significantly associated with regime change is chosen to be the transition variable which best determines the switching behavior. If all candidate variables are statistically significant, then the variable with the smallest p-value or the largest test statistic is chosen as the most appropriate transition variable. While carrying out this test for D4 RIN price, the candidate variables are own lagged price, lagged price of D6 RINs and weekly ethanol mandates. Results indicate the own lagged price of D4 RIN with a delay parameter of one is the most significant transition variable for D4 RIN prices (Table 14). All candidate transition variables are significant with a p-value of 0.0000. However, the first lagged own price of D4 RINs is the most significant as indicated with the largest F-statistic of 63.032. This result is in line with expectations for D4 RIN price regime change. When the ethanol blend wall is binding, obligated parties can circumvent the blend wall by over complying with the biodiesel mandate, thus generating excess D4 RINs, which can then be submitted towards meeting the implied ethanol mandate. Therefore, weekly ethanol mandates are thought to trigger regime change in D6 RIN prices, and the ratio of D6/D4 RIN price. However, the most appropriate transition variable for D4 RIN price is expected to be endogenous lagged own prices.

Identifying the most significant transition variable in D6 RIN price is carried out in the same manner. In this case, the most significant transition variable is expected to be exogenous and

specifically thought to be the weekly ethanol mandate. However, candidate variables are also tested including lagged price of D4 RINs and the endogenous lagged own price of D6 RINs. The test is carried out by iteration, just as is the case with D4 RINs and the RIN price ratio. Test results indicate the logarithm of weekly ethanol mandates, with a delay parameter of four, is the most significant transition variable for D6 RIN prices with an F-statistic of 68.316. The next most significant transition variable is the endogenous own price of D6 RIN with a delay parameter of three and an F-statistic of 55.328 while the first lag of D4 RIN price is found to have an F-statistic of 31.056 (Table 15). To summarize, regime change in the price ratio is best determined by the weekly ethanol mandate with a delay parameter of two lags. Regime change in D4 RIN price is best determined endogenously by D4 RIN price with a delay parameter of one lag. The regime in D6 RIN price is best determined by the weekly ethanol mandate with a delay parameter of four lags.

With strong evidence of nonlinearity in the individual RIN prices and the most significant transition variables determined, a STAR model is now fitted to each of the individual D6 and D4 RIN prices. Estimation is performed with non-linear least squares while specifying the transition function to be of the logistic form just as in the case of the price ratio. Estimation results of the D4 STAR model are presented in Table 16. The threshold variable $\hat{c}$ is estimated to be -0.097921, which is within the range of the transition variable $\Delta lnD4_{t-1}$. The values of the transition function are plotted, to evaluate the transition function and determine if the regime switch occurs smoothly or rapidly.

Figure 12, indicates the transition function for D4 STAR model takes on many values and the regime changes occur smoothly. While a SETAR model is not explicitly ruled out, it does not appear the transition occurs abruptly. In any case, the SETAR model is nested within and is a

special case of the STAR model. Table 17 compares the goodness of fit for the linear model and STAR model. Based on an examination of the criteria in this table, the nonlinear model provides a better fit for D4 RIN price.

Model diagnostics are performed for the D4 STAR model to test for remaining autocorrelation in the residuals, any remaining nonlinearity and the data is tested for parameter constancy. The tests are carried out in the same manner as in section 1.6.4 by specifying the auxiliary regressions of equations (1), (2), and (3). When testing for any remaining nonlinearity, test results indicate a failure to reject the null hypothesis of no remaining nonlinearity with a p-value of 0.4001. The test for parameter constancy is carried out by specifying the null hypothesis of no time varying parameters and results in a p-value of 0.6048, indicating a failure to reject the null hypothesis. Results strongly suggest the STAR model is an adequate fit for D4 RIN prices in terms of remaining linearity and parameter constancy. However, test results indicate there may be evidence of remaining autocorrelation. A p-value of 0.0495 indicates a failure to reject the null hypothesis of no autocorrelation up to the third lag, at the 99% confidence level. However, the null is rejected at the 95% confidence level (Table 10). Despite this, the model provides a good fit for in-sample observations and out-of-sample forecasts (Figure 13). In comparison with the linear model, the STAR model does not provide a large improvement when minimizing the mean absolute error and the mean squared error. Out-of-sample forecasts for the STAR model result in a MAE of 0.0266 while the linear model results in a MAE of 0.0249. Evaluating the MSE also indicates the STAR model shows little improvement over the linear model with MSE scores of 0.0012 and 0.0013. While the STAR model may provide a superior fit for in-sample observations, out-of-sample forecasts are less promising.

Evaluation of the D6 STAR model is carried out in the same way as with the D4 STAR model. First, the threshold parameter is checked to ensure the estimated value falls in the range of actual values. With an estimated value of 0.040806, the threshold parameter does indeed fall in the range of the minimum and maximum values of $\ln(ethmandate)_{t-4}$. Next, the transition function is plotted to examine the distribution and smoothness. Figure 14 illustrates the transition is not quite as smooth as exhibited in D4 RIN price. This is also evidenced by the relatively large value of gamma in the STAR model estimation results (Table 18). Recall, that as gamma goes to infinity, the transition becomes abrupt and the value of the transition function will only take on two values, zero and one. For D6 RIN price, the regime change is indeed a smooth transition, albeit not as smooth as in D4 RIN price.

Routine diagnostics are performed for the D6 STAR model, testing for parameter constancy and any remaining autocorrelation and nonlinearity. In all three cases, test results indicate a failure to reject the null hypotheses, thus concluding that the STAR model is adequate (Table 10). The goodness of fit is evaluated by comparing the sum of squared residuals, R-squared and AIC. The STAR model outperforms the linear model while comparing SSR and R-squared. However, AIC is minimized for the linear model with an AIC of -0.7916, compared to an AIC of -0.7450 in the STAR model. This indicates that while the STAR model provides a better fit, it also adds more complexity than is perhaps necessary. However, the D6 STAR model significantly outperforms the linear model during out-of-sample forecasts. The MAE, MSE and mean absolute percentage error (MAPE) are minimized under the STAR model. The STAR model results in a MAE, MSE and MAPE of 0.0003, 0.0097, and 99.21% respectively. The same scores for the linear model are 0.0020, 0.0287, and 296.02%. This improvement in out-of-sample prediction can be visually observed in the plot of fitted values for in-sample and out-of-sample forecasts in Figure 15.

1.7. Conclusion

The market for renewable identification numbers (RINs) is complex and constantly evolving. Understanding the behavior of the RINs can provide important insights into an industry, which may be critical to long term energy independence. Federal regulation, which determines annual renewable volume obligations, is often uncertain year to year, causing market participants, obligated parties and other agents to adapt their behaviors to the changing regulations. It is the behavior of these agents, which is reflected in the multiple regimes, and structural breaks of the RIN price.

Nonlinearities in the RIN market are examined by fitting threshold autoregressive and smooth transition autoregressive models to historical RIN data. Theoretical RIN price is primarily driven by the core value of the RIN, which is determined by the gap between the price of the biofuel and conventional fuel, also known as the blend margin. Therefore, ethanol and biodiesel blend margin variables are introduced as exogenous variables.

It is believed that coupling patterns observed in the RIN market are brought about by some proximity to the ethanol blend wall. In other words, there is some unobservable threshold, that when crossed, brings about a change in the behavior of the RIN market. Therefore, a key variable in this paper is the implied weekly ethanol mandate. This variable is constructed by taking the product of the fractional ethanol mandate and weekly gasoline and diesel supplies. This variable approximates the volume of ethanol blended on a weekly basis, excluding cellulosic ethanol and other advanced ethanol us and is found to be the most significant transition variable for D6 RIN price regime changes and regime change in the D6/D4 price ratio.

In regards to D4 RIN price, past values of D4 RIN price are found to be the most significant variable in determining regime change, indicating the process is self-exciting. This is in line with

expectations. The ethanol blend wall represents a challenge to producers and importers, which are obligated to meet annual blend mandates. Therefore, obligated parties are over complying with the biodiesel mandate to generate D4 biodiesel RINs that can be submitted for compliance with the overall mandate.

RIN prices are found to exhibit nonlinearity and regime changing behavior. STAR models are capable of taking on smooth transitions and abrupt transitions, making them well suited to model RIN prices.

1. A. Appendix

1. A.1. Tables

Table 1: RIN type by D-code and production pathway

Fuel (D Code)	Reduction in GHG emissions	Equivalence Value	Fuel Type
Cellulosic Biofuel (D3)	60%	1.7	Cellulosic ethanol, Renewable CNG/LNG, Naphtha, Renewable Gasoline
Cellulosic Diesel (D7)	60%	1.7	Cellulosic diesel, Renewable Jet Fuel
Biomass-Based Diesel (D4)	50%	1.5	Biodiesel, Renewable Diesel, Jet Fuels, etc.
Advanced Biofuel (D5)	50%	1.3	Sugarcane ethanol, Renewable Heating oil, etc.
Renewable Fuel (D6)	20%	1	Corn ethanol, etc.

Table 2: Optimal lag lengths of price ratio and individual prices

	Ratio of D6/D4		D4 RINs		D6 RINs	
LAG	AIC	CICOMP	AIC	CICOMP	AIC	CICOMP
1	-1267.46	-1250.34	-4920.12	-4903.60	-1267.46	-1250.34
2	-1265.54	-1243.56	-4918.65	-4895.81	-1265.54	-1243.56
3	-1264.36	-1237.53	-4925.84	-4896.66	-1264.36	-1237.53
4	-1262.38	-1230.68	-4925.37	-4889.86	-1262.38	-1230.68
5	-1263.18	-1226.64	-4924.45	-4882.60	-1263.18	-1226.64
6	-1263.17	-1221.77	-4930.00	-4881.82	-1263.17	-1221.77
7	-1264.27	-1218.01	-4928.84	-4874.33	-1264.27	-1218.01

Table 3: Test of serial correlation in D6/D4 RIN price

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.466909	Prob. F(1,300)	0.6274
Obs*R-squared	0.946436	Prob. Chi-Square(2)	0.6230

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Included observations: 305

Table 4: Test of heteroskedasticity in D6/D4 RIN price ratio

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0.762505	Prob. F(3,301)	0.5159
Obs*R-squared	2.300430	Prob. Chi-Square(3)	0.5124
Scaled explained SS	17.24378	Prob. Chi-Square(3)	0.0006

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Included observations: 305

Table 5: Tsay and Keenan tests for nonlinearity in D6/D4 price ratio

First-differenced Ratio			
Test	AR(p)	F-stat	p-value
Tsay	2	3.667	0.0126
Keenan	2	3.060	0.0811

Table 6: Linear model of D6/D4 RIN price ratio

Dependent Variable: DRATIO
Method: Least Squares
Included observations: 305

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001167	0.002097	0.556537	0.5783
DRATIO(-1)	0.289536	0.055945	5.175386	0.0000
DETHMRGN	-0.002343	0.016280	-0.143932	0.8857
DBDMRGN	-0.041172	0.016989	-2.423412	0.0160

R-squared	0.096347	Mean dependent var	0.001853
Adjusted R-squared	0.087340	S.D. dependent var	0.038250
S.E. of regression	0.036541	Akaike info criterion	-3.767712
Sum squared resid	0.401918	Schwarz criterion	-3.718921
Log likelihood	578.5761	Hannan-Quinn criter.	-3.748197
F-statistic	10.69744	Durbin-Watson stat	1.870313
Prob(F-statistic)	0.000001		

Table 7: Test for most significant transition variable of D6/D4 price ratio

Transition variable tests for D6/D4 price ratio
H0: The transition variable is not significant

Transition variable (STR)	F-statistic	P-value	Transition variable (STR)	F-statistic	P-value
ln(ethmandate) (-1)	11.250	0.000	dratio(-1)	43.420	0.000
ln(ethmandate) (-2)	62.432	0.000	dratio(-2)	19.547	0.000
ln(ethmandate) (-3)	26.188	0.000	dratio(-3)	49.249	0.000
ln(ethmandate) (-4)	36.263	0.000	dratio(-4)	27.493	0.000
ln(ethmandate) (-5)	6.854	0.000	dratio(-5)	24.132	0.000

Table 8: D6/D4 Price ratio STAR estimation results

Dependent Variable: DRATIO
Method: Least Squares (Gauss-Newton / Marquardt steps)
Included observations: 298 after adjustments
HAC standard errors & covariance using observed Hessian (Bartlett kernel, Newey-West fixed bandwidth = 6.0000)

	Coefficient	Std. Error	t-Statistic	Prob.
CONSTANT	0.001871	0.002086	0.897097	0.3704
DRATIO(-1)	0.168714	0.092607	1.821836	0.0695
DETHMRGN	-0.007861	0.011365	-0.691614	0.4897
DBDMRGN	-0.014684	0.018220	-0.805907	0.4210
CONSTANT’	-0.016779	0.018797	-0.892617	0.3728
DRATIO(-1)’	3.494861	0.932627	3.747331	0.0002
DETHMRGN’	0.330626	0.160887	2.055019	0.0408
DBDMRGN’	-0.237419	0.088252	-2.690237	0.0076
GAMMA	3.373670	1.349258	2.500389	0.0130
THRESHOLD	0.085397	0.005899	14.47766	0.0000

R-squared	0.222721	Mean dependent var	0.001813
Adjusted R-squared	0.198431	S.D. dependent var	0.038679
S.E. of regression	0.034630	Akaike info criterion	-3.855238
Sum squared resid	0.345371	Schwarz criterion	-3.731174
Log likelihood	584.4305	Hannan-Quinn criter.	-3.805576
F-statistic	9.169255	Durbin-Watson stat	1.859915
Prob(F-statistic)	0.000000		

Table 9: Linear, TAR and STAR model comparison for D6/D4 RIN price ratio

Linear model versus TAR and STAR model for D6/D4 RIN price ratio					
Linear Model		TAR		STAR	
R-squared	0.095417	R-squared	0.199361	R-squared	0.222721
SSR	0.402332	SSR	0.355751	SSR	0.345371
AIC	-3.773241	AIC	-3.839051	AIC	-3.855238

Table 10: Test results of model diagnostics for RIN price ratio and individual RIN price

Model Diagnostics for D6/D4 Price ratio and Individual RIN Price			
Test	D6/D4 Ratio	D4	D6
Remaining Autocorrelation (3rd-order)	7.314276 (0.062528)	7.833062 (0.049591)	6.104918 (0.106616)
Remaining Nonlinearity	22.09043 (0.966723)	28.21173 (0.400120)	49.09835 (0.071486)
Parameter Constancy	16.38977 (0.997944)	33.15026 (0.604858)	25.12394 (0.912921)

p-values in parenthesis

Table 11: Linear model of D4 RIN price

Dependent Variable: LRD4
Method: Least Squares
Included observations: 278

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.002357	0.005044	0.467340	0.6406
LRD4(-1)	0.186963	0.053949	3.465526	0.0006
DBDMRGN	0.011621	0.043803	0.265311	0.7910
LRD6	0.169851	0.028613	5.936199	0.0000

R-squared	0.149724	Mean dependent var	0.005477
Adjusted R-squared	0.140414	S.D. dependent var	0.090343
S.E. of regression	0.083760	Akaike info criterion	-2.107427
Sum squared resid	1.922334	Schwarz criterion	-2.055231
Log likelihood	296.9324	Hannan-Quinn criter.	-2.086486
F-statistic	16.08272	Durbin-Watson stat	1.846363
Prob(F-statistic)	0.000000		

Table 12: Linear model of D6 RIN price

Dependent Variable: LRD6
Method: Least Squares
Included observations: 276

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007532	0.009776	0.770487	0.4417
LRD6(-1)	0.127664	0.051907	2.459478	0.0145
DETHMRGN	-0.069348	0.074549	-0.930230	0.3531
LRD4	0.611461	0.108522	5.634451	0.0000

R-squared	0.136228	Mean dependent var	0.012578
Adjusted R-squared	0.126701	S.D. dependent var	0.173044
S.E. of regression	0.161711	Akaike info criterion	-0.791630
Sum squared resid	7.112886	Schwarz criterion	-0.739161
Log likelihood	113.2450	Hannan-Quinn criter.	-0.770575
F-statistic	14.29932	Durbin-Watson stat	1.911318
Prob(F-statistic)	0.000000		

Table 13: Tsay and Keenan tests for nonlinearity of Individual RIN prices

Log-differenced D4				Log-differenced D6			
Test	AR(p)*	F-stat	p-value	Test	AR(p)*	F-stat	p-value
Tsay	1	6.98	0.0086	Tsay	1	8.0260	0.0049
Keenan	1	6.22	0.0132	Keenan	1	7.8965	0.0053

*AR (p) lag orders are estimated by fitting an autoregressive model, choosing the lag order, which minimizes Akaike's information criteria (AIC).

Table 14: Tests of most significant transition variable for D4 RIN price

Transition variable tests for D4 RINs
H0: The transition variable is not significant

Transition variable	F-statistic	Transition variable	F-statistic	Transition variable	F-statistic
ln(ethmandate) (-1)	33.514	LRD6(-1)	18.167	LRD4(-1)	63.032^{++}
ln(ethmandate) (-2)	28.779	LRD6(-2)	6.938	LRD4(-2)	27.403
ln(ethmandate) (-3)	17.021	LRD6(-3)	12.099	LRD4(-3)	48.390
ln(ethmandate) (-4)	53.035^{++}	LRD6(-4)	16.162	LRD4(-4)	24.156
ln(ethmandate) (-5)	28.019	LRD6(-5)	22.298^{++}	LRD4(-5)	16.236

$^{++}$Indicates the largest test statistic for each candidate transition variable

Table 15: Tests of most significant transition variable for D6 RIN price

Transition variable tests for D6 RINs
H0: The transition variable is not significant

Transition variable	F-statistic	Transition variable	F-statistic	Transition variable	F-statistic
ln(ethmandate) (-1)	11.623	LRD6(-1)	19.671	LRD4(-1)	31.056^{++}
ln(ethmandate) (-2)	17.534	LRD6(-2)	15.486	LRD4(-2)	11.854
ln(ethmandate) (-3)	26.876	LRD6(-3)	55.328^{++}	LRD4(-3)	4.268
ln(ethmandate) (-4)	68.316^{++}	LRD6(-4)	20.829	LRD4(-4)	11.696
ln(ethmandate) (-5)	15.622	LRD6(-5)	34.132	LRD4(-5)	13.351

$^{++}$Indicates the largest test statistic for each candidate transition variable

Table 16: D4 STAR estimation results

Dependent Variable: LRD4
Method: Least Squares (Gauss-Newton / Marquardt steps)
Included observations: 278

	Coefficient	Std. Error	t-Statistic	Prob.
CONSTANT	-0.187345	0.102947	-1.819825	0.0699
LRD4(-1)	-0.546776	0.304523	-1.795513	0.0737
DBDMRGN	0.487435	0.264419	1.843418	0.0664
LRD6	0.037521	0.141266	0.265605	0.7907
CONSTANT’	0.223762	0.125983	1.776128	0.0768
LRD4(-1)’	0.542918	0.306883	1.769137	0.0780
DBDMRGN’	-0.589890	0.299760	-1.967878	0.0501
LRD6’	0.128052	0.162057	0.790166	0.4301
GAMMA	1.677048	0.989303	1.695181	0.0912
THRESHOLD	-0.097921	0.048294	-2.027586	0.0436

R-squared	0.211009	Mean dependent var	0.005477
Adjusted R-squared	0.184513	S.D. dependent var	0.090343
S.E. of regression	0.081584	Akaike info criterion	-2.139068
Sum squared resid	1.783778	Schwarz criterion	-2.008578
Log likelihood	307.3304	Hannan-Quinn criter.	-2.086716
F-statistic	7.963811	Durbin-Watson stat	1.988792
Prob(F-statistic)	0.000000		

Table 17: D4 linear and STAR model comparison

Linear model versus STAR model for D4 RIN price			
Linear Model		STAR Model	
R-squared	0.149046	R-squared	0.211009
SSR	1.923866	SSR	1.783778
AIC	-2.113824	AIC	-2.139068

Table 18: D6 STAR estimation results

Dependent Variable: LRD6
Method: Least Squares (Gauss-Newton / Marquardt steps)
Included observations: 271 after adjustments

	Coefficient	Std. Error	t-Statistic	Prob.
CONSTANT	0.008584	0.014224	0.603507	0.5467
LRD6(-1)	0.135789	0.099643	1.362748	0.1741
DETHMRGN	-0.062937	0.071172	-0.884292	0.3774
LRD4	0.500049	0.299988	1.666896	0.0967
CONSTANT’	-0.003813	0.023466	-0.162481	0.8711
LRD6’	-0.106844	0.125396	-0.852057	0.3950
DETHMRGN’	-0.071836	0.155464	-0.462076	0.6444
LRD4’	0.622346	0.316253	1.967878	0.0501
GAMMA	113.6799	420.2773	0.270488	0.7870
THRESHOLD	0.040806	0.012567	3.247034	0.0013

R-squared	0.150090	Mean dependent var	0.012770
Adjusted R-squared	0.120783	S.D. dependent var	0.174606
S.E. of regression	0.163722	Akaike info criterion	-0.745096
Sum squared resid	6.996049	Schwarz criterion	-0.612176
Log likelihood	110.9604	Hannan-Quinn criter.	-0.691727
F-statistic	5.121266	Durbin-Watson stat	1.917909
Prob(F-statistic)	0.000002		

Table 19: D6 linear and STAR model comparison

Linear model versus STAR model for D6 RIN price			
Linear Model		STAR Model	
R-squared	0.136228	R-squared	0.150090
SSR	7.112886	SSR	6.996049
AIC	-0.791630	AIC	-0.745096

1. A.2 Figures

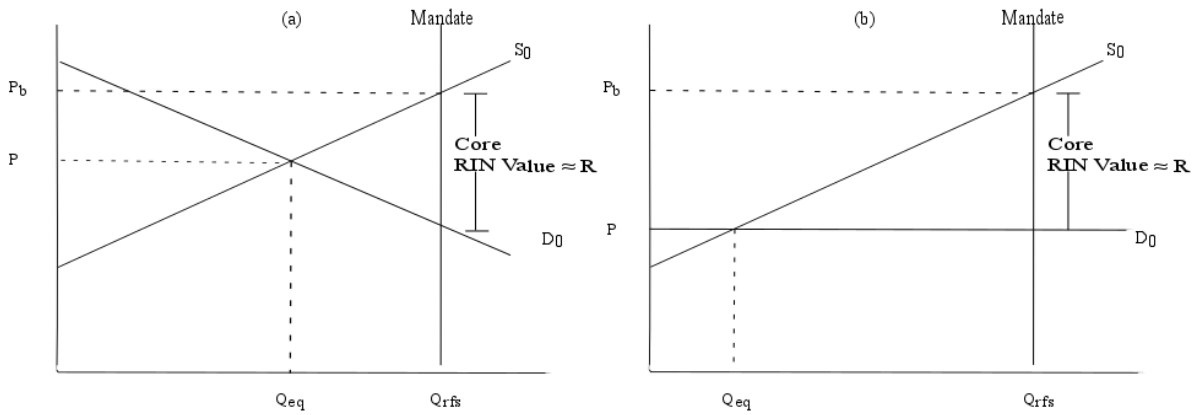

Figure 1: Core value of a RIN in ethanol (a) and biodiesel markets (b)

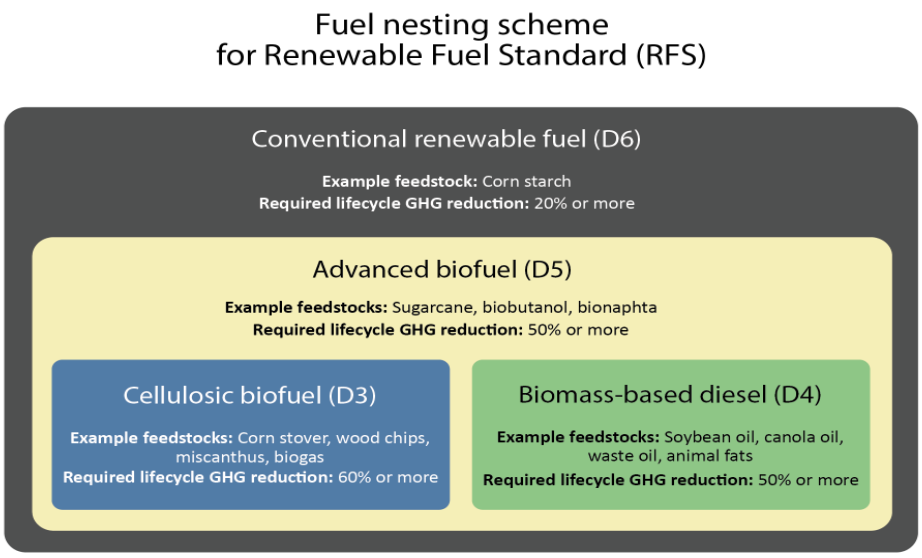

Figure 2: Nested structure of RINs (U.S. Environmental Protection Agency 2016)

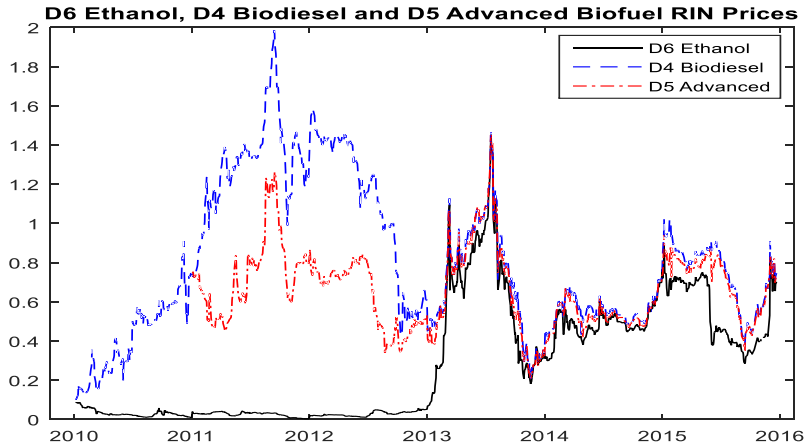

Figure 3: Historical RIN prices

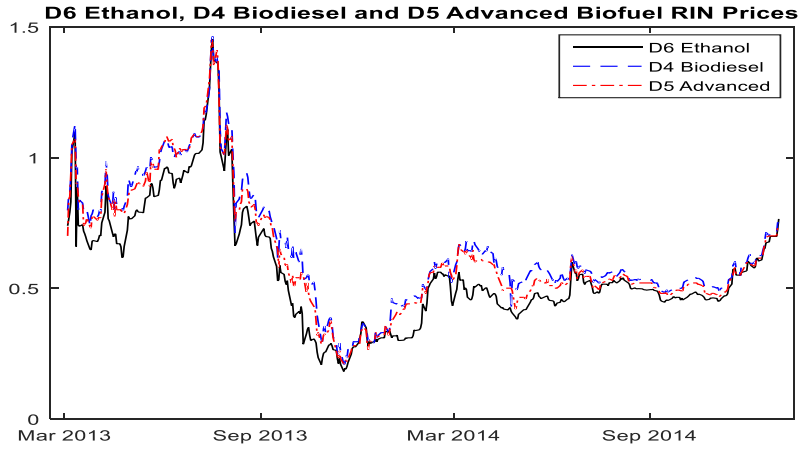

Figure 4: Coupled RIN prices

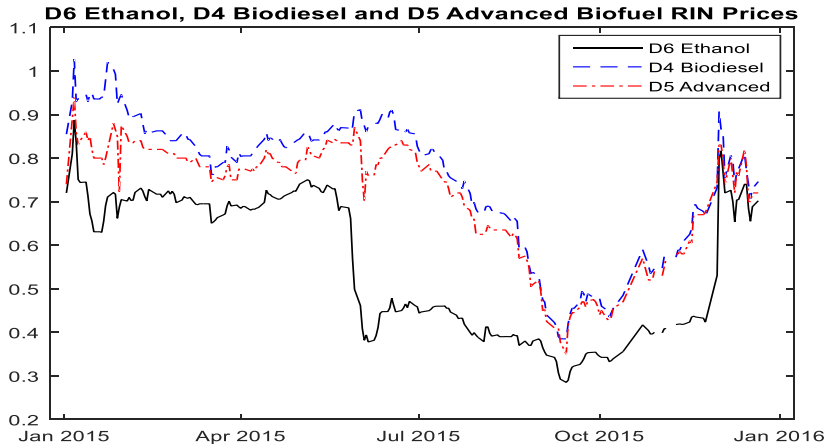

Figure 5: Decoupled RIN prices

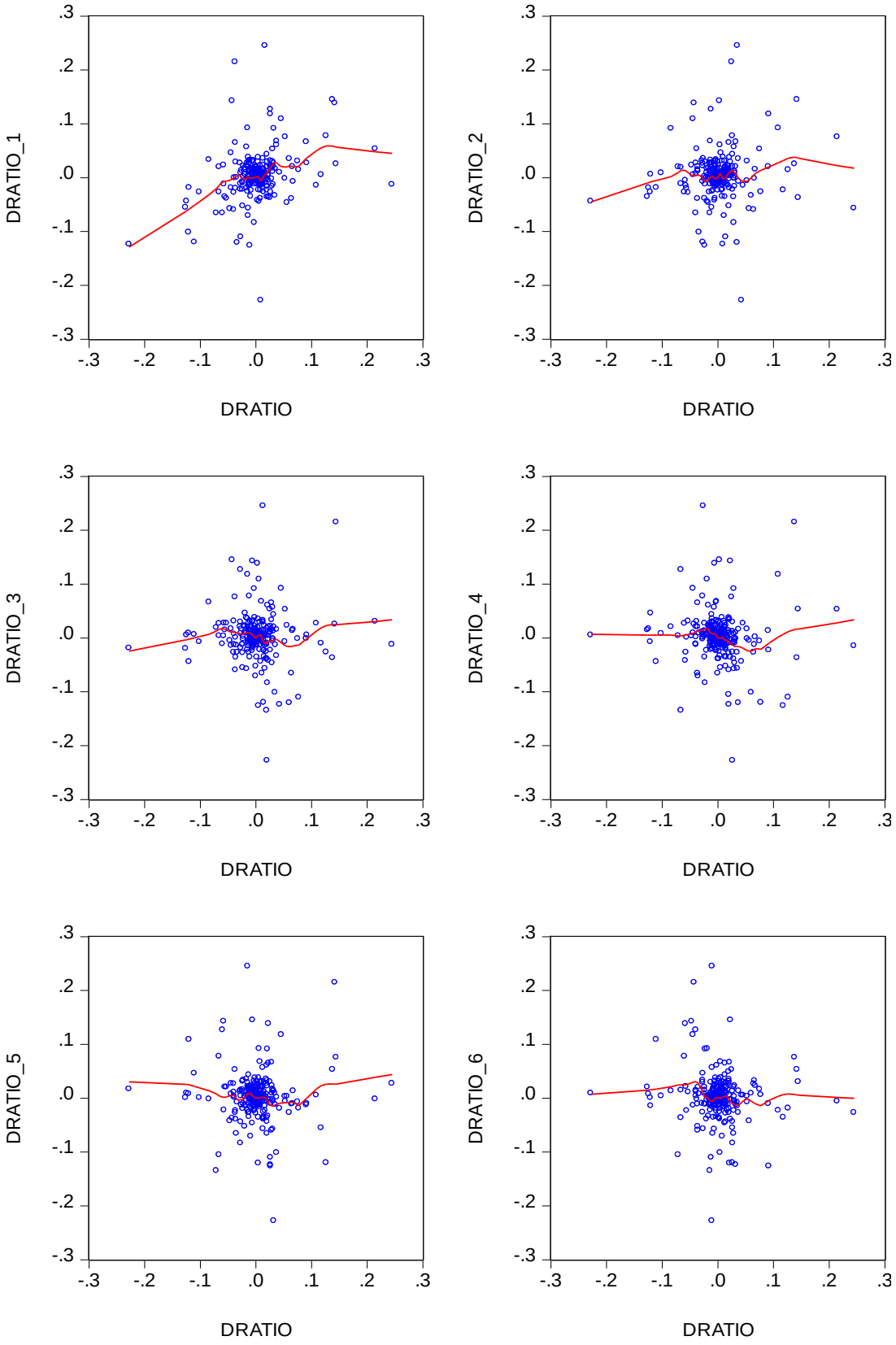

Figure 6: First-differenced Ratio of D6/D4 lag plots

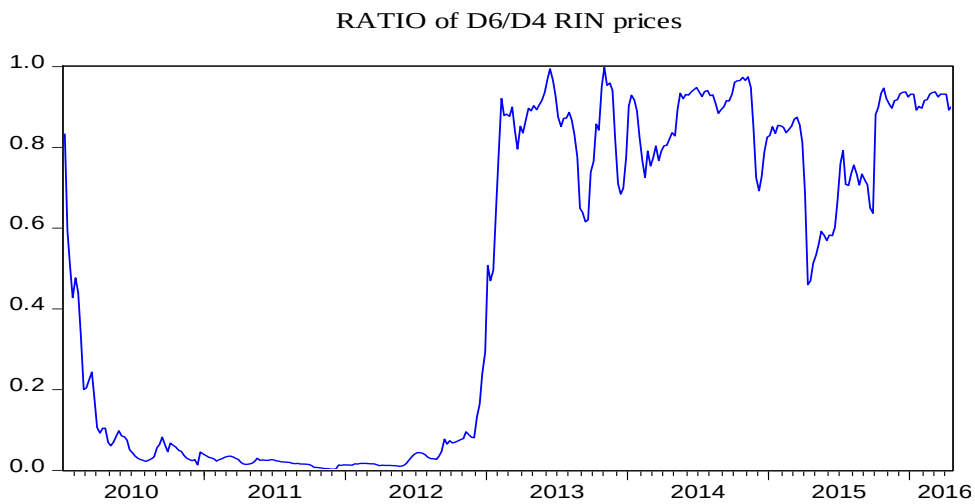

Figure 7: Historical Ratio of D6/D4 RIN prices

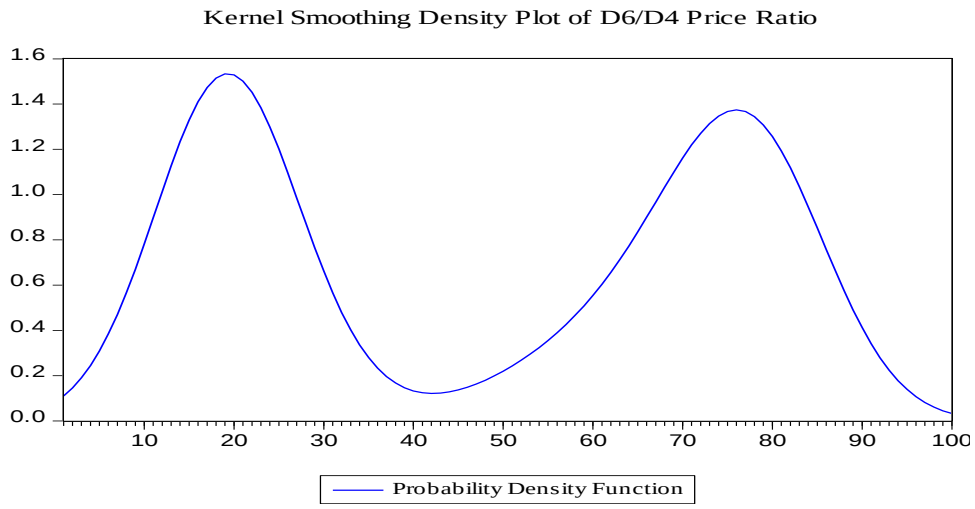

Figure 8: Kernel density plot to illustrate bi-modal distribution

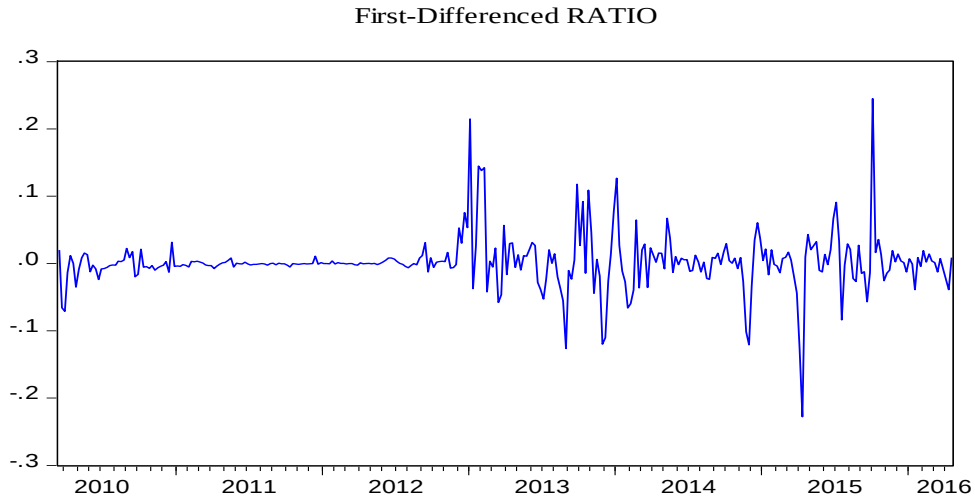

Figure 9: First-differenced ratio D6/D4 RIN prices

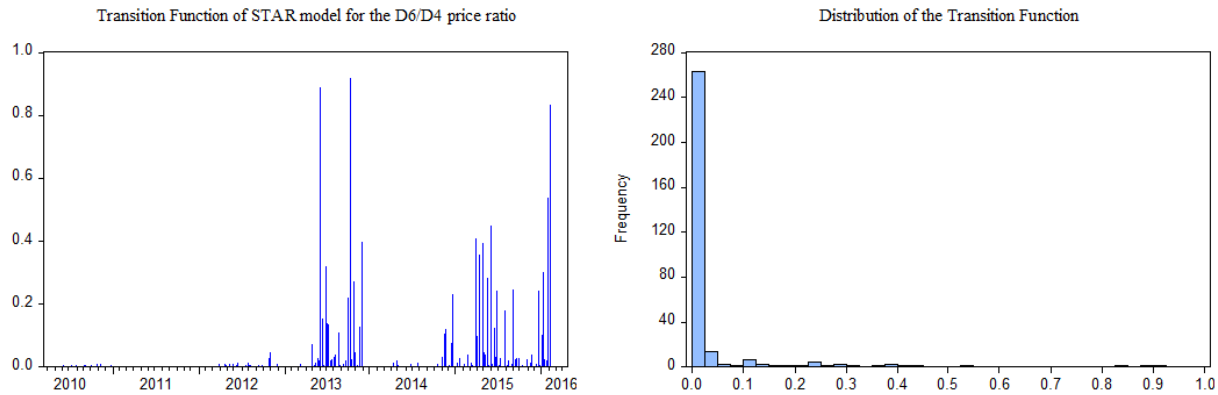

Figure 10: Evaluation of the transition function for D6/D4 RIN price ratio

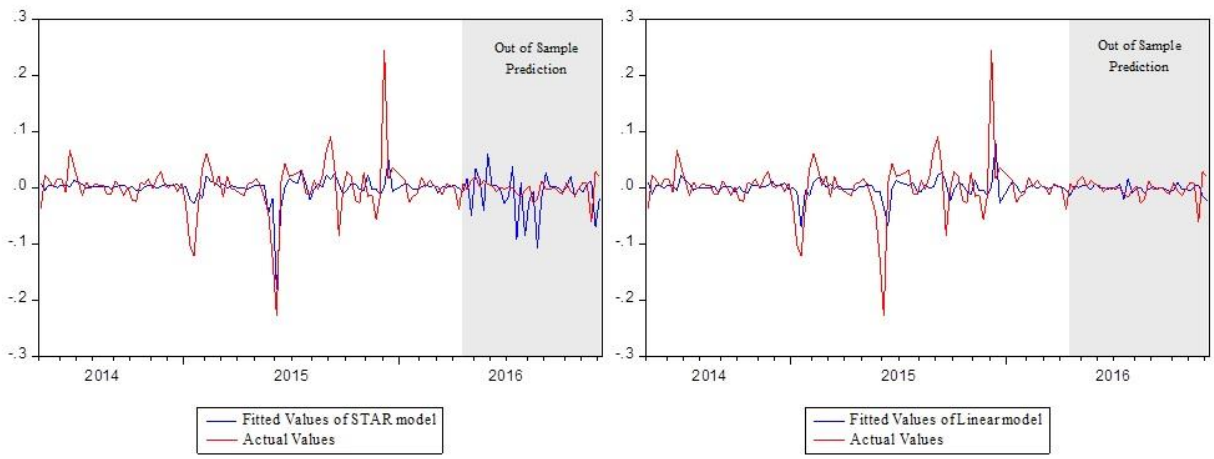

Figure 11: Fitted values compared to actual values of $\Delta(D6/D4)$ for the STAR model and the linear model

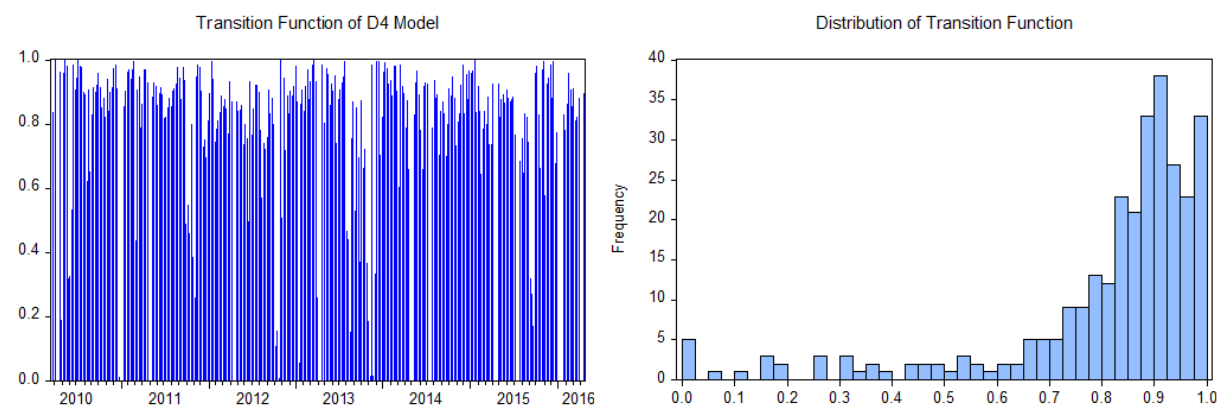

Figure 12: Evaluation of D4 Transition Function

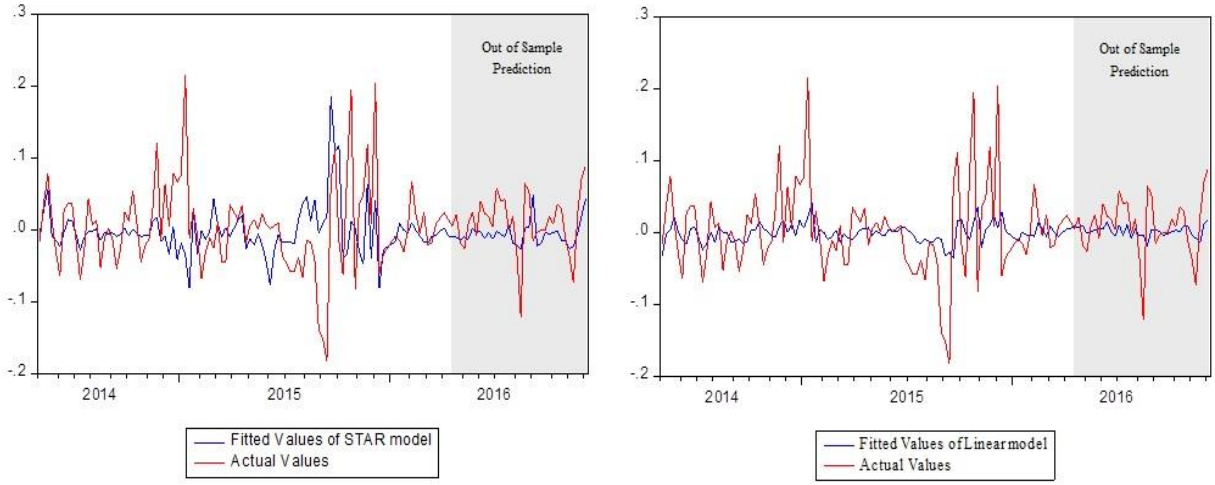

Figure 13: Fitted values compared to actual values of $\Delta ln(D4)$ price for the STAR model and the linear model

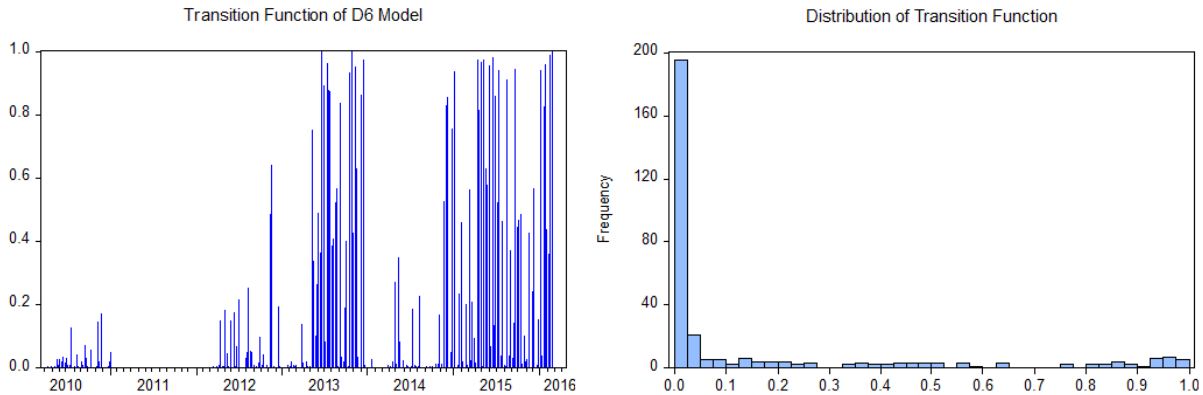

Figure 14: Evaluation of D6 Transition Function

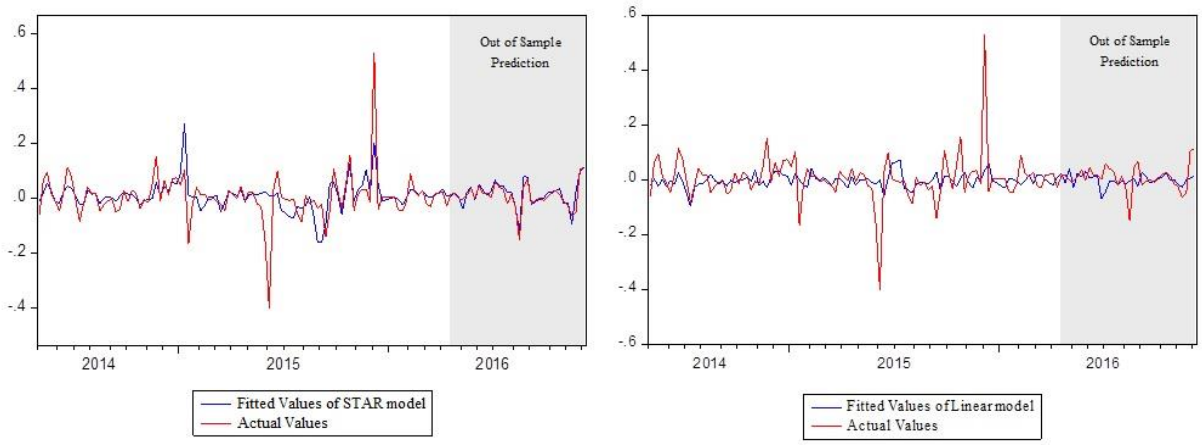

Figure 15: Fitted values compared to actual values of $\Delta ln(D6)$ price for the STAR model and the linear model

Chapter II: Renewable Fuel Standard Uncertainty and Tax Incentive Effects in the Optimal Investment Decisions of Second-Generation Biofuel Producers

Chapter II: Abstract

Investment in the production of advanced biofuel has been encouraged by the federal regulation, which enacted the Renewable Fuel Standard (RFS). However, investments and production of advanced biofuel has been lower than expected in recent years, leading regulators to reduce required volumes of advanced biofuel (U.S. Environmental Protection Agency 2015).

Real option analysis is used to help explain why uncertainty has reduced investment in advanced biofuel and eroded the requirements put forth by USEPA. This study examines the effect of uncertainty and irreversibility on a second-generation biofuel investment while considering the volatility of Renewable Identification Number (RIN) price, conventional fuel price and their correlations by applying a two stochastic variable approach. Producers of renewable fuel face an output price, which can be decomposed into the price of conventional fuel, plus the price of applicable RINs. Therefore, the firm is subject to uncertainty in both markets. Results have important policy implications, as uncertainties in the RIN market, created to encourage biofuel production, may be counteracting intended policy design.

2.1. Introduction

The Renewable Fuel Standard (RFS) was enacted to encourage investment in the production of advanced biofuel. However, production has been lower than expected in recent years. One potential explanation is that investors are hesitant to enter the biofuel market due to the unpredictability of conventional fuel markets which make the revenues and costs associated with biofuel production uncertain (Schmit, Luo, and Conrad 2011; McCarty and Sesmero 2014). However, uncertainty in the RFS itself has also been cited as one reason biofuel production has fallen short of expectations (Stock 2015). Regulators annually revise the RFS based on production capacity and projected demand. The annual announcement itself creates inherent uncertainty in the market. However, annual announcements have been delayed for up to 18 months in some years. In addition, regulators unexpectedly reduce required volumes of advanced biofuel when production is lower than expected (U.S. Environmental Protection Agency 2015). Blender tax credits provide another source of policy uncertainty (Irwin 2016). Blenders have been eligible for tax credits since 2005. However, policymakers have unexpectedly let these policies expire, only to subsequently reinstatement them some years later. These unexpected adjustments to biofuel policy add a second source of uncertainty to the biofuel-market entry decision. This paper represents the first attempt to isolate the effect of market and policy uncertainty in the decision to invest in biofuel production.

Potential entrants to the biofuel market must make irreversible investments in exchange for an uncertain stream of profits. This combination of uncertainty and irreversibility generate an incentive for a potential entrant to delay market entry to wait for updated information (Dixit 1989). This incentive is known as an option value and leads to prices, which trigger entry (exit) that are greater (less) than would be expected using traditional discounted cash flow methods.

Schmit, Luo et al. (2011), solve for optimal entry and exit trigger prices for a firm producing ethanol from corn, a first-generation biofuel. First-generation biofuels are produced from sugars such as those found in corn or sugar cane. The authors use a two-variable model to account for uncertainty in biofuel demand and operating costs. Their results indicate the two-variable model outperforms a one variable model in predicting industry expansion by including the individual components and their correlations as opposed to a simpler measure of aggregate uncertainty.

McCarty and Sesmero (2014), have conducted a real options analysis on the effect of conventional fuel price uncertainty in the investment of second-generation biofuel production.[11] Due to limitations in price data for second-generation biofuels, the authors assume prices will follow those of the gasoline market. The authors contribute to the literature by applying the methodology to a specific production pathway, which results in a second-generation biofuel known as bio-gasoline. Their results indicate trigger prices to be approximately 72% higher than traditional break-even prices.

Potential entrants to the market for second-generation biofuels are influenced by the observed price of both conventional fuel and the price of tradable credits, known as a renewable identification number. RINs trade on the open market and provide price support in the production of renewable fuels.[12] The price of the RIN will be reflected in the price of the biofuel.

The uncertain nature of RFS policy materializes in highly volatile RIN prices. The U.S. Environmental Protection Agency (EPA) has the authority to revise the annual RFS based on

---

[11] When originally conceived, the RFS targeted second-generation biofuels to account for 16 billion of the total 36-billion-gallon goal for 2022. Second generation biofuels are produced from non-food feedstocks, and thus do not have direct impacts on commodity food prices and are more environmentally friendly (Babcock 2015).
[12] McPhail, Westcott et al. (2011), provide a comprehensive review of the RIN market and theoretical RIN price.

projected demand and production capacities. There is precedent for these revisions. The EPA first reduced cellulosic mandates then issued cellulosic waiver credits in each year from 2010 to 2014 and did not finalize 2014 mandates until 2015. A report from the Congressional Research Service (CRS) highlighted uncertainty as the single underlying theme of the cellulosic mandate (Bracmort 2015). Meyer and Thompson (2012), show that cellulosic waivers introduce uncertainty into both the feedstock and biofuel markets. Major revisions in annual RFS obligations have been associated with sharp movements in RIN price. Lade, Lin et al. (2015), find that RFS policy shocks in 2013 led to sharp declines in RIN price, creating burdens for producers of renewable fuel, particularly for producers of advanced biofuel and biodiesel.

This study provides a real options framework to separate and analyze the effect of biofuel policy uncertainty in the form of highly volatile RIN price. By applying a two stochastic variable real options approach, the effect of uncertainty and irreversibility on a second-generation biofuel investment is examined while considering the volatility of RIN price, conventional fuel price and their correlations. Previous studies have not included both conventional fuel price and RIN price as two sources of uncertainty faced by the firm.

The addition of the RIN market certainly encourages investment, but the uncertainty associated with RIN price limits the effectiveness of the RIN program. This analysis shows that large uncertainties in the RIN market has significant effects on the firm's investment and operation decisions. Investment incentives such as bonus depreciation allowances and producer tax credits work to counteract uncertainty effects, but their efficacy is limited by the aggregate uncertainty faced by the firm.

2.2. The Fundamentals of Biodiesel RIN Price

To form a basic idea of how a RIN price is determined, suppose the market for biodiesel is represented in Figure 16 (Irwin (2014) & Stock (2015)). Biodiesel demand is perfectly elastic for biodiesel prices equal to conventional diesel prices, since the two are near perfect substitutes. Therefore, the demand price for biodiesel is assumed equal to the price of conventional diesel (P). These characteristics are also found in the market for other second-generation biofuels that are drop-in ready and have no blend wall.[13] Since the biodiesel market is well established, it serves as a guide throughout.

When the market is in equilibrium without any mandates, the equilibrium quantity supplied and consumed is given by Q_{eq}. Under the RFS, mandated quantities of biofuel must be blended into the nation's fuel supply. When the mandated quantity is greater than equilibrium quantity demanded, the mandate is said to be binding and will lead to a biodiesel supply price (P_b), which is greater than demand price (P). This creates a gap between the willingness to pay for biofuel and the marginal cost of supplying biofuel at the mandated quantities. RIN prices bridge this gap and ensure the mandate is met. This gap represents the core (intrinsic) value of the RIN.

An increase in the price of conventional fuel will cause an upward shift in the demand for biofuel, which would decrease the core value of RINs as depicted in Figure 17.

Similarly, shifts in the biodiesel supply curve will affect the core value of RINs. Policy shocks also affect the core value of RINs. For example, blender tax credits will cause an upward shift in

[13] The market for ethanol is not characterized by a horizontal demand curve due to the ethanol blend wall.

biofuel demand (Figure 18) and shifts in the mandate will affect the core value of the RIN (Figure 19).

As discussed in 1.2.1, RIN price is composed of a core value and a time value. The core value is determined by the blend margin. Departures from the blend margin indicate the presence of a time-value, which is influenced by factors outside the conventional fuel and biofuel prices. The time value is unknown *a priori* but can be found by taking the difference of the observed RIN price and the core value so that

$$time\ value = R - [(P_b - P) + blender's\ tax\ credit].$$

As, Irwin (2014) pointed out, the time-value reflects market expectations which are influenced by policy uncertainties. Policy uncertainty in annual mandates and the uncertain renewal of a blender tax credit, create incentives to bank RINs in periods where it is economical to do so. For example, if aggressive mandates are expected in the future periods, which increase the cost of compliance, demand in the current period may ramp up to take advantage of the relatively low cost current period. Similarly, if market participants expect the blender tax credit to be removed in the next period, demand for the biofuel and thus applicable RINs, will surge in the current low cost period. The policy uncertainties that influence the time value further add to the stochastic nature of RIN prices.

Figure 20 illustrates the differences between D4 biodiesel RIN prices and biodiesel blend margins. The vertical grey dotted lines represent EPA proposal for revisions and the vertical black solid lines represent EPA final rulemaking for the subsequent year's annual mandate. Each year the RFS mandates are to be announced no later than November 30, however, the EPA has consistently been late. The proposal for 2011 mandates included a 28% increase in annual percentage standards for advanced biofuels. A biodiesel blender's tax credit was also re-

introduced in 2011, after it previously was allowed to expire on December 31, 2009. The biodiesel blender's tax credit was first implemented in 2005, but lawmakers have allowed the policy to expire in 2010, 2012, 2014 and 2015. Each year the tax credit is set to expire on December 31, leaving the market uncertain as to whether it will be granted an extension or allowed to expire.

The blender's tax credit shifts demand for biodiesel and creates an additional incentive to acquire surplus RINs, which can be rolled over to the following year after the tax credit has expired. Recall that RINs may be saved and applied toward future mandates for up to two years. Only 20% of an obligated party's mandate can be met with vintage RINs, but when the tax credit is active, blenders are able to exceed their current year mandates at a discounted cost to apply surplus RINs in costlier future periods.

In July of 2011, EPA proposed significant increases in annual percentage standards and in December of 2011, EPA finalized those standards, which resulted in a 15% increase in total renewable percentage standards, a 55% increase in advanced biofuel and a 32% increase in biodiesel percentage standards for 2012. The blender's tax credit was allowed to expire in 2012, yet Figure 20 illustrates there is still significant departure from the blend margin during the first half of 2012. In fact, D4 RIN prices were on average $0.34 higher than the blend margin until June of 2012, indicating the policy shock of an upward revision in RFS mandates had a lingering effect on RIN price through the first half of 2012.

The blender tax credit was reinstated in early January of 2013 and roughly coincided with the announcement of proposed RFS volumes for 2013, about 36 weeks late. Nearly eight months later the EPA finalized the RFS volumes for 2013, which happened to coincide with sharp declines in RIN price. The finalized RFS volumes did not differ from the proposed volumes. By

examining the data, it becomes clear the sharp down turn in RIN price in the summer of 2013 had more to do with a sharp decline in the price of biodiesel.

Through 2014 and the first half 2015, Figure 20 illustrates that D4 RIN prices closely followed the blend margin until the early summer of 2015 During this time there were no RFS revisions and the blender tax credit was not in effect. It was not until June of 2015 that EPA first proposed the RFS volumes for 2014, 2015 and 2016. The 2015 mandates were nearly identical to 2014 and not surprisingly, D4 RIN prices departed from the blend margin, but this time, fell below the blend margin, indicating the market expected an oversupply of RINs in 2015.

The biodiesel blender's tax credit was reinstated once again in 2016, and the 2016 RFS volumes were finalized in December of 2015. For 2016, the final total renewable percentage standard grew by more than 11% from 2014 to 2016 while the biodiesel percentage standard increased 16.5%. These aggressive mandates and reinstatement of tax credits led to RIN prices, which began trending upward while blend margins remained relatively flat.

From the above discussion, it is clear that RFS policy and blender tax incentive uncertainties are additional sources of stochasticity that contribute to RIN price volatility, in addition to stochastic influences through conventional fuel price. This necessitates the use of a two-variable real-option model, which captures the uncertainties faced in both the conventional fuel market and the RIN market. The two-variable model is also superior in its ability to isolate market uncertainties and policy uncertainties. Therefore, one can hold conventional fuel volatilities constant while varying RIN volatilities to examine the effects of reduced policy uncertainty.

2.3. A Model of Optimal Entry Into and Exit From the Biofuel Market

Producers of renewable fuel face an output price, which is composed of the price of conventional fuel, plus the price of applicable RINs, as discussed in the previous section. . Therefore, the firm producing renewable fuel decides to enter the market based on conditions observed in both the conventional fuel markets and the RIN market and, the firm is subject to uncertainty in both markets.

Should an inactive firm decide to enter the market for renewable fuel the firm will incur an initial cost of investment, which is sunk. For example, investment in the production of a biofuel requires a large capital investment to build a new production facility including, newly constructed structures, storage tanks, boilers, crushing and other industrial equipment. An active firm has the option to either continue operation or exit operations by incurring an additional sunk cost.

Given the option to delay, the firm's decision to enter or exit is impacted by the irreversibility and uncertain nature of the investment. The plant can be in two different states, inactive ($i = 0$) or active ($i = 1$). In an inactive state, the firm incurs no costs since the investment has not yet been initiated and therefore the building of the plant has not begun. The firm must decide when to switch from the inactive state where $i = 0$ to the active state where $i = 1$ by choosing t_1. When the firm decides to become active it incurs the one-time sunk cost, k, to enter the market and pays a stream of operating costs w, in exchange for a stream of revenue $P + R$, assuming quantity is normalized to one. Operating costs w, consist of the feedstock costs, transportation of that feedstock to the conversion facility and conversion costs. The biofuel producer's output price is the price of conventional fuel, P, plus the RIN price, R. The conventional fuel could be

gasoline, diesel or jet fuel, depending on the advanced biofuel produced. Similarly, the type of biofuel, production pathways, and feedstocks dictates the particular RIN type.

An active plant has the option to exit the market with a one-time payment. Once active, the firm must decide when to switch states from active where $i = 1$ to inactive where $i = 0$ by choosing t_0. The payment to exit may be positive if the sunk costs (k), incurred from exit exceed any salvage values (s), from the sale of capital on a secondary market, or negative if the salvage value of remaining capital exceeds any sunk costs. For example, when $s \leq k$ the one-time payment to exit the market is positive. The firm's flow of payments in each state is represented as

$$\pi_{i=0,1}(P,R)\begin{cases}0 & \text{if the firm is inactive} \quad and\ i=0\\ P+R-w & \text{if the firm is active} \quad and\ i=1\end{cases}.$$

Given the discount rate δ, the firm's optimal entry and exit decision satisfies

$$V_0(P_0,R_0) = \max_{t_1} E_0\left\{\int_0^{t_1} \pi_0(P(t),R(t))e^{-\delta t}dt + \ [(V_1(P(t_1),R(t_1)) - k)e^{-\delta t_1}]\right\} \quad (2)$$

$$V_1(P_0,R_0) = \max_{t_0} E_0\left\{\int_0^{t_0} \pi_1(P(t),R(t))e^{-\delta t}dt + \ [(V_0(P(t_0),R(t_0)) - s)e^{-\delta t_0}]\right\} \quad (3)$$

subject to dP, dR and $P(0) = P_0,$ and $R(0) = R_0$.

The firm chooses the optimal time to enter or exit the market to maximize the expected net present value of the firm's investment option. The trigger prices which induce the firm to enter or exit are thus implicitly defined and represented as P_h and P_l respectively.

2.3.1 Market Dynamics

The price of a gallon of conventional fuel, P, follows a generalized Ito process such that the change in price

$$dP = \alpha_p(P)dt + \sigma_p(P)dz_p \tag{4}$$

where $\alpha_p(P)$ is the deterministic trend in the price process and $\sigma_p(P)$ represents the instantaneous standard deviation. The trend changes over the time increment dt . The change in z_p , dz_p , follows a Wiener process such that $dz_p = \varepsilon_t\sqrt{dt}$ where, ε_t is a normally distributed random variable with a mean of zero and a standard deviation of unity. Furthermore ε_t is serially uncorrelated so that $E(\varepsilon_t, \varepsilon_s) = 0 \ \forall\, t \neq s$ and thus the values of dz_p for any two different intervals of time are independent, following a Markov process with independent increments. It is also assumed that the discount rate δ is greater than the drift rate α_p, which indeed must hold otherwise investment would never be optimal, as the growth rate would outpace the discount rate. Hence, it would always be possible to do better by waiting longer.

The price at which RINs are traded is determined on a secondary market and prices have exhibited high degrees of volatility throughout RIN history. One primary cause of RIN price volatility is uncertainty in annual RFS obligations and uncertainty in blender tax credit renewal. This uncertainty is captured by assuming RINs follow a stochastic process. In this case, RIN prices are assumed to follow a generalized Ito process, just as is the case for the price of conventional gasoline.

$$dR = \alpha_r(R)dt + \sigma_r(R)dz_R \tag{5}$$

where $\alpha_r(R)$ is the deterministic trend in the price process, $\sigma_r(R)$ represents the standard deviation and the same assumptions of equation (4) apply. The two stochastic processes are potentially correlated in their Wiener processes such that $E[dz_P dz_R] = E[\varepsilon_{tP}, \varepsilon_{tR}]dt = cov(dz_P, dz_R)/\sqrt{var(dz_P)var(dz_R)} = \rho_{pr}dt$.

2.3.2 The Firm's Decision to Enter

The idle firm has expected net present value denoted as $V_0(P, R)$. Since the firm is inactive in this state, $V_0(P, R)$ represents the value of the option to enter the market. Over the range of prices P and R where it is optimal for an idle firm to remain idle, the asset of the investment opportunity must be willingly held. Since there are no operating profits being generated, the only return is the expected return from the option value $E_t[dV_0(P, R)]dt^{-1}$.This is also referred to as the capital appreciation, of the firm's asset as prices increase, where in this case, the asset is the firm's option to enter the market.

The required return from the investment is represented by the function $\delta V_0(P, R)$and the no-arbitrage condition of efficient markets sets these two returns equal.

$$\delta V_0(P, R) = E_t[dV_0(P, R)]dt^{-1} \tag{4}$$

Under the efficient markets theorem, equation (4) must hold and implicitly defines the entry trigger prices for conventional fuel and RINs. Equation (4) essentially says that over an infinitesimal period dt, the total expected return on the investment opportunity is equal to its expected rate of capital appreciation and facilitates the evaluation of the affect that changes in output price have on the value of an inactive firm.

Using Ito's Lemma $dV_0(P, R)$ is expanded using a Taylor series expansion. It is assumed that P and R are both continuous-time stochastic processes as represented by equation (4) and (5). Consider that the function $V_0(P, R)$ is at least twice differentiable in P and R. Total differentiation of $V_0(P, R)$ to higher-order terms, and making the necessary substitutions yields

$$dV_0(P,R) = \frac{\partial V_0}{\partial P}\alpha_p(P)dt + \frac{\partial V_0}{\partial P}\sigma_p(P)dz + \frac{\partial V_0}{\partial R}\alpha_r(R)dt + \frac{\partial V_0}{\partial R}\sigma_r(R)dz$$
$$+ \frac{1}{2}\frac{\partial^2 V}{\partial P^2}\left(\sigma_p(P)\right)^2 dt + \frac{1}{2}\frac{\partial^2 V}{\partial R^2}\left(\sigma_r(R)\right)^2 dt$$
$$+ \frac{\partial^2 V}{\partial P \partial R}(\sigma_p(P)\sigma_r(R))\rho_{pr}dt \tag{6}$$

Now, substituting equation (6) into equation (4) the following is obtained

$$\delta V_0(P,R) = E_t\left[\frac{\partial V_0}{\partial P}\alpha_p(P)dt + \frac{\partial V_0}{\partial P}\sigma_p(P)dz_p + \frac{\partial V_0}{\partial R}\alpha_r(R)dt\right.$$
$$+ \frac{\partial V_0}{\partial R}\sigma_r(R)dz_R + \frac{1}{2}\frac{\partial^2 V_0}{\partial P^2}\left(\sigma_p(P)\right)^2 dt + \frac{1}{2}\frac{\partial^2 V_0}{\partial R^2}\left(\sigma_r(R)\right)^2 dt$$
$$\left. + \frac{\partial^2 V}{\partial P \partial R}(\sigma_p(P)\sigma_r(R)\rho_{pr})dt\right] dt^{-1}. \tag{7}$$

Given that $E(dz) = 0$ the middle terms $\frac{\partial V_0}{\partial P}\sigma_p(P)dz_p = \frac{\partial V_0}{\partial R}\sigma_r(R)dz_r = 0$ and equation becomes the second-order homogenous partial differential equation

$$\frac{\sigma_p(P)\sigma_r(R)\partial^2 V_0}{\partial P \partial R}(\rho_{pr}) + \frac{\sigma_p(P)^2}{2}\frac{\partial^2 V_0}{\partial P^2} + \frac{\sigma_r(R)^2}{2}\frac{\partial^2 V_0}{\partial R^2} + \frac{\partial V_0}{\partial P}\alpha_p(P)$$
$$+ \frac{\partial V_0}{\partial R}\alpha_r(R) - \delta V_0(P,R) = 0 \tag{8}$$

2.3.3. The Firm's Decision to Exit

The value of the active firm is denoted as $V_1(P,R)$, which is the expected discounted value of the active firm's net earnings plus their option to exit. Over the range of prices P and R, where it is optimal for an active firm to remain active, the firm is earning $(P + R - w)$. The active firm receives a dividend equal to the flow of profits they earn while active in the market $(P + R - w)$ and capital gains arise from the expected future appreciation of the firm. Similar to equation (5), efficiency in markets requires

$$\delta V_1(P,R) = (P + R - w) + E_t[dV_1(P,R)]dt^{-1}. \tag{9}$$

Applying Ito's Lemma to expand $dV_1(P,R)$ and following similar procedures as in equations (6)-(8) results in the second-order non-homogenous partial differential equation

$$\frac{\sigma_p(P)\sigma_r(R)\partial^2 V_1}{\partial P \partial R}(\rho_{pr}) + \frac{\sigma_p(P)^2}{2}\frac{\partial^2 V_1}{\partial P^2} + \frac{\sigma_r(R)^2}{2}\frac{\partial^2 V_1}{\partial R^2} + \frac{\partial V_1}{\partial P}\alpha_p(P) \tag{10}$$

$$+ \frac{\partial V_1}{\partial R}\alpha_r(R) - \delta V_1(P,R) = w - P - R.$$

Solving (8) and (10) will implicitly define the prices, which trigger the firm to enter and exit the market respectively. Partial differential equations such as (8) and (10) often require the use of numerical techniques to approximate a solution. In this case, analytical solutions are possible if one is willing to make rather strong assumptions regarding the operating cost (w).[14] Although operating costs are certainly a function of conventional fuel price, it is not likely to be linear in nature. Such a strong assumption would be prohibitive with little gain, particularly in light of the fact that numerical techniques have been developed to overcome the challenges associated with partial differential equations. For example, Brekke and Øksendal (1994) developed the method of variational inequalities for solving optimal entry and exit models where no closed form solution can be found. This will be revisited in section 2.4.3.

2.4. An Application to a Second-generation Biofuel Investment

To demonstrate the influence of market and policy uncertainty, the paper focuses on producers of second-generation biofuels that generate biodiesel D4 RINs. The four RVOs are nested within each other so that fuels with higher GHG reductions can be used to meet the standards for a

[14] If w, is a linear function of conventional fuel then (8) and (10) may be transformed into an ordinary differential equation, consisting of one stochastic variable which is a ratio of RIN price over conventional fuel price. One can then solve for the ratio of prices, which induce the firm to enter or exit the market.

lower GHG reduction. For example, cellulosic biofuels with a lifecycle greenhouse-gas emission that results in a 60% reduction from the baseline can be submitted for the cellulosic biofuel category, the biomass-based diesel category, the advanced category or the conventional renewable fuel category. Therefore, the cellulosic D7 and D3 RINs are worth at least as much as D4, D5, and D6 RINs. Similarly, fuels in the biomass-based diesel D4 category are worth at least as much as D5 and D6 RINs. However, production of cellulosic biofuel has not been in significant quantity. As of March 10 2016, the total RIN count was 2,842 million. Of that, only 14.75 million were D3 RINs and zero D7 RINs were generated. Therefore, historical data for D3 and D7 RIN price is lacking and thus D4 RIN prices are evaluated in this study.

2.4.1 Estimating Stochastic Price Processes

Two parameters are estimated for each geometric Brownian motion, the drift rate α, and the standard deviation σ. Next, the covariance of the two stochastic processes ρ_{pr} is estimated. Data on conventional fuel price is obtained from the Energy Information Association and RIN price data is obtained from EcoEngineers, a provider of daily and historical RIN data.

To estimate the stochastic processes, the assumption that the two prices follow a geometric Brownian motion (GBM) it tested. Recall, the change in z , dz_p , follows a Wiener process such that $dz = \varepsilon_t\sqrt{dt}$ where, ε_t is a normally distributed random variable with a mean of zero and a standard deviation of unity. Since the variance of ε_t is one, the variance of $dz = \left(\sqrt{dt}\right)^2 * Var[\varepsilon_t] = dt$ and the Wiener process is normally distributed with a mean of zero and variance of $dt = t_k - t_{k-1}$. The Wiener process implies ΔlnP_t and ΔlnR_t are independent of autoregressive lags and normally distributed. To examine the independence or potential lack thereof, the autocorrelation and partial autocorrelation are plotted for both the transformed series of conventional fuel price ΔlnP_t and RIN price ΔlnR_t.

Based on Figure 21, it does not appear that independence is tenable for ΔlnR_t. This implies $\in_{R_t}$is not white noise in $\Delta lnR_t = \varphi_1 \Delta lnR_{t-1} + \in_{R_t}$. Therefore, additional lags are added to the specification until $\in_{R_t}$ becomes white noise in $\Delta lnR_t = \sum_{i=1}^{p-1} \varphi_i \Delta lnR_{t-i} + \in_{R_t}$, where (p) is the lag order

A Portmanteau test for white noise is used to determine that the third lag produces white noise in the residual term. Given that $\in_{R_t}$ is white noise, the Augmented Dickey Fuller (ADF) test is specified to verify the process is a unit root and thus satisfies GBM properties. By simply including an additional term of lnR_{t-1} and a constant term, one can test for a unit root using Case II of the Augmented Dickey Fuller such that

$$\Delta lnR_t = \gamma_0 + \gamma_1 lnR_{t-1} + \sum_{i=1}^{p-1} \varphi_i \Delta lnR_{t-i} + \in_{R_t}.$$

The null hypothesis in Case II is that $\gamma_1 = 0$. After conducting the ADF test under Case II a p-value of 0.5763 is obtained, signaling a failure to reject the null hypothesis and thus reaching a conclusion that the series is indeed unit root. This implies the estimated autoregression includes a constant term, but the true process follows a unit root with no drift. The result is consistent with GBM, but the sensitivity of such tests to null specification motivates a second test.

Based on an examination of D4 RIN price time series in Figure 20, the data appear to have a non-zero intercept and follow a time trend. Therefore, Case IV of the ADF test is specified such that

$$\Delta lnR_t = \gamma_0 + \gamma_1 lnR_{t-1} + \gamma_2 t + \sum_{i=1}^{p-1} \varphi_i \Delta lnR_{t-i} + \in_{R_t}.$$

In Case IV, the null hypothesis is that $\gamma_1 = \gamma_2 = 0$ and γ_0 is allowed to be any value. In this case, test results are even stronger in failing to reject the null with a p-value of 0.9957. D4 RIN prices are thus appropriately modeled as a GBM. Therefore, drift and variance parameters can be deduced from the regression of

$$\Delta lnR_t = c + \sum_{i=1}^{p-1} \varphi_i \Delta lnR_{t-i} + \in_{R_t}$$

Recall that $\mu_r = (\alpha_r - \frac{1}{2}\sigma_r^2)$ such that, $\Delta lnR(t) = \mu_r dt + \sigma_r dz$ and $\Delta lnR(t) = (\alpha_r - \frac{1}{2}\sigma_r^2)dt + \sigma_r dz$. If $dt = 1$ then $E[\Delta lnR(t)] = \mu_r$ and taking the expectation of both sides of the above regression

$$E[\Delta lnR_t] = E\left[c + \sum_{i=1}^{p-1} \varphi_i \Delta lnR_{t-i} + \in_{R_t}\right]$$

results in $\mu_r = c + \mu_r \sum_{i=1}^{p-1} \varphi_i$. Therefore, $\mu_r = \frac{c}{1-\sum_{i=1}^{p-1} \varphi_i} = (\alpha_r - \frac{1}{2}\sigma_r^2)$ and the standard deviation can be read directly from the regression results root mean squared error (Schmit, Luo et al. 2011). OLS results indicate the constant c is not statistically different from zero. Therefore $\left(\hat{\alpha}_r - \frac{1}{2}\hat{\sigma}_r^2\right)$ is set equal to zero and the drift parameter is solved for by plugging the root mean squared error in for $\hat{\sigma}_r$.

$$\left(\hat{\alpha}_r - \frac{1}{2}\hat{\sigma}_r^2\right) = 0$$

$$\hat{\alpha}_r = \frac{1}{2}\hat{\sigma}_r^2$$

$$\hat{\alpha}_r = \frac{1}{2}(0.04793)^2 = 0.0011486$$

Recall that dt was set to equal one, therefore the drift and variance parameters are annualized following the convention of converting 1-day parameters to h-day parameters. This requires scaling by a factor of $\sqrt{h}$ resulting in a drift rate of $\hat{\alpha}_r = 0.0011486 * \sqrt{252} = 0.0182$ and a standard deviation of $\hat{\sigma}_r = 0.04793 * \sqrt{252} = 0.7608$.

This same process is carried out for the drift and standard deviation of conventional fuel price. By examining the correlogram (Figure 22) of ΔlnP_t it appears that, independence is tenable.

To formally test for independence a Portmanteau white noise test is conducted with the null hypothesis that white noise is present in the residuals $\in_{P_t}$ of $\Delta lnP_t = \varphi_1 \Delta lnP_{t-1} + \in_{P_t}$. The Portmanteau Q-test statistic is 43.079, resulting in a p-value of 0.3411, signaling a failure to reject the null hypothesis.

Based on an examination of diesel price against time (Figure 23), the data appear to have a non-zero intercept and time trend. So an ADF test of Case IV is specified such that

$$\Delta lnP_t = \gamma_0 + \gamma_1 lnP_{t-1} + \gamma_2 t + \in_{P_t}.$$

The null hypothesis in Case IV is that $\gamma_1 = \gamma_2 = 0$ and γ_0 is allowed to take any value. The above ADF test with zero lags produces a p-value of 0.7915 signaling a failure to reject the null hypothesis of unit root. Variations of the Case IV ADF test with additional lags are also tested and results signal a failure to reject the null hypothesis in all cases.

To obtain the drift and variance parameters an OLS regression of $\Delta lnP_t = c + \varphi_1 \Delta lnP_{t-1} + \in_{P_t}$ is conducted just as was the case for D4 RIN price. Once again, the constant term is not statistically different from zero. Reading the root mean squared error directly from the OLS results and plugging into the expression for the mean, the drift rate is found to be

$$\hat{\alpha}_p = \frac{1}{2}(0.01844)^2 = 0.00017.$$

With a 1-day drift rate of 0.00017, the annual drift rate becomes $\hat{\alpha}_p = 0.00017 * \sqrt{252} =$ $.0026986$ and the annual standard deviation becomes$\hat{\sigma}_p = 0.01844 * \sqrt{252} = 0.2927$.

Next, the correlation coefficient of conventional fuel price and D4 RIN price is obtained by estimating the correlation between the residuals of the fitted OLS regressions on ΔlnR_t and ΔlnP_t. Theory suggests RIN prices are negatively correlated with conventional fuel price. As the price of conventional fuel increases, substitution effects take hold and shift the demand for alternative fuels, which in turn causes a decrease in the price of RINs (Figure 17). This negative relationship between RIN price and conventional fuel price is borne out in the data as evidenced by the estimated correlation of -0.0129.

2.4.2 Cost Parameters

Cost parameters are based on the production of a green-diesel using purchased pennycress seed, which is crushed by the biorefinery to extract the oil. Typically, a techno-economic analysis is required to estimate cost parameters of a production pathway.[15] Due to the detailed and specialized information required to conduct a TEA, cost parameters are sourced from previous literature. Two primary cost parameters utilized in this study are operating costs and capital costs (initial investment cost). The cost associated with producing renewable fuel depends on the

[15] Techno-economic analysis is a method used to assess the technical and economic performance of a particular production pathway (Brown and Brown 2013). TEA's represent a simplified version of commercial scale projects, that allows the biorefinery's production pathway to be evaluated for feasibility. The primary method of developing a TEA is through a detailed process model. Performance information is collected on the technologies under consideration. Appropriate production scenarios are identified and the process model is designed using process-engineering software. Based on the process design, capital investments are determined and a discounted cash flow analysis is performed. This allows the investment and production cost of a biorefinery to be determined.

technological process selected by the firm. One such technological process that shows considerable promise is a Green Fuel Technology capable of producing three different types of alternative fuel through the conversion of eight or more different feedstocks. A firm deciding to invest will employ this technology to produce the most profitable alternative fuel. Presumably, this would imply that the firm produces the alternative fuel, which generates the most valuable RIN. Of the three types of RINs under consideration in this study, D4 Biodiesel RINs have the greatest value due to the nested structure of RINs discussed in section 4.[16] D4 RINs are generated through the production of biodiesel, green-diesel and renewable jet fuel. Firms employing the Green Fuel Technology are able to produce either of these fuels. Producers and importers of conventional fuel are required to submit a specified number of RINs each year based on their annual production (or import) of conventional fuel. Therefore, the Renewable Fuel Standard supports demand for either biodiesel or green-diesel. On the other hand, producers and importers of conventional jet fuel are not obligated to submit RINs.

Green-diesel is a second-generation alternative fuel that can be blended at any proportion with conventional diesel. Chemically, green-diesel is identical to conventional petroleum based diesel, so there is no blend limit. Vehicles can operate on 100% green-diesel without any engine modifications. It requires no changes to infrastructure and it has a higher cetane rating than conventional diesel (UOP 2016). A promising feedstock for this conversion technology is

[16] Based on this nested structure cellulosic biofuel (D3) and cellulosic diesel (D7) RINs are more valuable than biodiesel D4 RINs. However, D3 and D7 RINs are not considered in this study due to a lack of significant historical production and thus lacking historical data. The first year D3 and D7 RINs were generated was 2012, but of the 15,306,658,432 total RINs generated in 2012, only 21,810 (0.0001425%) of those were D3 or D7 RINs. In 2015 the total RIN count, including D3, D4, D5, D6 and D7 was 17,908,121,504 RINs. Of the 17.9 billion 2015 RINs, D3 and D7 RINs made up just 141,557,292, or 0.7905%. As of March 10, 2016 the total RIN count was 2,842,575,051 and only 14,755,743 D3 RINs and zero D7 RINs were generated.

pennycress oil. Pennycress is a non-food, winter annual cover crop that produces high quality oil seeds. As a winter cover crop, it does not displace any land for food crops, and it provides all the soil benefits associated with other cover crops. Additionally the high quality oil seed has potential to generate significant farm revenue for land normally left fallow (Moser et al. 2009).

English, Menard et al. (2016), provide a thorough review of operating and capital costs for a variety of production pathways, including green fuel production pathways. The authors examine four different conversion processes: 1) hydro-processing of seed crush or purchased oils, 2) pyrolysis to hydro-processing, 3) gasification and 4) pyrolysis of biomass. Since pennycress is the chosen feedstock under consideration for this study, cost parameters represent the hydro processing of pennycress oil, resulting from onsite crushing of pennycress seed.

Based on a conversion facility, which demands 96.3 million pounds of pennycress oil, with a production capacity of 13 million gallons, the capital costs are estimated to be $3.93 per gallon. Operating costs are based on feedstock production, transportation and conversion costs and are estimated to be $4.35 per gallon (English, Menard et al. 2016). The following table summarizes the cost parameters and their source.

2.4.3 Numerical Methods

This section relies on a function approximation and collocation method to solve the firm's problem of optimal entry and exit timing. Fackler (2008), provides a general approach and MATLAB implementation which is applied to the optimal entry and exit problem of second-generation biofuel producers.

The firm's problem in (2) and (3) can be written more succinctly as

$$V_i(P,R) = \max\{\pi_i(P,R) + E\left[(dV_i(P,R))\right] dt^{-1}, V_j(P,R) - C_{ij}\} \tag{11}$$

where C_{ij}, is the cost of switching states between active or inactive. Brekke and Øksendal (1994), showed that (11) can be expressed as a set of variational inequalities as long as the value functions $V_0(P,R)$ and $V_1(P,R)$ are stochastically complete which corresponds to the high contact condition of optimal stopping. First note that (11) implies that

$$V_i(P,R) \geq \pi_i(P,R) + E\left[(dV_i(P,R))\right]dt^{-1}$$

and

$$V_i(P,R) \geq V_j(P,R) - C_{ij}.$$

Apply Ito's lemma to $E\left[(dV_i(P,R))\right]dt^{-1}$ as in (6) and (7) the firm's decision to enter can be represented as

$$\frac{\sigma_p \sigma_r \partial^2 V_0}{\partial P \partial R}\left(\rho_{pr} PR\right) + \frac{{\sigma_p}^2}{2}\frac{\partial^2 V_0}{\partial P^2}P^2 + \frac{{\sigma_r}^2}{2}\frac{\partial^2 V_0}{\partial R^2}R^2 + \frac{\partial V_0}{\partial P}\alpha_p P + \frac{\partial V_0}{\partial R}\alpha_r R = \delta V_0(P,R).$$

Combining this with the value matching condition

$$V_0(P,R) = V_1(P,R) - k,$$

the optimal value function thus satisfies the conditions,

$$\frac{\sigma_p \sigma_r \partial^2 V_0}{\partial P \partial R}\left(\rho_{pr} PR\right) + \frac{{\sigma_p}^2}{2}\frac{\partial^2 V_0}{\partial P^2}P^2 + \frac{{\sigma_r}^2}{2}\frac{\partial^2 V_0}{\partial R^2}R^2 + \frac{\partial V_0}{\partial P}\alpha_p P + \frac{\partial V_0}{\partial R}\alpha_r R \leq \delta V_0(P,R) \tag{12}$$

and

$$V_0(P,R) \geq V_1(P,R) - k. \tag{13}$$

One of these two conditions must hold with equality at each point in the state space. Whichever of these two hold with equality determines the optimal firm decision. The left hand side of condition (12) is the expected return the firm will gain from delaying investment. The right hand

side of condition (12) is the firm's required return to continue delayed investment. If condition (12) holds with inequality the expected returns from delaying, fall short of the firm's required return to remain inactive through continued to delay. A similar story holds true for condition (13). The left hand side of condition (13) is the expected present value of delaying investment and the right hand side is the expected present value for the active firm, minus the cost to enter the market. In other words, condition (13) simply compares the expected payoffs between the inactive firm and the active firm. The switching point occurs when condition (13) holds with equality. So if the value matching condition holds with strict equality such that $V_0(r_H) = V_1(r_H) - k$, then the optimal firm decision is to invest the sunk cost k and enter the market.

To find the optimal value function in each state, suppose the firm's value function $V_i(P, R)$, where $i = 0,1$, can be approximated by $\phi(P, R)\theta_i$ where $\phi(\cdot)$ represents a family of n-basis functions and θ_i is an n-vector of approximating coefficients. The approximating coefficients are then fixed by requiring $\phi(P, R)\theta_i$ to satisfy the Bellman equation at n-collocation nodes. The collocation nodes are simply points along the real number line between the upper and lower limits.

The approximate differential operator is defined as

$$\beta_i(P, R) = \delta V_i(P, R) - \alpha_r R \frac{\partial V_i}{\partial R} - \alpha_p P \frac{\partial V_i}{\partial P} - (R)^2 \frac{{\sigma_r}^2}{2} \frac{\partial^2 V_i}{\partial R^2} - (P)^2 \frac{{\sigma_p}^2}{2} \frac{\partial^2 V_i}{\partial P^2} - \left(\rho_{pr} PR\right) \frac{\sigma_p \sigma_r \partial^2 V_i}{\partial P \partial R}.$$

The set of n-basis functions for a family of approximating functions form the $n \times n$ matrix represented by Φ and B represents $n \times n$ matrix of $\beta_i(P, R)$ evaluated at n-nodal points (Fackler 2008).

The problem is written as an extended vertical linear complementary problem (EVLCP) as developed by Gowda and Sznajder (1994).[17] To solve the EVLCP, a smoothing Newton algorithm is employed while specifying the family of approximating solutions to be a piecewise linear function using upwind finite differences to approximate first and second order differentiation. Piecewise linear functions are most appropriate for entry and exit problems because of the inherent discontinuities in the second derivative of the value function where it becomes optimal for the firm to switch states from idle to active, or from active to idle. The approximation state space spans a wide range of possible prices. The lower bound is set to \$0.00 and the upper bound is set to \$20 per gallon for both conventional fuel and RIN price.[18] Between each lower and upper bound, 100 nodes are specified, so that the matrices of nodal points for both conventional fuel price and RIN price are both 100×100.[19]

2.5. Results

In a single parameter model, break-even entry price is defined as $\frac{W_h}{\delta-\alpha_p} \geq \frac{w}{\delta} + k$ where parameters are as previously defined. The price which triggers investment in this single variable break-even analysis is $\frac{W_h}{0.10-0.0027} \geq \frac{4.35}{0.10} + 3.93$ or that $W_h \geq \$4.62$ per gallon. Similarly, the single variable break-even exit price is defined as $\frac{W_l}{\delta-\alpha_p} \leq \frac{w}{\delta} + s \rightarrow \frac{W_l}{0.10-0.0027} \leq \frac{3.38}{0.10} + 0.9825$ or that $W_l \leq \$4.33$.

---

[17] Interested readers should also refer to (Fackler 2008).

[18] An important choice in specifying the model is the choice of lower and upper limits on the approximation interval and the number of nodal points. If the interval is too wide, a large number of node points will be required to obtain accurate solutions. Conversely if the approximation interval is too narrow, such that the process starts near an optimal switching point, the solutions may be inaccurate (Fackler 2008).

[19] Therefore M_1 and M_2 are each $10{,}000\times10{,}000$ matrices, while q_1 and q_2 are each $10{,}000\times1$ matrices.

In a two variable model, break-even entry price is defined as $\frac{W_h}{\delta-\alpha_r-\alpha_p} \geq \frac{w}{\delta} + k$ where the output price is now a composite of both the conventional fuel price and D4 RIN price, as is the case under a binding RIN market. Using the drift rate for both D4 RIN and conventional diesel price, the two-variable break-even prices which triggers investment is $\frac{W_h}{0.10-0.0027-0.0182} \geq \frac{4.35}{0.10} +$ 3.93 or that $W_h \geq \$3.75$. Two-variable break-even exit price is defined as $\frac{W_l}{\delta-\alpha_r-\alpha_p} \leq \frac{w}{\delta} + s$ and is calculated to be $3.52 per gallon. Therefore, when uncertainty and irreversibility are not considered, RIN prices reduce the entry price and reduce the gap between the entry and exit prices. From this perspective, RIN markets are accomplishing the goals set forth by policy makers. However, RIN markets have been highly volatile since their inception. The volatility in RIN markets exacerbates the volatility already present in the market for transportation fuels, and entering into production of biofuels requires significant capital investment, which is at least partially irreversible. The trigger price, which induces investment, is significantly larger when uncertainty and irreversibility is considered. Before examining the results of the real options approach using two stochastic variables, the trigger prices under a single variable real option analysis are examined.

In the absence of a RIN market, firms would face a single output price. Using this rational a real option analysis with a single stochastic variable is performed using the same parameters from conventional diesel to represent the parameters of a drop-in ready alternative, and the same cost parameters found for the production pathway discussed earlier, which also represents the production pathway of a drop-in ready alternative.

2.5.1 Results of the Single Stochastic Variable Case

Assuming the above parameters, the single stochastic variable case shows that uncertainty increases the price required for a firm to enter the market. Furthermore, the price, which induces the firm to exit the market, decreases under uncertainty, creating a gap between the entry and exit price that is larger than under the break-even analysis. This gap represents hysteresis in the market, which is a zone of inaction for the firm. The greater the uncertainty the firm is subject to, the larger this hysteresis effect becomes. In this case, the entry price is found to be $5.31 per gallon and the exit price is found to be $3.79 per gallon. Recall that the break-even entry (exit) price was found to be $4.62 ($4.33) per gallon.

2.5.2 Results of the Two-Stochastic Variable Case

In the two-variable case, the firm is subject to uncertainty in both conventional fuel price and RIN price.

Figure 24 traces out the combination of conventional fuel price and RIN price, which triggers entry and exit actions. For combinations of conventional fuel price and RIN price above the entry threshold, a firm will choose to enter the market. For combinations below the exit threshold, firms will choose to exit the market.

The downward slope of the threshold curves reflect the fact that a decline in conventional fuel price must correspond with an increase in RIN price to satisfy the threshold. As lower conventional fuel prices prevail, higher RIN prices are required to trigger firm entry and lower RIN prices are required to trigger firm exit.

Combinations between these two thresholds represent a zone of inaction. In this zone, a firm currently inactive will remain inactive and a firm currently active will remain active. This zone of inaction is wider than traditional economic theory would suggest. For example, it is well

known in economic theory that a shutdown price occurs when price falls below average variable cost. In reality, firms will often remain in business even after the price falls below average variable cost. In fact, this may be the optimal decision based on the chance that prices return to previous levels. In other words, the firm's optimal decision is dependent on the path of prices in previous periods. This path dependency is an example of economic hysteresis, and so the zone of inaction is known as the hysteresis region.

Notice the zone of inaction widens with lower conventional fuel price and higher RIN price. In other words, the ratio of relative prices is not constant over the entire range of prices. For example, if RIN prices are \$0.00, the entry and exit thresholds are found to be \$6.67 and \$3.43 respectively. This implies that if RIN prices are zero, firms would require a conventional fuel price of \$6.67 per gallon to enter the market. Conversely, firms would exit the market if conventional fuel price fell to \$3.43 per gallon. The narrowest range between the entry curve and the exit curve occur where $P + R = \$6.46$ for the entry trigger and $P + R = \$3.43$ for the exit trigger. In this case, the narrowest range of inaction occurs when RIN prices (R) are \$0.60 per gallon, resulting in a conventional fuel price (P), which triggers entry of \$5.86 for a combined minimum entry price threshold $(P + R)$ of \$6.46. This entry price threshold is satisfied for a number of combinations of $P + R$. For example, if conventional fuel prices are \$5.25 per gallon the RIN price that would induce the firm to enter the market is found to be \$1.21 resulting in a combined output price of \$6.46 per gallon. The same is true for exit price thresholds. If conventional fuel price were to fall to \$2.83, a RIN price of \$0.60 would induce the firm to abandon operations and exit the market. Supposing RIN price is \$1.21 under a more binding mandate, the corresponding conventional fuel price, which triggers exit, is found to be \$2.22.

Notice in Figure 24 that the slopes of the entry and exit threshold curves are not constant. Intuitively this result seems logical. RIN prices are substantially more volatile than conventional fuel price, so with less price support in the form of conventional fuel price, an even higher RIN price is required to overcome the price uncertainty. It is helpful to compare Figure 24 to the case of a perfectly certain RIN price. Assume RIN price still grows at the annualized rate of 0.0182 found in section 2.4.1, but the variance parameter is set to zero. Figure 25 demonstrates that the non-constant slopes in Figure 24 are attributable to differing volatilities.

When variance parameters for conventional fuel and RIN prices are set to zero, the slope of the entry threshold is constant. The exit threshold is non-constant as conventional fuel price declines which is attributable to the negative correlation and larger growth rate of RIN price.[20]

To illustrate an advantage of the two stochastic variable model, results are compared to entry and exit trigger prices assuming RIN prices are zero (single-variable model) and assuming entry and exit if determined by a discounted cash flow approach (break-even analysis) as described in section 2.5. In the single variable model, RIN price is assumed zero so that the firm's marginal revenue is determined solely by conventional fuel price similar to methodology in (McCarty and Sesmero 2014). This indeed is the case under a non-binding RFS mandate or in the case where RIN markets cease to exist.

The entry trigger price when both conventional fuel price and RIN price are stochastic is substantially larger than the threshold predicted under the one-variable real option model and the traditional break-even analysis models. Similarly, the exit trigger price is lower than the exit

[20] This is easily confirmed by setting the growth rate of both prices to be equal.

price predicted in the one-variable and break-even models. Accounting for the aggregate volatility and correlations between conventional fuel price and RIN price exacerbates the uncertainty effect.

The addition of the RIN market encourages investment, but the uncertainty associated with RIN price limits the effectiveness of the program. Figure 26 illustrates how the threshold curves shift and flatten when RIN price volatility is varied from lower to higher levels of volatility. The ability to hold one price volatility constant while varying the other is an advantage of the two-variable model. Doing so illustrates how individual volatilities can affect the firm's investment decision over the range of prices, which satisfy the entry and exit threshold curves. One can also see, in Figure 26, the zone of inaction widens as RIN volatility increases.

Entry and exit threshold curves pivot in opposite directions with increases in volatility. This occurs because increases in uncertainty increase the value of preserving the option to invest in future periods, thus increasing the threshold to enter the market, and increase the expected value of ongoing operations, resulting in a lower threshold to abandon operations. Similar intuition applies when varying conventional fuel price volatility while holding RIN price volatility constant (Figure 27). However, in this case, the intercepts along the vertical access account for much of the variation when conventional fuel price volatility is varied. Increasing conventional fuel price volatility causes the entry threshold curve to steepen while the exit threshold curve flattens.

While entry and exit thresholds are affected by varying volatilities and their correlations, they are most sensitive to changes in the firms cost structure. Varying the firm's capital cost, k, one can see that a smaller capital cost of $1 per gallon will narrow the range of inaction. Because the firm

is modeled to have a capital salvage value equal to 25%, the effect of varying capital cost is most largely seen in varying entry thresholds (Figure 28).

Firm operating cost has perhaps the largest effect on entry and exit thresholds, shifting both the entry and exit threshold curves upwards with increasing operating costs. Figure 29 illustrates how the two threshold curves shift with decreasing and increasing operating cost. If operating costs are reduced to $2 per gallon, the minimum entry trigger price is found to be $3.64 per gallon while the maximum exit trigger is found to be $1.62 per gallon.

Investigating how volatility and cost parameters affect the firm's investment decision leads to interesting implications for policy makers whom wish to incentivize production of biofuels. The RIN market was established to provide a flexible way of complying with the renewable fuel standard. However, this analysis has shown that large uncertainties in the RIN market has significant effects on the firm's investment and operation decisions. Policymakers wishing to counteract the uncertainty effects have various tools at their disposal, namely tax credits and bonus depreciation allowances.

2.5.3 Tax Incentives

Producers of second-generation biofuel may qualify for two policy incentives that encourage investment, beyond the price support provided by RINs.[21] The first incentive is a tax credit of $1.01 for every qualifying gallon of second-generation biofuel produced. The second incentive

[21] To be eligible, second-generation biofuel producers must be registered with the IRS where second-generation biofuel is defined as any liquid fuel produced from lignocellulosic or hemicellulosic matter on a renewable or recurring basis. Renewable sources of lignocellulosic or hemicellulosic matter include dedicated energy crops and trees, wood and wood residues, plants, grasses, agricultural residues, fibers, animal wastes and other waste materials, including solid municipal solid waste (H.R. 110-627 2008)

provides a 50% depreciation allowance to recover a portion of the cost of qualifying production property. The depreciation allowance may only be claimed during the first year of operations such that it takes the form of a bonus depreciation allowance. Bonus depreciation allowances accelerate the depreciation of qualified property, lowering the cost of capital required for investment in those assets and increasing cash flow to the firm making the investment (Guenther 2015). Claiming the bonus depreciation, lowers the firm's depreciable basis by 50%, thereby lowering annual depreciation costs and increasing cash flow.[22]

Based on IRS property requirements, a second-generation biofuel producer may elect to deduct half of the firm's capital conversion costs including components of refining, packaging, crushing and plant infrastructure such as the boiler, cooling tower and fire protection. Capital transportation costs are not included in this depreciation allowance. Assuming a 13-million-gallon production, the depreciation allowance reduces the per gallon capital costs from $3.93 to $1.97 per gallon. Following a straight-line depreciation method, the operating costs are also adjusted to account for the reduction in ongoing depreciation costs.

Second-generation biofuel producers also qualify for a $1.01 tax credit for each gallon produced. This results in a operating credit of $13,130,000, thereby reducing per gallon operating costs by $1.01. However, the operating costs are also reduced by the difference in per unit depreciation costs that results from the depreciation allowance. With no depreciation allowance, the firm's

[22] Firms can take a 50% special depreciation allowance for qualified second-generation biofuel plant property. To qualify for the depreciation allowance, the property must meet the following requirements. (i)The property is used in the United States solely to produce second-generation biofuel. (ii)The original use of the property must begin with you after December 20, 2006. (iii) Property must have been acquired by purchase after December 20, 2006, with no binding written contract for acquisition in effect before December 21, 2006. (iv) The property must be placed in service for use in trade or business or for the production of income before January 1, 2017.

ongoing depreciation costs are $3,409,002 or $0.26 per gallon. After adjusting for the depreciation allowance the ongoing per unit depreciation costs are $0.13 per gallon. Therefore, the total reduction in operating costs is $1.14, resulting in a new operating cost of $3.21 per gallon.

Under the subsidized capital and operating costs, the firm's break-even entry and exit prices are found, using the same approach as previously shown, to be $3.32 and $3.17 respectively. However, when accounting for the annualized growth rate of RIN prices, the two-variable break-even entry and exit prices are found to be $2.70 and $2.58 respectively. The break-even entry and exit prices do not account for the uncertainty in conventional fuel price, RIN price and their correlation as previously discussed. To account for these uncertainties, the subsidized capital and operating costs are analyzed in the two-variable real option framework.

Figure 30 illustrates the policy effects on the firm's optimal entry and exit trigger prices. Here it can be seen that the depreciation allowance and tax credit partially counteract the uncertainty effects of the RIN market.

The subsidized entry trigger of $4.65 is close to the single variable break-even entry price of $4.61. However, the upside of RIN price uncertainty still exists in the form of a lower exit trigger. The subsidized exit trigger price of $2.63 is approximately $1.70 lower than the single variable break-even exit price.

The policy incentives do indeed lower the entry and exit thresholds and partially counteract RIN price uncertainty effects. It appears then; the two policy incentives are important tools to incentivize investment in second-generation biofuels, particularly in the face of highly volatile RIN markets. However, the policy effect is not enough to put the entry and exit trigger prices on

par with the two-variable break-even model. To reach parity with the two-variable break-even result, policymakers would need to increase the depreciation allowance and the per unit tax credit. Presumably a per unit tax credit is costlier as the credits may be claimed for as long as the policy is active. Whereas the depreciation allowance can only be claimed in the first year of operations. Therefore, policymakers ought to be interested in the least costly mix of these two incentives to reach the desired policy effect. Finding the least costly mix of policy incentives would imply that the tax credit should be minimized so that in conjunction with the depreciation allowance just satisfies the desired policy effect. To approach the entry price of $3.75 in Table 26 policymakers could cease the depreciation incentive and institute a per unit tax credit of $2.20 which would result in an unchanged capital cost of $3.93 and a subsidized operating cost of $2.15 per gallon. This scenario would result in an entry trigger price of $3.83 in the two-variable real option model. However, the same result is reached with an alternative mix of policy incentives. By increasing the depreciation allowance to 60% and the tax credit to $1.44 per gallon, the subsidized capital and operating costs become $1.58 and $2.75 respectively. Under this mix of policy incentives, the two-variable real option model again predicts an entry trigger price of $3.83 per gallon.

Analyzing the efficiency and equity effects of the tax credit and depreciation allowance is beyond the scope of this study. However, under the assumption that capital is fixed in the short term, tax incentives are likely to divert capital away from more productive uses toward the tax favored uses (Guenther 2015). Whether or not the welfare benefits outweigh the potential losses associated with the capital diversion remains a topic for future studies. This study provides a framework for analyzing the efficacy of second-generation biofuel producer incentives, but

makes no claims about the efficiency, equity, administration or overall effect on economic growth.

2.6. Conclusion

This study separates and analyzes the effect that RFS and blender's tax credit uncertainty has on firm investment decisions. Expiration and reinstatement of blender tax credits, along with annual delays, revisions and waivers issued by the EPA create uncertainty throughout the biofuels supply chain. This uncertainty leads to volatility in the price of RINs, which in turn increases the aggregate uncertainty faced by producers of second-generation biofuels.

Using a two-stochastic-variable real option model, the effect of policy uncertainty is isolated from general market uncertainties. In this framework, the price, which triggers the firm to invest and enter the market, is identified. Numerical solutions also provide threshold curves, which represent all combinations of conventional fuel price and RIN price, which induce the firm into action.

Investigating how the price volatility of both conventional fuel price, RIN price and their correlations effect the firm's decision to enter the market, remain idle or to abandon the market if already active, provides important insights into an industry, which may be critical to long term energy independence and national security. Furthermore, results could have important policy implications. If high levels of price and policy uncertainty depress investment in second-generation biofuels, then perhaps, the RFS would be more effective by preventing delays in annual announcements, limiting annual revisions, and setting the volume obligations over a longer horizon. Any effort in these areas would work to reduce policy uncertainty that has been characteristic of the RFS.

Results indicate the price, which triggers the firm to invest and enter the market, is significantly higher under the real options model with two stochastic variables. The underlying cause is price volatility and irreversible capital costs. In the market for second-generation biofuels, the price volatility faced by potential market entrants comes in the form of both conventional fuel price and RIN price. Conventional fuel prices have historically exhibited large volatilities, but biofuel policy uncertainty creates high volatility in RIN price adding to the aggregate uncertainty faced by potential market entrants. Clearly, the RIN program has successfully encouraged investment in biofuels. However, high levels of policy uncertainty lead to volatile RIN prices, which work to limit the effectiveness of the RIN program.

The increased aggregate uncertainty creates additional headwinds for second-generation biofuel producers. Investment incentives such as bonus depreciation allowances and producer tax credits work to counteract uncertainty effects, but these policy incentives would be more effective in stimulating investment if the firm faced less aggregate uncertainty.

2. A. Appendix

2. A.1 Tables

Table 20: OLS results of D4 RIN price

Ordinary Least Squares result for the equation: $\Delta lnR_t = c + \sum_{i=1}^{p-1} \varphi_i \Delta lnR_{t-i} + \epsilon_{R_t}$

Dependent						
ΔlnR_t	Coef.	Std. Err.	t	P > t	[95% Conf. Interval	
Independent						
ΔlnR_{t-1}	0.08366	0.0254	3.300	0.001	0.033927	0.1333873
ΔlnR_{t-2}	0.01104	0.0254	0.430	0.664	-0.03886	0.060936
ΔlnR_{t-3}	0.07701	0.0253	3.040	0.002	0.027295	0.1267253
c	0.00099	0.0012	0.810	0.417	-0.00141	0.0033868
Observations	1543					
F(3,1539)	7.16					
Prob > F	0.0001					
Root MSE	0.04793					

Table 21: OLS results of conventional fuel price

Ordinary Least Squares result for the equation: $\Delta lnP_t = c + \varphi_1 \Delta lnP_{t-1} + \epsilon_{P_t}$

Dependent						
ΔlnP_t	Coef.	Std. Err.	t	P > t	[95% Conf. Interval	
Independent						
ΔlnP_{t-1}	-0.0209821	0.025505	-0.82	0.411	-0.07101	0.0290455
c	-0.0004234	0.000469	-0.90	0.367	-0.00134	0.0004969
Observations	1,545					
F(1, 1543)	0.680					
Prob > F	0.41080					
Root MSE	0.01844					

Table 22: Parameter estimation results

Parameter	Definition	Value	Scale
α_p	Drift Rate Conventional Fuel	0.0027	Per year
α_r	Drift Rate D4 RINs	0.0182	Per year
σ_p	Standard Deviation Conventional Fuel	0.2927	Per year
σ_r	Standard Deviation RINs	0.7609	Per year
ρ_{pr}	Correlation of Fuel and RINs	-0.0129	Per year

Table 23: Cost parameters sourced from prior literature

Parameter	Definition	Value	Scale	Source
δ	Discount Rate	0.10[23]	per year	(Brown and Brown 2013)
w	Operating Cost	4.35	per gallon	(English et al., 2016)
k	Capital Cost	3.93	per gallon of total capacity	(English, Menard et al. 2016)
s	Salvage Value	0.98	per gallon of total capacity	(Schmit, Luo et al. 2009)

Table 24: Comparison of entry and exit prices

Method	Entry	Exit
Single Variable Break-Even	$4.6149	$4.3281
Two Variable Break-Even	$3.7517	$3.5183
Single Variable Real Option	$5.3141	$3.7921
Two Variable Real Option	$6.4646	$3.4343

Table 25: Comparison of production costs with tax incentives

Pennycress fed Green Jet Fuel Production Pathway		
	Costs - with Tax Incentives	Costs-No incentives
Name Plate Capacity (gallons)	13,000,000	13,000,000
Capital Transportation ($)	$118,915	$118,915
Capital Conversion ($)	$25,494,055	$50,988,111
Total Capital Costs ($)	$25,612,970	$51,107,026
Total Capital Costs ($/gallon)	$1.97	$3.93
Operating cost Conversion ($)	$10,804,002	$12,508,503
Operating Farmgate Cost of Feedstock ($)	$43,335,000	$43,335,000
Operating Cost of Transportation ($)	$666,396	$666,396
Operating Credit ($)	-$13,130,000	$0
Total Operating Cost	$41,675,398	$56,509,899
Total Operating Costs ($/gallon)	$3.21	$4.35

[23] A discount rate of 10% is a common assumption in Techno-Economic Analyses. It is chosen as the rate, which will make the present value of cash proceeds equal to the present value of cash outlay's over the life of the plant. In other words, this is the break-even internal rate of return for the production pathway under study in the TEA.

Table 26: Comparison of entry and exit prices with current policy incentives

Method	Entry	Subsidized Entry	Exit	Subsidized Exit
Single Variable Break-Even	$4.61	$3.32	$4.33	$3.17
Two Variable Break-Even	$3.75	$2.69	$3.52	$2.58
Single Variable Real Option	$5.31	$3.77	$3.79	$2.81
Two Variable Real Option	$6.46	$4.65	$3.43	$2.63

2. A.2 Figures

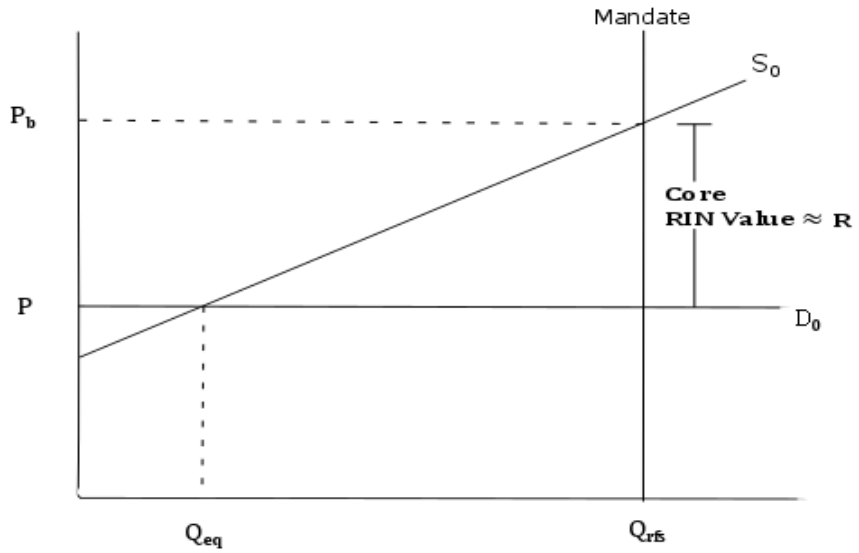

Figure 16: Core Value of the RIN

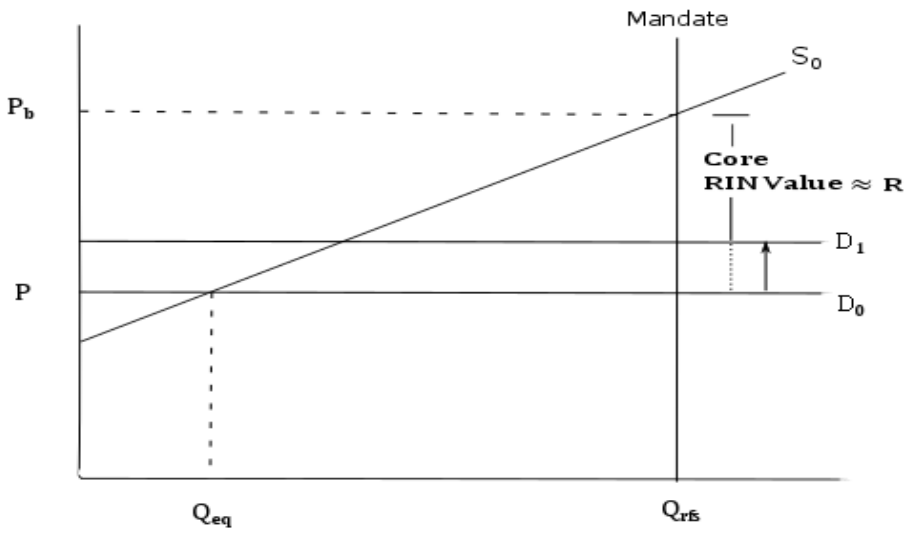

Figure 17: Outward demand shift causes decrease in RIN value

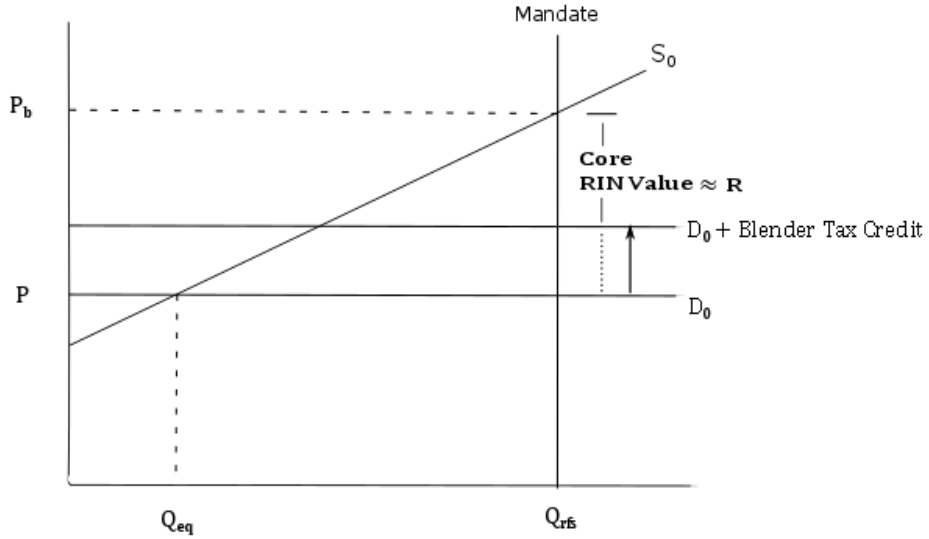

Figure 18: Effect of a blender tax credit on the core value of a RIN

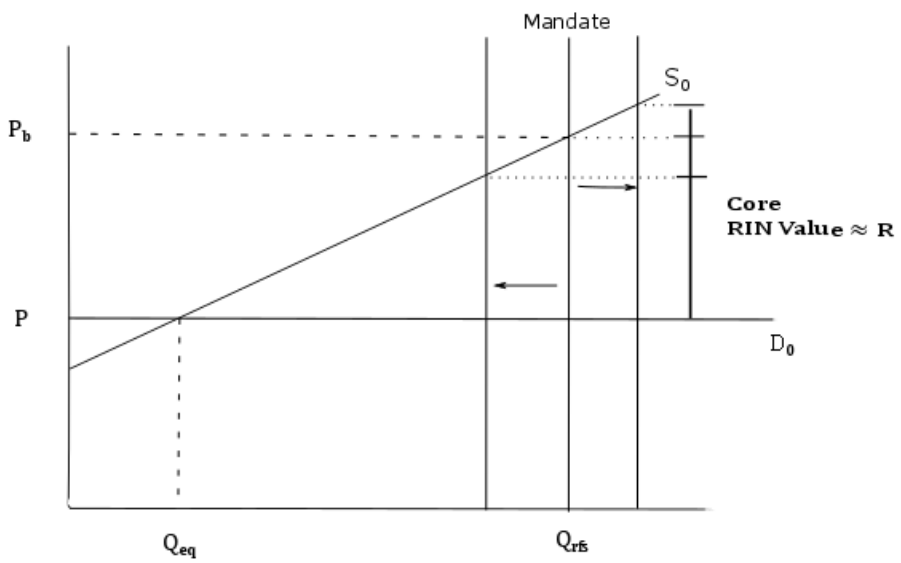

Figure 19: Effect of a shift in the RFS mandate on the core value of a RIN

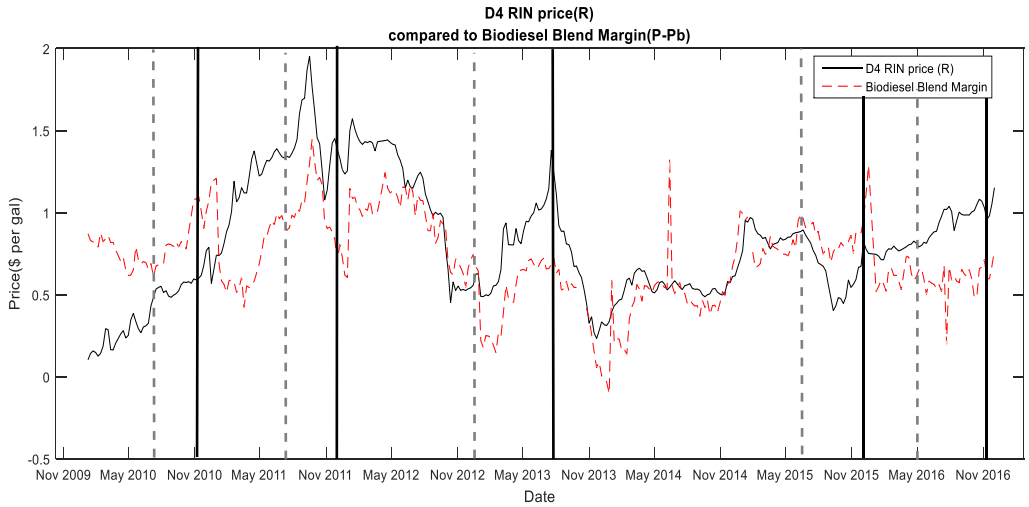

Figure 20: Historical biodiesel RIN prices and biodiesel blend margin

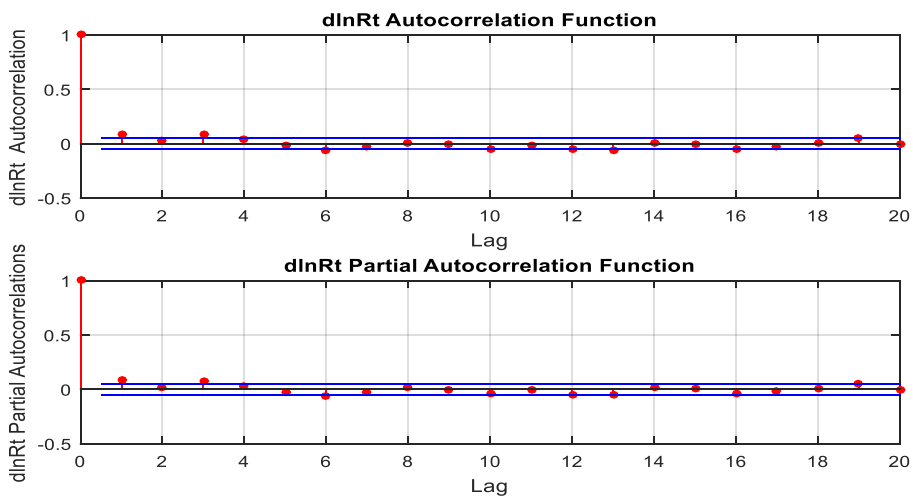

Figure 21: ACF and PACF of ΔlnR_t

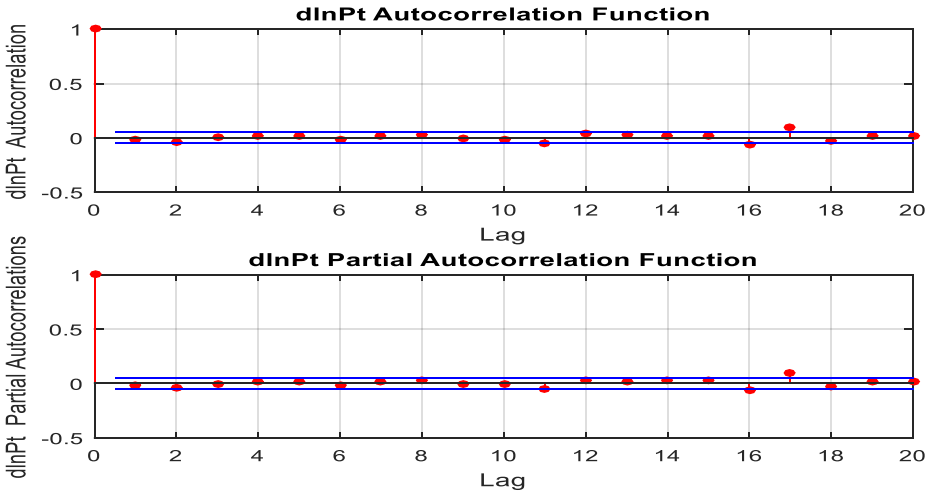

Figure 22: ACF and PACF of ΔlnP_t

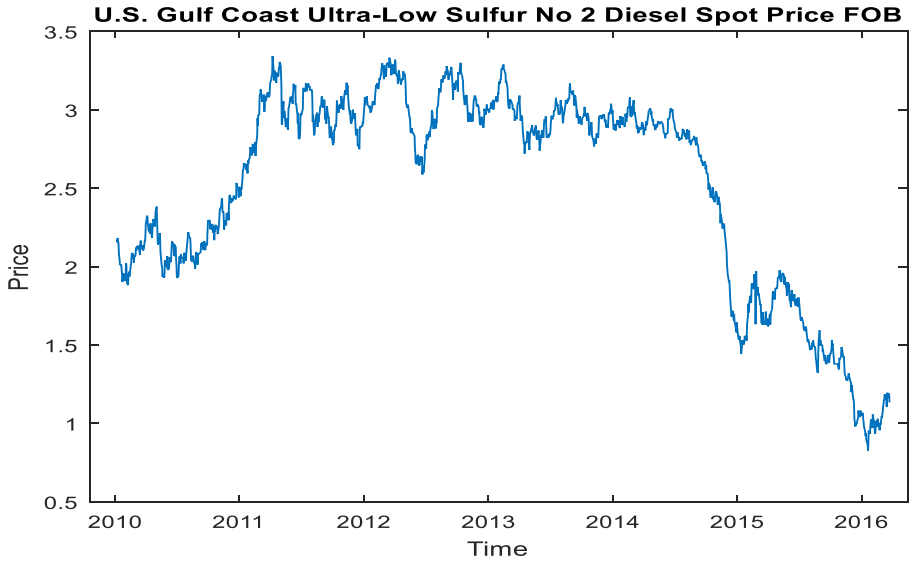

Figure 23: Historical Diesel Price

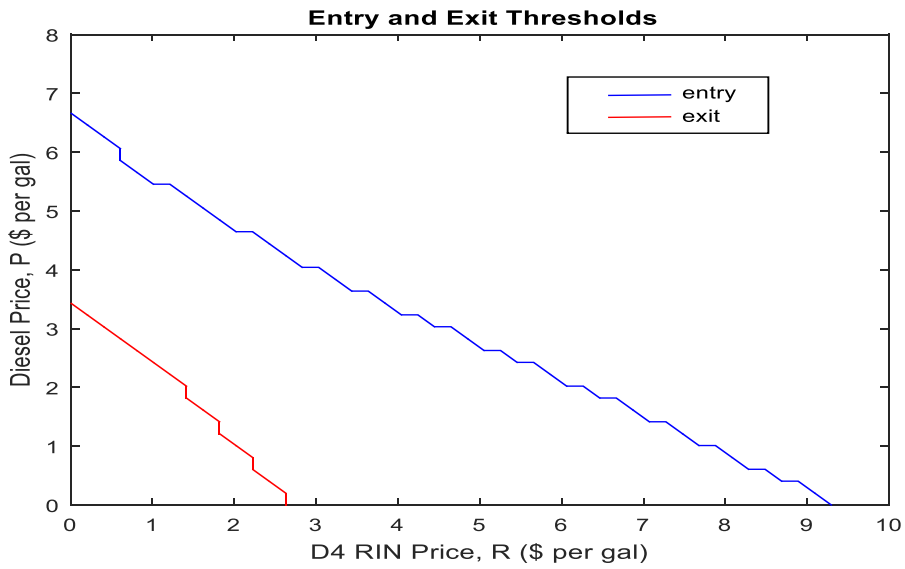

Figure 24: Two-variable entry and exit threshold curves

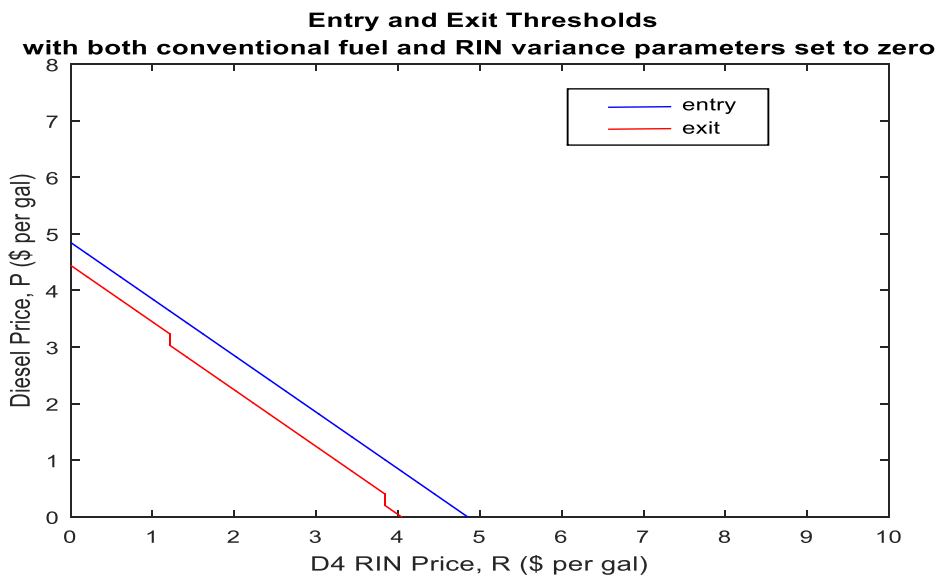

Figure 25: Two-variable entry and exit threshold curves with fuel price and RIN price variance set to zero

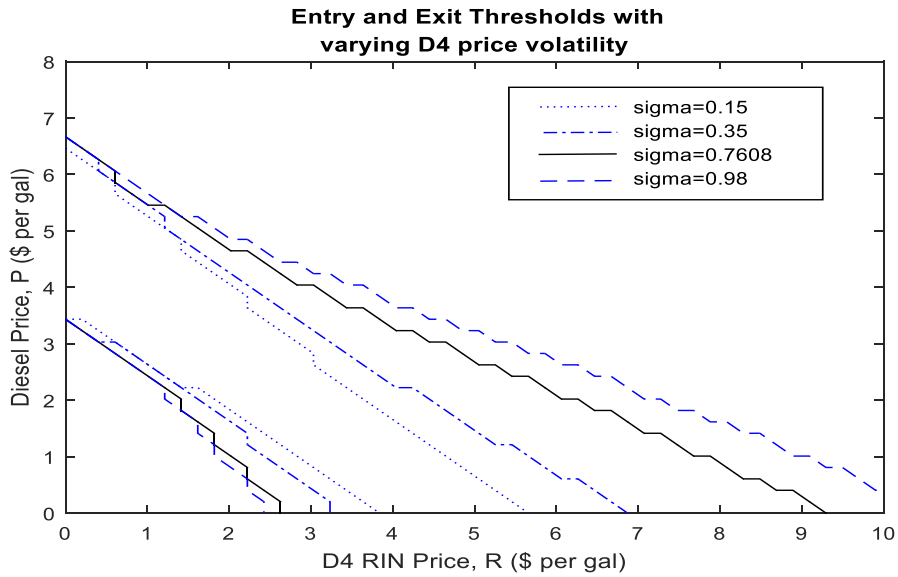

Figure 26: Threshold curve sensitivity holding conventional fuel price volatility constant

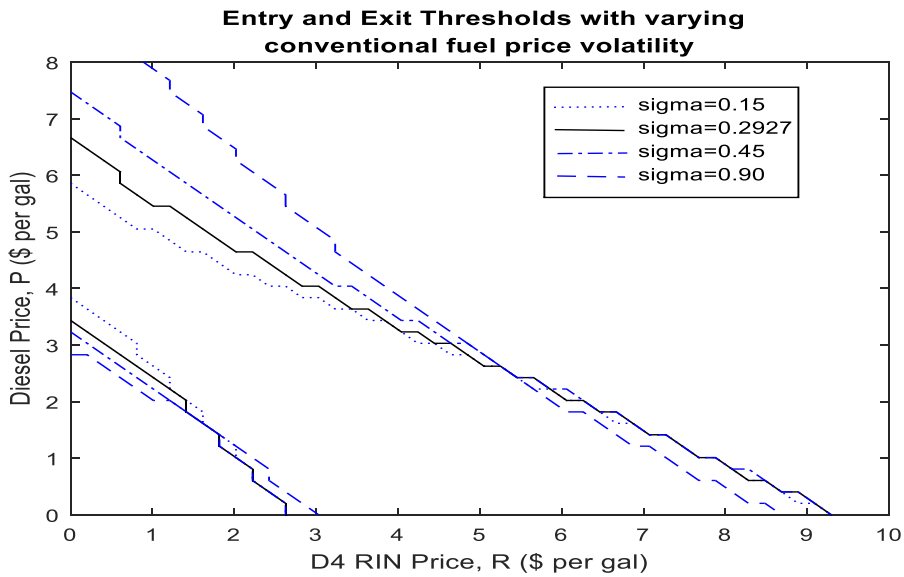

Figure 27: Threshold curve sensitivity holding RIN price volatility constant

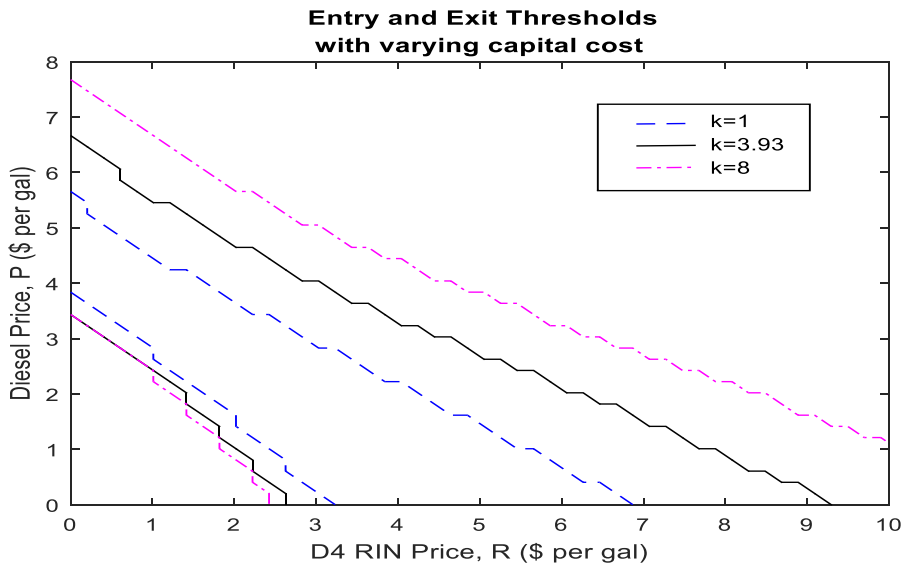

Figure 28: Effect of changes in the firms per gallon capital cost on entry and exit thresholds

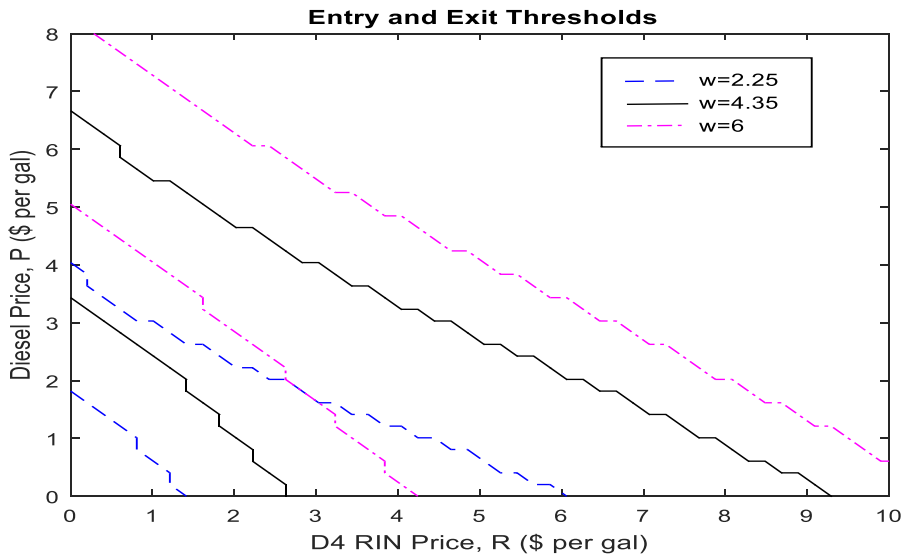

Figure 29: Effect of changes in the firm's operating cost on entry and exit thresholds

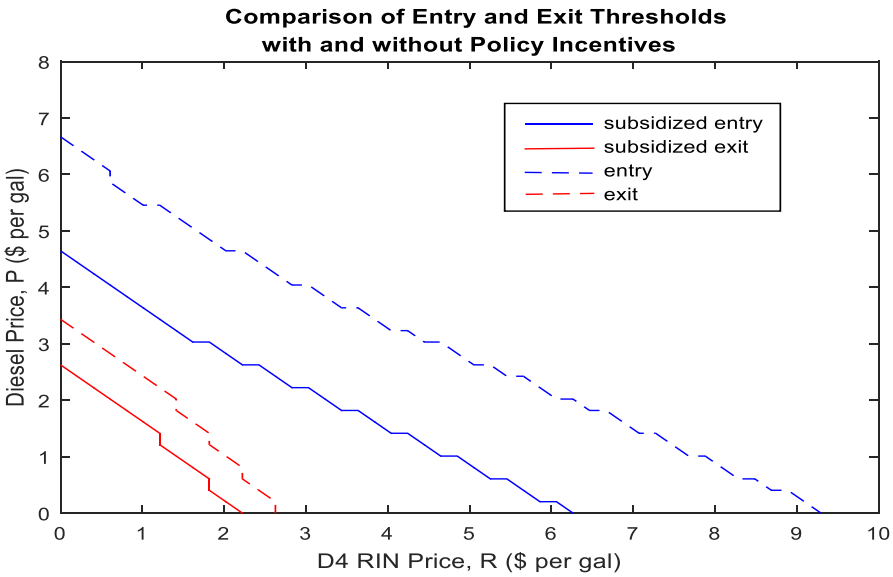

Figure 30: Entry and exit threshold curves while accounting for producer policy incentives

Chapter III: Potential for Pennycress to Support an Alternative Jet-Fuel Industry

Chapter III: Abstract

This study examines the economic feasibility in pennycress production and its potential to supply a renewable aviation industry. Pennycress is an oilseed plant that is being considered as a second-generation biofuel feedstock. The plant produces an oilseed, which is of high quality and high oil content. The plant lifecycle fits with traditional crop rotations such as a corn and soybean rotation. Therefore, pennycress may provide soil and environmental benefits associated with cover crops while also providing a cash benefit to producers in the form of a valuable oilseed. Pennycress has never been produced at commercial scale. Therefore, this study provides an economic analysis at the producer level and simulates county level supply curves across the contiguous United States, using a using a partial equilibrium simulation model of the US agricultural sector. County level results are then aggregated to the National level to estimate the potential supply of alternative jet-fuel produced from pennycress oil.

3.1. Introduction

The commercial aviation industry consumed over 16.2 billion gallons of jet fuel in 2014, accounting for approximately 25% of operating expenses (U.S. Department of Transportation BTS 2014). With such a large proportion of operating expenses coming directly from fuel consumption, fuel price uncertainty can have a substantial impact on airline operating strategies. Airlines use various strategies to manage fuel price uncertainty including adjustments in aircraft size and utilization, flight-route and destination offerings, vertical integration and financial hedges (Davidson, Newes et al. 2014). The use of renewable jet fuel has the potential to decrease fuel-price uncertainties, thereby decreasing fuel-hedging costs, and decreasing profit volatility. The United States Government is actively supporting the development of renewable jet fuels. Two examples of this are the Farm-to-Fly 2.0 program, which focuses on supply chain and infrastructure, and the Commercial Aviation Alternative Fuels Initiative (CAAFI), which supports research and development, environmental assessment, fuel testing, demonstration and commercialization. As of 2015, the standard setting organization, ASTM International has approved three renewable jet fuels for use in aviation (USDOT FAA 2015).

This paper examines the potential for pennycress to support a renewable jet fuel industry in the United States. Pennycress (Thlaspi Arvense) is commonly referred to as “stinkweed” or “French-weed” and found throughout the United States (CABI 2016). Pennycress is being evaluated as a potential feedstock for biodiesel and biojet fuels. The crop has the potential to supply both oil and biomass to the biofuels market. Following harvest, seed crushing and pre-processing; pennycress offers a suitable oil to allow conversion to a Hydro-processed Ester and Fatty Acid (HEFA) fuel. Pennycress provides the benefits of a winter cover crop while also supplying producers with direct economic benefits. As a winter cover in double crop rotations, pennycress

is established in the fall and harvested in mid-to-late-spring. Pennycress can offer the benefits of weed reduction, excess nutrient uptake and management, and hydrological regulation by aiding in the drying of excessive soil moisture. The plant is grown on fallow winter land, providing a valuable cash crop which avoids the pitfalls of the food-versus-fuel debate (Moser, Knothe et al. 2009).

The U.S. Environmental Protection Agency (EPA) recently conducted an assessment of greenhouse gas emissions attributable to the production and transportation of pennycress oil, to be used in the making of biodiesel and biojet fuels. The EPA indicates that based on its analysis, pennycress could qualify as biomass-based diesel or advanced biofuel if typical fuel production process technologies are used (USEPA 2015). The EPA anticipates approval for the generation of renewable identification numbers, following existing rule making related to camelina feedstock production. Pennycress is in the brassica family, as is camelina, so while no new rule making is expected, pennycress producers will likely obtain approval through the EPA's feedstock pathway petition process. Approval of production pathways for the generation of RINs is currently being conducted on a case-by-case basis. The approval will likely include adherence to invasiveness monitoring and reporting requirements, as the plant has some potential for invasiveness (USEPA 2015). Indeed, there are currently nine states with restrictions of pennycress cultivation due to the potential for invasiveness. These states include, IN, KS, MN, NE, NV, OH, SD, and WA. Previously Michigan had been included in this list until it was deregulated in 2015 (USDA-APHIS 2015). Despite having a high-risk classification, pennycress is not likely to have an impact when incorporated as a component of traditional row crop rotations. This is because herbicide applications that occur during the corn or soybean planting provide adequate control of any leftover pennycress seed (Sindelar, Schmer et al. 2015).

3.2. Farm-Gate Costs

Large-scale cultivation has not yet been undertaken for pennycress, thus management practices and stand establishment practices are still in research and development. Variety trials have been undertaken throughout the Midwest, particularly in Western Illinois, Minnesota, and Nebraska. A variety of planting methods have been tested, including drill, broadcast and aerial seeding. Aerial seeding over standing corn canopy or a broadcast with light incorporation after corn harvest has been successful. Aerial or broadcast seeding rates of 5lbs/acre have generated yields ranging from approximately 1400lbs/acre to 2200lbs/acre with an average yield of1800lbs/acre. To establish a stand of pennycress ready for harvest in May-June and not interfere with soybean planting, aerial seeding over corn canopy may generate the best results. However, planting earlier than September 1 should be avoided, because if the plant flowers too early in winter months, survival rates are low. Phippen, John et al. (2010), conducted a study of five different September planting dates and found that seed oil content was not significantly different among the various planting dates. However, stand establishment declined as temperatures declined throughout the fall.

Application of nitrogen and sulfur requirements are found to be minimal (Moser, Knothe et al. 2009). Growth chamber experiments show, seed yield is not significantly influenced by the application of nitrogen above 50lbs per acre (Mason, Phippen et al. 2013). Some current field trials have not required the use potassium or phosphorus, but commercial scale operations may require the application of these fertilizers to replace loss nutrients (USEPA 2015).

Plants of the brassica family are susceptible to the fungus following an unusually wet spring. Insects and pests however, are not likely to pose a risk. This is because the pennycress plant completes its life cycle in the spring when insects are emerging.

There is not yet an established market for pennycress and commercial scale production has not been widely established. However the Agricultural Marketing Resource Center provides an estimate of $0.15 per pound of seed (AGMRC 2015). Enterprise budgets are developed to estimate the per acre cost of production for pennycress, based on a yield of 1600 (Table 37). Based on numerous research trials a yield of 1600 pounds per acre seems reasonable and conservative (Fan, Shonnard et al. (2013), Sedbrook, Phippen et al. (2014), Agricultural Marketing Resource Center (AGMRC) (2015)).

No insecticide or herbicide applications are required for a successful pennycress crop. As a member of the mustard family and a winter annual, insect pressure is insignificant due to its natural chemistry and the temperatures of the growing season thus limiting an insecticide requirement (Arvens Technology Inc 2010). Crop insurance for Pennycress is not currently available and not included in variable costs. Therefore, variable costs include those costs for seed, nitrogen, repair and maintenance, fuel and lube, labor, operating interest, and machinery rental. Fixed costs include capital recovery costs for machinery and equipment, which are combine, with grain head, 215hp tractor with grain cart and semi-tractor trailer for hauling.

Based on the production of 1600 pounds per acre and a market price of $0.15 per pound, estimated revenues are $240 per acre. Seed costs are also projected by the AGMRC and estimated to be $2.50/ 5lbs. Total variable expenses are estimated to be $67.47 per acre assuming aerial seeding of five pounds per acre. Total Fixed costs are estimated to be $18.24 per acre. Breakeven prices assuming a yield of 1600 pounds per acre is estimated to be $0.05 per pound, 10 cents below the estimated market price provided by AGMRC. Break even yield for a market price of $0.15 per pound is estimated to be 571 pounds per acre. The break-even price for a yield

of 1600 pounds per acre does not account for the uncertainty and sunk costs in the decision to invest in pennycress production.

3.3. A Model of Optimal Decisions in the Production of Pennycress

Pennycress has never been commercially produced so there are significant risks involved the enterprise. Not surprisingly, a large portion of that risk stems from uncertainty in yields and price. There is currently no market for pennycress; so naturally, there are no prevailing prices, which producers can gauge, and no historic price data, which can be used to form expectations. Similarly, there are no historic yield data and it is quite possible that yield results from experimental plots will not translate to yield results under commercial conditions. Furthermore, there is considerable risk to the yields of the follow-on crop, such as soybeans. If the harvesting of pennycress is later than expected, it could delay the planting of soybeans, which could affect soybean yields (UIUC ACES 2013). Previous research has found mixed results in this area. Johnson, Kantar et al. (2015), found the effect of pennycress on soybean yields to be dependent upon location. Planted, following the corn harvest, pennycress can be harvested in late May to mid-June. Sindelar, Schmer et al. (2015), have conducted a review of 17 studies where soybean planting was delayed. The authors found that all but four locations resulted in a yield reduction, varying from -4% to -28%.

Phippen and Phippen (2012), also note that planting soybean could be delayed due to poor field conditions and/or late pennycress harvest. Furthermore, the authors find that a delayed planting of soybean until mid-June results in reduced oil content and a 20% reduction in soybean yields. However, the yield penalties associated with delayed planting may be offset by accrued soil benefits in subsequent years. The authors conducted a two-year study to examine the effect of pennycress residues on soybean yields. In the first year, soybean yields were only affected by

planting date. In the second year, soybean yields were significantly affected by pennycress residues, but the effect was positive. Two years of pennycress residue resulted in greater moisture levels and other benefits to the soil. This led to increased soybean yields in the second year. The fact that pennycress residue benefits did not materialize until year 2 is in line with a large body of cover crop research. Cover-crop benefits accrue gradually and are typically not observed in the first year (Clark 2012).

Probabilistic yield penalties translate to an irreversible cost to investing in the production of pennycress. However, in years two and beyond, pennycress residue effects take hold. Soil benefits accrue, and year-two yield penalties are offset. Therefore, a producer who invests in the production of pennycress can expect to incur a one-time sunk cost, stemming from the first-year yield penalties.

The irreversible nature of probabilistic yield penalties, coupled with uncertain pennycress revenues create an option value in delaying investment in pennycress production. Producers may wish to delay investment until uncertainties are resolved or until the expected gains from investing now, outweigh the value of delaying into the future. The producer can be in one of two states, either in traditional crop rotation where $(i = 0)$ or in a pennycress double-cropping rotation where $(i = 1)$. The producer decides when, if ever, to switch states from traditional rotation to double-cropping rotation by choosing t_1. If the producer is already double-cropping pennycress, they must decide when, if ever, to abandon pennycress production by choosing t_0. When the producer is actively double-cropping pennycress, they earn an additional flow of payments $\pi_i(P, Y)$ which is defined as

$$\pi_{i=0,1}(P, Y) = \begin{cases} 0 & if\ no\ pennycress\ is\ produced\ and\ i = 0 \\ P \cdot Y - w - c & from\ the\ production\ of\ pennycress\ and\ i = 1 \end{cases} \quad (4)$$

In (4), P is the price of pennycress per pound of seed, Y is the quantity yielded per acre, w is the per acre operating costs and c is the per acre capital recovery costs of machinery used in the planting and harvesting of pennycress. As discussed above, the first-year yield penalties to soybean represent a sunk cost (k) that the farmer expects to incur when investing in pennycress. This is treated as a one-payment the farmer must make in order to invest in the production of pennycress. However, the decision to plant pennycress is an annual decision. Assuming the farmer is free of any binding contractual obligations, they may choose to forego subsequent pennycress production while incurring no costs to this decision to abandon. Therefore, the farmer's investment cost is k and the exit cost is zero.

Given the discount rate δ, the farmer's optimal entry and exit decision satisfies

$$V_0(P_0, Y_0) = \max_{t_1} E_0 \left\{ \int_0^{t_1} \pi_0(P(t), Y(t)) e^{-\delta t} dt + \ [(V_1(P(t_1), Y(t_1)) - k) e^{-\delta t_1}] \right\} \quad (5)$$

$$V_1(P_0, Y_0) = \max_{t_0} E_0 \left\{ \int_0^{t_0} \pi_1(P(t), Y(t)) e^{-\delta t} dt + \ [(V_0(P(t_0), Y(t_0))) e^{-\delta t_0}] \right\} \quad (6)$$

subject to dP, dY and $P(0) = P_0,$ and $Y(0) = Y_0$.

3.3.1 Market Dynamics for Pennycress

The price of pennycress seed, P, follows a generalized Ito process such that the change in price

$$dP = \ \alpha_p(P)dt + \sigma_p(P)dz_p \quad (7)$$

where $\alpha_p(P)$ is the deterministic trend in the price process and $\sigma_p(P)$ represents the instantaneous standard deviation. The trend changes over the time increment dt . The change in z_p , dz_p , follows a Wiener process such that $dz_p = \ \varepsilon_t \sqrt{dt}$ where, ε_t is a normally distributed

random variable with a mean of zero and a standard deviation of unity. Furthermore ε_t is serially uncorrelated so that $E(\varepsilon_t, \varepsilon_s) = 0 \ \forall \ t \neq s$ and thus the values of dz_p for any two different intervals of time are independent, following a Markov process with independent increments. It is also assumed that the discount rate δ is greater than the drift rate α_p, which indeed must hold otherwise investment would never be optimal, as the growth rate would outpace the discount rate. Hence, it would always be possible to do better by waiting longer.

Pennycress yields are also assumed to follow a generalized Ito process, just as is the case for the price of pennycress.

$$dY = \alpha_y(Y)dt + \sigma_y(Y)dz_y \tag{8}$$

where $\alpha_y(Y)$ is the deterministic trend in the price process, $\sigma_y(Y)$ represents the standard deviation and the same assumptions of equation (7) apply. The two stochastic processes are potentially correlated in their Wiener processes such that $E[dz_P dz_y] = E\left[\varepsilon_{tP}, \varepsilon_{ty}\right]dt = cov(dz_P, dz_y)/\sqrt{var(dz_P)var(dz_y)} = \rho_{py}dt$.

3.3.2 The Farmer's Decision to Produce Pennycress

The farmer who remains in traditional crop rotations has expected net present value denoted as $V_0(P, Y)$. Since the farm is in traditional crop rotation, $V_0(P, Y)$ represents the value of the option to produce pennycress. In this state, the only additional value stemming from pennycress is the expected gain in the value of retaining the option to invest denoted as $E_t[dV_0(P, Y)]dt^{-1}$.This is also referred to as the capital appreciation of the farmer's asset as prices and yields increase, where in this case, the asset is the farmer's option to begin pennycress production.

The required return from the investment is represented by the function $\delta V_0(P,Y)$and the no-arbitrage condition of efficient markets sets these two returns equal.

$$\delta V_0(P,Y) = E_t[dV_0(P,Y)]dt^{-1} \tag{9}$$

Under the efficient markets theorem, equation (9) must hold and implicitly defines the price and yield combinations, which trigger entry and exit. Equation (9) essentially says that over an infinitesimal period dt, the required return on the investment opportunity is equal to its expected rate of capital appreciation and facilitates the evaluation of the affect that changes in output price and yield have on the value of a farmer who has not yet adopted pennycress.

Following Ito's Lemma, $dV_0(P,Y)$ is expanded using a Taylor series expansion. It is assumed that P and Y are both continuous-time stochastic processes as represented by equation (7) and (8). Through Ito's Lemma, equation (9) becomes the second-order partial differential equation

$$\frac{\sigma_p(P)\sigma_r(Y)\partial^2 V_0}{\partial P \partial Y}(\rho_{py}) + \frac{\sigma_p(P)^2}{2}\frac{\partial^2 V_0}{\partial P^2} + \frac{\sigma_y(Y)^2}{2}\frac{\partial^2 V_0}{\partial Y^2} + \frac{\partial V_0}{\partial P}\alpha_p(P) + \frac{\partial V_0}{\partial Y}\alpha_y(Y) - \delta V_0(P,Y) = 0 \tag{10}$$

3.3.3 The Farmer's Decision to Exit

The value of the farmer who is actively producing pennycress is denoted as $V_1(P,Y)$, which is the expected discounted value of the net earnings plus their option to exit. Over the range of prices P and yields Y, where it is optimal to remain in production of pennycress, the project is earning $(P * Y - w - c)$. The farmer receives a dividend equal to the additional flow of profits earned from producing pennycress $(P * Y - w - c)$ and capital gains arise from the expected future appreciation of the farm operation. Similar to equation (9) efficiency in markets requires

$$\delta V_1(P,Y) = (P * Y - w - c) + E_t[dV_1(P,Y)]dt^{-1}. \tag{11}$$

Applying Ito's Lemma to expand $dV_1(P,Y)$ equation (11) results in the second-order partial differential equation

$$\frac{\sigma_p(P)\sigma_r(Y)\partial^2 V_1}{\partial P \partial Y}\left(\rho_{py}\right) + \frac{\sigma_p(P)^2}{2}\frac{\partial^2 V_1}{\partial P^2} + \frac{\sigma_r(Y)^2}{2}\frac{\partial^2 V_1}{\partial Y^2} + \frac{\partial V_1}{\partial P}\alpha_p(P) + \frac{\partial V_1}{\partial Y}\alpha_y(Y) - \delta V_1(P,Y) = c - w - P * Y. \quad (12)$$

Solving (10) and (12) will implicitly define the price/quantity combination, which induce the farmer to enter and exit production of pennycress. Partial differential equations such as (10) and (12) often require the use of numerical techniques to approximate a solution. The same numerical methods detailed in 2.4.3 are applied here to approximate the solution. Before this approach can be carried out, the model must be parameterized.

3.3.4 Parameterization of the model

To estimate the parameters of the Ito process, one must first determine which Ito process best fits the data. Geometric Brownian motions (GBM) in yield and price data are first considered. There is currently no market for pennycress, and thus no historical price or yield data. However, pennycress is similar to that of industrial rapeseed and there is a well-established global market for rapeseed. Therefore, historical rapeseed prices and yields are used to estimate drift and variance parameters. Rapeseed yields vary dramatically in different parts of the world. For example, German rapeseed yields have reached nearly 4000 pounds per acre in recent years. However, Canadian rapeseed yields recently reached an all-time high of 2000 pounds per acre. Agronomic research on pennycress places estimated yields to be in the range of 800 pounds per acre, up to 2200 pounds per acre. Therefore, historical Canadian yields are used here. Historical price data from global spot markets is utilized to represent a matured global market for

pennycress seed. Therefore, the historical spot price for rapeseed free on board (FOB) at Rotterdam is utilized.

To determine if GBM is the appropriate process for the chosen data, unit root tests are carried out. In both cases, the null hypothesis of a unit root is rejected at the 99% confidence level. Therefore, GBM is excluded from the analysis and geometric mean reversion (MR) processes are the next to be considered. The geometric MR process is written as

$$\frac{dP}{P} = \lambda_p(\bar{P} - P)dt + \sigma_p dz_p \tag{13}$$

where λ_p is the rate of reversion, σ_p is the standard deviation, $\bar{P}$ is the long run average and the change in z_p , dz_p . Recall that dz_p follows a Wiener process such that $dz_p = \varepsilon_t\sqrt{dt}$ where, ε_t is a normally distributed random variable with a mean of zero and a standard deviation of unity. To estimate the rate of reversion parameter, standard deviation and long run mean, one can rearrange (13) to current period form.

$$\frac{P_t - P_{t-1}}{P_t} = \lambda_p(\bar{P} - P_t) + \sigma_p\varepsilon_t \tag{14}$$

Ordinary least squares can be used to estimate (14) in the form of $y = a + bx + e$. In performing the regression, set $y = \frac{P_t - P_{t-1}}{P_t}$ and allow $x = P_t$. From this, one can see that

$$a + bx + e_t = \lambda_p\bar{P} - \lambda_p P_t + \sigma_p\varepsilon_t$$

The parameters of interest are recovered such that $\hat{\lambda}_p = -\hat{b},\ \bar{P} = \frac{\hat{a}}{\hat{\lambda}_p}$ and $\hat{\sigma}_p = sd(\hat{e}_t)$. The rate of reversion and the standard deviation are then annualized using the convention of converting 1-day parameters to h-day parameters. This requires scaling by a factor of $\sqrt{h}$ resulting in a reversion rate of $\hat{\lambda}_p = 0.0101 * \sqrt{252} = 0.1603$ and a standard deviation of $\hat{\sigma}_p = 0.0130 *$

$\sqrt{252} = 0.2064$. This estimation procedure is the same for the historical rapeseed yields as well. The rapeseed data however, are already in annual terms so there is no need to scale by a factor of $\sqrt{h}$. Next, the correlation coefficient for yield and price are estimated after annualizing price. Doing so results in a weakly negative relationship between yield and price with a correlation coefficient of -0.2943. Estimation results of the model parameters are summarized in Table 27.

3.3.5 Cost Parameters

Costs of pennycress production include variable costs and fixed costs per acre. Fixed costs come in the form of machinery ownership and operating costs in the harvesting and transportation of pennycress. Specifically, these fixed costs are the capital recovery costs of a 215hp tractor with grain cart, combine-harvester with grain head and semi-tractor trailer. One of the primary determinants of capital recovery costs is the field capacity or the number of acres per hour that the machine can effectively cover. Since these costs are only incurred after the farmer has begun pennycress production, they are distinctly different from sunk investment costs. In this analysis, variable costs of $67.47 per acre are utilized for determining optimal trigger prices and in simulating pennycress supply (Table 37).

Sunk investment costs result from expected yield reductions in soybean production. As discussed previously, yield reductions vary and are dependent on a number of factors. Therefore, a probabilistic yield reduction and expected cost is estimated by examining historic soybean yields. Data is gathered from the United States Department of Agriculture for the state of Illinois' historic soybean yields. Illinois is used in this analysis because of its large and well-established production of corn and soybeans, in addition to having favorable conditions for pennycress production. The data is annual and ranges from 1977 to 2016. Higher frequency data is unavailable as crop yields are annual in nature. The early 1970's saw a boom in farm

production and yields. Agricultural exports surged and farm incomes jumped from $2.3 billion in 1972 to $19.6 billion in 1973. Furthermore, congress had recently passed the Plant Variety Protection Act of 1971. The legislation allowed private firms to patent new hybrid seeds. Following this, corn and soybean hybrids were developed and the race to increased yield was on (Lawton 2016). During the growing season of 1974, soybean yields took an unexpected hit, despite the push to increase production. Spring rains delayed plantings, a summer drought affected growth and an early freeze damaged crops. As demand for seed grew, the 1975 soybean seed supplies were limited, which resulted in poor qualities and higher costs for producers. In response, the market saw an influx of new soybean varieties from both public sources and private. However, 1976 soybean production was still down over 18% compared to 1975. It was not until 1977 when producers had ample supply of quality seeds and inflated global prices had waned (FAO 1976, Lawton 2016). From 1940 to 1976, Illinois soybean yields averaged 26.32 bushels per acre. From 1977 to 2016, the average yield was 42.54 bushels per acre. Therefore, any yield data, prior to 1977 could be misleading and 1977 is chosen as the start date for the analysis of soybean yields.

Previous literature finds that soybean yield reductions vary from -6.5% to -28%. Schnitkey (2017), assumes a soybean yield of 61 bushels per acre for Northern Illinois, and high-productivity regions of Central Illinois. That figure is rounded down to 60 bushels per acre, resulting in probabilistic yield reductions ranging from -3.9 bushels to -16.8 bushels per acre, with a midpoint of -10.35. Using the historic distribution of Illinois soybean yields, the probability of occurrence is determined for the specified range and midpoint of yield reductions. Visually examining the distribution in a histogram does not provide clear evidence. Therefore, the theoretical distribution, which best fits the observed yield data is found by iteratively

performing a maximum likelihood estimation while specifying a number of candidate distributions. For each iteration, the distribution fit is scored by calculating the information criteria score. Each fit is scored using Akaike's information criterion (AIC), consistent AIC (CAIC), Schwarz Bayesian criterion (SBC), and Bozdogan's Information complexity criterion (ICOMP). All criteria minimize the log-likelihood function while penalizing for the number of parameters. However, ICOMP also accounts for the covariance structure and thus, for collinearity between the factors and dependence among the parameter estimates. Therefore, ICOMP is used as the determining criteria while the others are carried out simply for comparison purposes. Candidate distributions include Normal, Lognormal, Exponential, Gamma, Weibull, Extreme Value and Generalized Extreme Value. Under the ICOMP criteria, the lognormal distribution is found to be the best fitting for historic soybean yields (Table 28).

To examine the fit of the theoretical distribution over the observed data, the histogram of historical yield data is overlaid by the theoretical lognormal curve. This illustrates that the lognormal distribution appears to be a good fit as confirmed by the ICOMP criteria. The parameters are recovered from the maximum likelihood estimation while specifying a lognormal distribution. Next, the lognormal cumulative distribution function is calculated for each value of the yield reduction to calculate the probability that yield takes on a value less than or equal to 56.1, 49.7 and 43.2 (Table 29). The expected yield penalty is then calculated and multiplied by the assumed price per bushel of $10.60 to obtain the expected cost of soybean yield reduction (Schnitkey 2017).

The expected cost of soybean yield penalties is synonymous with the one-time investment payment (k) the farmer must make to enter pennycress production. With operating costs and investment costs defined, the discount rate must be determined to solve the farmer's decision

problem. The discount rate is the opportunity cost of investing in pennycress production. If the project is funded by debt capital such as a farm loan, this rate reflects the cost of capital. If both equity and debt are used, the discount rate is the weighted average cost of equity funds and debt funds. This is also known as the weighted-average cost of capital (WACC) and is formulated as

$$\delta = k_e W_e + k_d W_d (1 - \tau)$$

where k_e is the cost of equity capital, k_d is the cost of debt capital, W_e and $W_d$are the proportion of equity funds and proportion of debt funds in the business (Clark 2009). The parameter τ represents the marginal tax rate. In terms of financial measures, the cost of equity is the rate of return on equity (ROE), and the cost of debt is the interest expense divided by the farm debt. The proportion of equity funds is the equity to asset ratio, and the proportion of debt funds is the debt to asset ratio. Therefore, these parameters can be calculated using historic data of farm financials in the United States. Following Clark (2009), data is gathered from the University of Minnesota's Finnish Biodiversity Information Network (FINBIN), Farm Financial Database for all crop farms in all States, over the years 2009 to 2015. The average of all farms over this period is utilized to satisfy the equation for the discount rate or the WACC. (CFFM FINBIN 2015). The average ROE (k_e) is found to be 9.80%, the average equity to asset ratio (W_e) is 62.08%, interest expense to debt ratio (k_d) is 4.15% and the average debt to asset ratio (W_d) is 38%. Based on the average net income of all crop farms from 2009 to 2015, the marginal tax rate (τ) is 28% (The Tax Foundation 2013). Plugging these figures into the WACC equation results in a discount rate (δ) of 7.22%. The discount rate, operating costs and one-time investment costs are summarized in Table 30.

3.3.6 The Price of Pennycress Which Triggers Entry and Exit

Using the parameter values estimated in the previous sub-section, numerical approximation is carried out to identify the combination of yield and price, which induce entry and exit into pennycress production. Estimation results in threshold curves, which represent an entry threshold of $P * Y = \$110.94$ and an exit threshold of $P * Y = \$60.81$

Figure 30, shows the entry and exit threshold curves which represent all combinations of price and yield which induce the farmer to enter into pennycress production or exit pennycress production. At a yield level of 1600 pounds per acre, the price, which triggers the farmer to produce pennycress, is found to be approximately $0.08 per pound. A yield of 1200 pounds per acre results in a trigger price of approximately $0.10 per pound (Table 31). The Agricultural Marketing Resource Center has determined that $0.15 per pound would be a reasonable market price for pennycress. This analysis shows that, a price of $0.08 to $0.15 per pound of pennycress is a reasonable expectation for yields ranging from 800 pounds per acre to 1600 pounds per acre.

Yields will not be constant across the United States and individual farm decisions will vary by region. The decision to adopt pennycress will be influenced by opportunity cost in regions that have historically held comparative advantages in the production of competing crops. Adopting a corn-pennycress-soybean rotation in some areas, may mean farmers would be forced to shift acreage away from other crops such as wheat, oats or other winter crops. Furthermore, there are certain regions across the United States where pennycress is not a suitable crop due to climate conditions. To address these considerations, an analysis of historical climate data is carried out and a partial equilibrium model is utilized which accounts for competitive land allocation decisions throughout the agricultural sector.

3.4. Geographic Suitability

Pennycress is very cold tolerant and can grow in wide range of conditions. However, the plant is sensitive to high heat conditions. Pennycress has been known to survive very cold conditions with the ability to survive Canadian Prairie winters where temperatures can drop to negative 30 degrees Celsius. Sharma, Cram et al. (2007), found that Pennycress exhibits a higher freezing tolerance than *Arabidopsis thaliana* and *Brassica napus* (Canola), both of which are closely related to pennycress. With a three-week cold acclimation period, pennycress was freeze tolerant at -16.8 degrees Celsius or less than 2 degrees Fahrenheit. Temperatures above 85 degrees Fahrenheit pose a risk to pennycress and are found to have an impact on pod and seed numbers (Parker and Phippen 2012). Higher temperatures of 92 degrees Fahrenheit and above cause the plant to abort flowering (Sedbrook, Phippen et al. 2014).

Therefore, geographic regions across the United States, which poses a risk to pennycress production yield, are identified using a lower bound of 5^0F and upper bound of 88^0F. Monthly climate data, including monthly average high and low temperatures, ranging from January 2000 to December 2014 are obtained. The data consist of all climate divisions within the Continental United States (CONUS), representing 48 states and 344 climate divisions. Next, the months of June, July and August are removed from the sample. Then the frequency of reaching the lower and upper bound temperature is calculated within each climate division and each year, for the months September through May. The frequencies are then aggregated over the years 2000 through 2014. Using the aggregated frequencies, the median number of times an upper or lower bound temperature is reached, is calculated for each of the 344 climate divisions. This information is then disaggregated to the county level using a spatial union. For those counties,

which fall under more than one climate division, the county-climate division pairing with the largest geometric area is retained.

Median values range from zero to three over the sample. Therefore, it is assumed that a median of one will imply a 25% yield reduction, a median of two will imply a 50% yield reduction and a median of three will imply a 75% yield reduction. Using these assumptions, the potential yields for each county in the 48 contiguous states are estimated. For example, under the 1600 pound per acre assumption, a county with a median of one is expected to have a potential yield of 1200 pounds per acre, a 25% reduction from the 1600-pound assumption. In a county with a median of two, the potential yield is expected to be 800 pounds per acre, a 50% reduction in the 1600-pound assumption. Those counties with a median of three are expected to have a potential yield of 400 pounds per acre, a 75% reduction from the 1600-pound assumption (Figure 38).

3.5. Land Allocation and Supply Simulation

Estimated county level yields and variable costs are used to generate national and regional supply curves using a partial equilibrium model known as the Policy Analysis System (POLYSYS). The POLYSYS model simulates changes in land management, crop production, farm income, and commodity prices and estimates the resulting impacts on the U.S. Agricultural sector. The development of POLYSYS is a result of joint efforts among the University of Tennessee's Agricultural Policy Analysis Center, U.S. Department of Agriculture's Economic Research Service and Oklahoma State University's Great Plains Agricultural Policy Analysis Center (Ray, Ugarte et al. 1998). The model has theoretical underpinnings to an equilibrium displacement model (Appell, Fu et al.). EDM's are used to evaluate exogenous shocks to the equilibrium market, which result in proportional changes to variables, which are endogenously determined by the system. The equilibrium market is represented by a published baseline of

current and projected agricultural sector characteristics. The published baseline is provided by the United States Department of Agriculture (USDA). The data and assumptions used in establishing the published baselines are updated regularly and available to the public.

Each simulation of POLYSYS imposes one or more changes to the baseline scenario, acting as an exogenous shock on the equilibrium, which results in proportional adjustments to the endogenous variables, agricultural supply, demand, prices, and income. Thus, each simulation can be thought of as a series of multi-period, annual adjustments to the EDM.

While POLYSYS is conceptually an EDM, the system of equations is organized of four interdependent modules, which make up its core. The four core modules include national livestock supply and demand, regional crop supply, national crop demand and a national agricultural income module. Key inputs into the model include a baseline agricultural outlook, national land use, agricultural management practices (i.e. crop inputs and yields), livestock sector data (i.e. forage requirements), agricultural sector policy data such as government payments, and economic indicators such as demand elasticities and land use shift coefficients.

The crop supply component of POLYSYS is constructed at the conterminous United States county level. The crop supply module is thus composed of 3110 independent linear programming models (LP) which represent land allocation decisions in each county. First, estimates of the available acreage, which may be used to enter into crop production, switch production among competing crops, or move out of crop production are determined for nine model crops including corn, grain sorghum, oats, barley, wheat, soybean, cotton, hay, and rice, once the available acreage is determined, it is allocated among competing crops based on the maximization of expected returns. Enterprise budgets for each crop are used in the determination of expected returns for that crop and region. Since land allocation is based on maximizing

expected returns, POLYSYS incorporates a price expectations function. The function can take the form of a simple naïve price expectations, a weighted 3-year lagged price, or utilize a rational expectations function. In this case, the naïve price expectation is used, meaning each year's planting decision is based on previous year prices. The national prices are adjusted to county prices using a historic national-regional price relationship.

Each regional LP model is solved independently and aggregated up to obtain national crop supply estimates. Crop supply estimates are utilized by the crop demand module in the determination of crop prices. It also estimates crop usage by food, feed, industrial, exports and stock carryover. The summation of food, feed and industrial use represents domestic demand. In this module, demand is a function of price shifters, cross-price shifters and other variable non-price shifters. The simulation procedure is simplified by anchoring the analysis to the baseline and utilizing demand elasticities and price flexibilities. While the crop supplies are estimated at each county, crop demands and prices are found at the national level. Given a fixed supply curve, the demand module simultaneously solves for price and demand where price is the intersection of supply and demand.

The third core module, the livestock module, incorporates information from the crop demand and supply modules to estimate supply, demand and prices for cattle, hogs, sheep, broilers, turkeys, eggs and milk at the national level. This module estimates production at a national level to generate percent changes in livestock quantities, which are then combined with direct and cross-price flexibilities to estimate expected livestock prices.

The national agricultural income module estimates cash receipts, production expenses, government outlays, net returns, and net realized incomes. A set of identity equations and econometric models are used to estimate the crop and livestock receipts and expenses accruing to

producers. Simultaneously receipts and expenses accruing to government programs are also estimated. These are then combined to arrive at the net realized income.

Results of the national agricultural income module can be fed into another auxiliary module known as the community economic impacts module. This is particularly useful in estimating impacts to local communities by using the Impact Analysis for Planning (IMPLAN) model to measure direct, indirect, and induced economic activities. [24]

In this study, POLYSYS is used to estimate crop supply, including the supply of pennycress, at the county level for the 48 continental United States. As with all crop supply decisions in the framework, the simulation is based on a one-year enterprise budget. Pennycress is introduced to the POLYSYS framework as a double-crop enterprise of corn-pennycress-soybean, which competes for land allocation. Furthermore, soybean yields are modeled to have a 6.5% reduction in yield when preceded by pennycress. These restrictions are imposed to reflect the estimated decrease in yields resulting from later planting times. The decision to adopt pennycress production is an additional net return to the corn or soybean producer, and the total planted acreage for the pennycress/corn/soybean enterprise is evenly split among pennycress and corn/soybean

Regional supply results are then aggregated to the national level to examine the supply potential of pennycress and resulting impacts on price and acreage of corn, soybeans, grain sorghum, oats, barley, wheat, cotton and rice. Based on simulation results, the supply of renewable jet fuel can be calculated by assuming a usage rate of supplied pennycress oil. The simulation is carried out

[24]Interested readers should refer to (Ray, Ugarte et al. 1998) for a more detailed explanation of the POLYSYS modules.

for years 2016-2039, where 2016 is the baseline year. A range of possible pennycress prices are utilized to draw out pennycress supply curves of this period.

3.5.1 Results of Land Allocation and Supply Simulation

In the baseline scenario market price for pennycress is $0.00 and there is no production. Pennycress comes into production with a market price of just $0.05 per pound of pennycress seed. Over the simulation period of 2016-2039, with a market price of $0.05, an average of 6.7 million acres are planted with a National average per acre yield of 1222 lbs, resulting in an estimated 7,774 million pounds of pennycress seed supplied to the market. This translates to an estimated 2,721 million pounds of vegetable oil and 197 million gallons of jet fuel produced.[25] If the long run average price of pennycress were $0.15 per pound through years 2039, average total acres planted are estimated to be 19.60 million acres for a total supply of 23,611 million pounds of seed and an average of 597 million gallons of fuel, over the years 2016-2039. However, as the market matures and demand for pennycress oil increases, prices may increase. Therefore, POLYSYS is used to simulate price driven supply curves ranging up to $0.50 per pound of seed produced.

Table 32 illustrates the simulated supply of pennycress seed and resulting supply of renewable aviation fuel in millions of gallons for the years 2020 and 2039. Simulation results indicate a slight decline moving from years 2020 to 2039. This is indicative of a decline in projected commodity prices that is a key design feature of POLYSYS.

---

[25] With an oil content of 35%, 7,774,000,000*0.35=2,721,000,000 pounds of oil. There are estimated to be 49.4 pounds of jet fuel in every 100 pounds of vegetable oil (Pearlson, Wollersheim et al. 2013). Therefore, 2,686,000,000*0.494=1,344,000,000 pounds of jet fuel. A pound of renewable jet fuel weighs 6.84 pounds, so 1.344 million pounds is converted to gallons by dividing by 6.84, which results in 197 million gallons of renewable jet fuel.

Pennycress has the potential to supply a large contribution of renewable aviation fuel to the market, particularly above the $0.15 threshold. However, increases in price have a diminishing effect on supply. The representative supply curves under 1600 pound scenario are presented in Figure 34, for the years 2020 and 2039.

Supply of renewable fuel sharply turns upward for pennycress prices above $0.20, indicating a production capacity in the corn and soybean rotation, for the regions identified. As pennycress production becomes more profitable, acreage is shifted away from other crops while corn and soybean acreage increases (Figure 35 & Table 38).

Table 38 presents the simulated results for eight primary crops including corn, grain sorghum, oats, barley, wheat, soybeans, cotton and rice. These results belong to the scenario with an assumed yield potential of 1600 pounds per acre and a market price of $0.15 per pound of pennycress seed. The baseline figures represent the case where there is no market for pennycress. With an existing market price of $0.15, an annual average of 18.6 million acres of pennycress are harvested. With the addition of pennycress acreage, comes a slight increase in harvest corn and soybean acreage. In this scenario, harvested acreage of corn and soybean includes single-crop and double-cropping methods. Therefore, total harvested acreage increases by 3.2% and 5% over the baseline scenario for corn and soybean respectively. This leads to an increase in production and total supply of corn, increasing from 14.89 billion bushels to 15.25 billion bushels. Soybean production increases from 40.4 billion bushels to 41.3 billion bushels.

Corn prices subsequently fall by approximately -8% to an average price of $3.40 per bushel, down from $3.70 per bushel. The price of soybean remains relatively flat with an average price of $8.90 per bushel in the base case and $8.80 per bushel in the alternate case.

Despite the estimated decrease in corn prices, corn producers experience a 25.3% increase in net returns. This result is driven by decreasing variable expenses in corn production of -8.24% and a slight increase in the value of production of 2.63%. Soybean producers experience a 10.3% increase in net returns, driven primarily by flat prices and increased production.

Significant changes also occur in other commodities. Grain sorghum, wheat and cotton are estimated to decline in total supply by -5%, -6.2% and -12% respectively. The decreased total supplies coincide with a subsequent increase in average prices of 9.9%, 11.0% and 21.4% respectively. The net return effects on sorghum, wheat and cotton are mixed. Sorghum net returns decrease by -66.6%, and cotton net returns decrease by -9.4%. However, the net returns for wheat increase by 30.5%. Barley and oat producers also experience a decline in net returns of -33.4% and -20.1% respectively.

The loss in net returns to barley, oats, sorghum and cotton are offset by increased net returns in corn, wheat, soybean, rice and the addition of pennycress. The production of pennycress is estimated to provide an additional $3.541 billion in new farm revenues and net returns of approximately $2.219 billion. Therefore, total net returns are estimated to increase by 23.3% over the baseline.

Increases in the price of pennycress will continue the trend of increased acreage in corn and soybean while sorghum, oat, barley, wheat, cotton and rice decline. At $0.50 per pound, pennycress production adds $15.243 billion to farm revenues and $13.502 billion in additional net returns. Therefore, total net returns across all crops in Table 38 increase from $38.795 billion to $60.797.1 billion, an increase of 56.7% (Table 39).

3.5.2 Regions of Pennycress Production

Figure 40 illustrates those counties where pennycress comes into production by year 2039. Of the 3109 counties in the contiguous states, 72.2% are expected to plant pennycress acreage for a total of 24.432 billion pounds of pennycress seed by year 2039. The minimum production is found to be in Keweenaw County, Michigan; however, the production is economically insignificant with only roughly .1026 acres planted and approximately 54.8 pounds of pennycress produced. The maximum production level is found in McLean County, Illinois, with 123,253 acres planted and 87.82 million pounds of pennycress produced. This production level is quite large in comparison to the median production of pennycress is 1.22 million pounds per county. McLean County Illinois has 653,874 cropland acres and the average farm size is 465 acres. This translates to an average, of approximately 265 farms in McLean County producing pennycress.

McLean County is located in an ideal pennycress-growing region according to historical climate data and pennycress characteristics. As seen in Figure 37, McLean County is located in an area, which is modeled as having no risk to pennycress yields, and potential yields are assumed 1600 pounds per acre.

Regions where pennycress is most economically feasible are densely located throughout the Midwest Corn Belt. Corn and soybean are the predominant crop rotations in this region, and incorporating pennycress into the rotation of corn-pennycress-soybean requires little change in comparison to areas that traditionally grow other crops. The median pennycress production across all states is 469.3 million pounds. However, for the Corn Belt States of Iowa, Illinois, Nebraska, Minnesota, Michigan, Ohio, Kansas and Indiana, the median pennycress production per county is 1,935.2 million pounds. In fact, these eight Corn Belt States account for 13.996

billion pounds of pennycress collectively, or 57.3% of the estimated total 24.432 billion pounds of pennycress production in 2039.

Conversion facilities would likely benefit from being located throughout the Corn Belt, particularly in the early years of pennycress market establishment. Technology for the conversion of pennycress to advanced biofuel is under development and there is an increasing demand by the aviation industry for renewable jet fuels. Based on techno-economic literature, plant-gate costs for the production of green-jet fuel can be estimated.

3.6. Plant-Gate Costs

Production of renewable fuel can occur through a number of technological processes. One such technological process, which shows considerable promise, is a Green Fuel Technology capable of producing three different types of alternative fuel through the conversion of eight or more different feedstocks. This technology can produce green-gasoline, green-diesel or a drop-in green jet fuel. Furthermore, co-products such as propane and naphtha are a result of the process.

Green-jet fuel is a second-generation alternative fuel that can be blended at 50/50 ratio with conventional jet fuels. This fuel can reduce greenhouse gas emissions by 65-85% compared with conventional jet fuels and has shown higher energy density in flight, allowing aircraft to fly further on less fuel. It is also a drop-in replacement fuel which requires no changes to aircraft technology or fuel infrastructure (UOP 2016). Plant-gate costs are therefore based on the production of a green-jet fuel using purchased pennycress oil. Typically, a techno-economic analysis is required to estimate cost parameters of a production pathway. Techno-economic analysis (TEA) is a method used to assess the technical and economic performance of a particular production pathway (Brown and Brown 2013). TEA's represent a simplified version of commercial scale projects, that allows the biorefinery's production pathway to be evaluated

for feasibility. The primary method of developing a TEA is through a detailed process model. Performance information is collected on the technologies under consideration. Appropriate production scenarios are identified and the process model is designed using process-engineering software. Based on the process design, capital investments are determined and a discounted cash flow analysis is performed. This allows the investment and production cost of a biorefinery to be determined.

Due to the detailed and specialized information required to conduct a TEA, cost parameters are sourced from previous literature. Two primary cost parameters utilized in this study are operating costs and capital costs (initial investment cost). English, Menard et al. (2016), provide a thorough review of operating and capital costs for a variety of production pathways. For green fuel production pathways, the authors examine four different conversion processes: 1) hydro-processing of purchased oil, 2) pyrolysis to hydro-processing, 3) gasification and 4) pyrolysis of biomass. Since pennycress is the chosen feedstock under consideration for this study, cost parameters represent the hydro processing of purchased pennycress seed.

A fully integrated facility which performs the crushing, extraction and hydroprocessing of pennycress oil is modeled. Table 33 summarizes the plant-gate costs of the representative facility which produces green-jet fuel. The facility produces 4000 barrels per day which translates to 61,320,000 gallons per day of product. This requires the crushing and extraction of 742.6 mil pounds of seed. Therefore, the facility is modeled as having two crushing and extraction units which each have a maximum capacity to process 900 tons per day.

Assuming the plant maximizes jet production, 33 million gallons of jet fuel are produced along with co-products of 4.46 mil gallons of naphtha, 15.7 mil gallons of diesel and 4.04 mil gallons of propane. Capital costs per gallon of jet fuel are $0.86 and operating costs per gallon of jet fuel

are $3.50, including a per gallon feedstock cost of $2.71. Total capital and operating costs per gallon is $4.36.[26]

Therefore, absent of any policy incentives, a single conversion facility which produces 33 million gallons of jet fuel would require $4.36 per gallon to cover operating expenses. Utilizing the operating cost and capital cost required to produce an advanced biofuel at this representative conversion facility, a cash flow analysis can be conducted to determine breakeven prices and the economic feasibility of producing a second-generation biofuel from pennycress feedstock. However, much like the pennycress producer, the biofuel producer faces uncertainty and irreversibility in their decision to invest in such a project. Therefore, the price, which induces the firm to invest in the production of renewable aviation fuels, or other second-generation biofuels, is likely to be significantly higher than a traditional breakeven price. Applying the model of chapter two indicates this price to be approximately $3.86, for a firm not reliant on the RIN subsidy. However, firms reliant on the RIN subsidy might require an even higher price to trigger investment. As was illustrated in chapter two, the high volatility in RIN price may have counteracting effects and actually delay investment with potential entrants requiring approximately $4.44 per gallon. Producer tax incentives such as the producer tax-credit and bonus depreciation work to offset the headwinds of the RIN market. When accounting for these incentives, the price that triggers investment is estimated to be $3.63 per gallon.[27]

---

[26] Capital and operating costs in the production of green-diesel, naphtha and propane are presented in Table 34, Table 35 and Table 36.
[27] See section 2.5.3 for a detailed explanation

3.7. Conclusion

This study uses enterprise budgeting, real-option modeling and partial equilibrium simulations to examine the potential for pennycress to supply a renewable aviation industry. Pennycress is modeled as a winter cover crop, which also provides a cash benefit while fitting into the traditional corn and soybean rotation. Enterprise budgets are developed to estimate net returns of pennycress, based on a yield of 1600 pounds per acre. Suitable regions of growth are identified based on agronomic characteristics. Regions that pose climatic risk to pennycress are penalized in the form of yield reductions, based on the median number of times an extreme temperature is reached during the pennycress growth season. Temperatures of 5 degrees Fahrenheit or less and 88 degrees Fahrenheit or more are deemed extreme for pennycress. Yield reductions reduce net return yet still allow these areas to come into production, particularly where corn and soybean have comparative advantages.

A real option analysis is developed to model the decision of adding pennycress to the corn and soybean rotation. In this framework, both yield and price are modeled as a stochastic mean reverting process. Historic yield of Canadian rapeseed and rapeseed spot price (Rotterdam FOB) is used parameterize the model. This analysis identifies the optimal price, which triggers investment into pennycress production while accounting for the yield and price uncertainty, as well as irreversible sunk costs. Yields ranging from 800 to 1600 pounds per acre coincide with an entry trigger price of $0.15 to $0.08 per pound. Previous research places the potential market price within this range at $0.15 per pound.

Using a partial equilibrium model known as POLYSYS, supply curves are simulated using a range of prices from $0.00 to $0.50 per pound. Supplies are estimated at the county level and aggregated to the national level to estimate the total amount of pennycress and renewable

aviation fuel supplied. Price is allowed to increase at five-cent increments. Pennycress first comes into production at $0.05 with 6.65 million acres planted. As the simulated price increases, the market quickly responds by doubling planted acres at $0.10 per pound. The simulation is run for over a 25-year period with the terminal year occurring in 2039.

Assuming a market price of $0.15 per pound of pennycress seed, planted-acreage is estimated to reach 20.17 million acres, producing 24,245 million pounds of pennycress seed and 8,486 million pounds of pennycress oil by 2020. This amount of pennycress oil would feed 88 conversion facilities with a nameplate capacity of 13 million gallons of drop-in green fuel, which each demand 96.3 million pounds of pennycress oil. With 24,245 million pounds of pennycress seed, approximately 613 million gallons of renewable aviation fuel is produced.

By the year 2039, 20.30 million acres would be planted while assuming a market price of $0.15 per pound of pennycress seed. This would generate approximately 24,432 million pounds of pennycress or 8,551 pounds of pennycress oil, enough to supply roughly 89 conversion facilities with a nameplate capacity of 13 million gallons. These results are flat compared to the year 2020, which is due to the projections of commodity price embedded in the simulation model POLYSYS. These projections are regularly updated by United States Department of Agriculture (USDA), Food and Agricultural Policy Research Institute (FAPRI) and the U.S. Congressional Budget Office (CBO).

Based on the estimates it appears that pennycress has the potential to supply approximately 600 to 800 million gallons to an alternative aviation fuel industry. The results are believed to be conservative based on current projections of corn and soybean price, conservative yield assumptions and an imposed 6.5% soybean reduction. However, previous research has shown that subsequent years of pennycress production can actually improve soybean yields as the soil

benefits of a cover crop accrue. The model does not account for soil or other environmental benefits through the adoption of a cover crop in corn and soybean rotations. This limitation is acknowledged and left for future research. Another limitation is the use of current projections of commodity price. As projections are updated, the effect of rising or falling prices in corn, grain sorghum, oats, barley, wheat, soybean, cotton and rice will alter the results. If future prices of corn and soybean rise, the supply of pennycress would rise as producers look to meet market demands of corn and soybean. Furthermore, the maximum yield potential is assumed 1600 pounds per acre for pennycress. Over the regions, which come into production, average yields are approximately 1,267 pounds per acre. However, previous agronomic research has found yield to be as high as 2200 pounds per acre. Genetic research into this crop is also ongoing. Pennycress shares a similar genetic structure with Arabidopsis. The large body of research around Arabidopsis may allow for rapid advances in pennycress genetic varieties with improved yield and other desirable characteristics, making pennycress an even more viable option than what is found here.

3. A. Appendix

3. A.1 Tables

Table 27: Stochastic yield and price parameter estimations

Parameter	Definition	Value	Scale
$\hat{\lambda}_p$	Reversion Rate Rapeseed Price	0.1603	Per year
$\hat{\lambda}_y$	Reversion Rate Rapeseed Yield	0.0182	Per year
$\hat{\sigma}_p$	Standard Deviation Rapeseed Price	0.2064	Per year
$\hat{\sigma}_y$	Standard Deviation Rapeseed Yield	0.1123	Per year
ρ_{pr}	Correlation of Yield and Price	-0.2943	Per year

Table 28: Best of fit distribution for historical soybean yield

	Historical Soybean Yield Best of Fit (1977-2016)						
	EV	EXPO	GAMMA	GEV	LOGNORM	NORM	WEIBUL
ICOMP	280.69	380.03	269.27	269.59	268.74*	268.92	273.43

** Minimum score*

Table 29: Probability of occurrence for yield reductions and expected cost

Percent	Reduction	Probability of Occurring	Probabilistic Yield Penalty
6.50%	3.9	0.9601	3.74
17.25%	10.35	0.8464	8.76
28.00%	16.8	0.5693	9.57
Expected Yield Penalty	22.07	Expected cost of yield penalty	$233.95 per acre

Table 30: Summary of pennycress cost of production parameters

Parameter	Definition	Value	Scale
δ	Discount Rate	0.0722	per year
w	Operating Cost	$67.47	per acre
k	One-time Sunk Cost	$233.95	per acre

Table 31: Entry and exit trigger prices for pennycress production

Yield and Price Combinations						
Yield	800	1000	1200	1400	1600	1800
Entry Trigger Price	$0.15	$0.12	$0.10	$0.09	$0.08	$0.07
Exit Trigger Price	$0.05	$0.04	$0.04	$0.04	$0.03	$0.03

Table 32: Simulated pennycress production (figures in millions)

1600 pound per acre yield scenario						
	2020			2039		
Price	Supply(lbs)	Planted Acres	Fuel (gal.)	Supply(lbs)	Planted Acres	Fuel (gal.)
$0.00	0	0	0	0	0	0
$0.05	7674	6.65	194.0	8146	6.94	205.9
$0.10	17728	14.67	448.1	17736	14.54	448.3
$0.15	24245	20.17	612.9	24432	20.30	617.6
$0.20	27352	22.98	691.4	26905	22.49	680.1
$0.25	28936	24.36	731.4	28367	23.82	717.1
$0.30	29897	25.18	755.7	29215	24.53	738.5
$0.40	31103	26.26	786.2	30737	25.89	777.0
$0.50	32010	27.15	809.1	31559	26.66	797.7

Table 33: Summary of plant-gate costs in the conversion of pennycress oil to green-jet

Pennycress fed Green Jet Fuel Production Pathway	
	Costs - fully integrated
Facility demand (pounds of seed)	595,645,419
Price ($/pound)	$0.15
Yield (pounds/acre)	1,600
Production (Gallons)	32,999,746
Name Plate Capacity (gallons)	61,320,000
Capital Feedstock to bio-oil ($)	$21,328,404
Capital Transportation ($)	$118,915
Capital Cost of Hydroprocessing ($)	$9,805,686
Total Capital Costs ($)	$31,253,005
Total Capital Costs ($/gallon)	$0.95
Operating cost Conversion ($)	$6,614,435
Cost of Feedstock ($/gallon)	$2.71
Operating Cost of Transportation ($)	$666,396
Operating Cost of Hydro-Processing	$36,011,556
Brownfield or other operating Credit ($)	($1,293,184)
Total Operating Cost	$41,999,203
Total Operating Costs ($/gallon)	$3.98
Total Capital and Operating Costs ($/gallon)	$4.93

Table 34: Summary of plant-gate costs in the conversion of pennycress oil to green-diesel

Pennycress fed Green Diesel Production Pathway	
	Costs - full integrated
Facility demand (pounds of seed)	124,118,177
Price ($/pound)	$0.15
Yield (pounds/acre)	1,600
Production (Gallons)	15,703,879
Name Plate Capacity (gallons)	61,320,000
Capital Feedstock to bio-oil ($)	$42,656,808
Capital Transportation ($)	$118,915
Capital Cost of Hydroprocessing ($)	$9,805,686
Total Capital Costs ($)	$52,581,409
Total Capital Costs ($/gallon)	$0.86
Operating cost Conversion ($)	$13,228,870
Cost of Feedstock ($/gallon)	$1.19
Operating Cost of Transportation ($)	$666,396
Operating Cost of Hydro-Processing	$36,011,556
Brownfield or other operating Credit ($)	($1,293,184)
Total Operating Cost	$48,613,638
Total Operating Costs ($/gallon)	$1.98
Total Capital and Operating Costs ($/gallon)	$2.84

Table 35: Summary of plant-gate costs in the conversion of pennycress oil to green-propane

Pennycress fed Green Propane Production Pathway	
	Costs - full integrated
Facility demand (pounds of seed)	8,736,970
Price ($/pound)	$0.15
Yield (pounds/acre)	1,600
Production (Gallons)	4,043,969
Name Plate Capacity (gallons)	61,320,000
Capital Feedstock to bio-oil ($)	$42,656,808
Capital Transportation ($)	$118,915
Capital Cost of Hydroprocessing ($)	$9,805,686
Total Capital Costs ($)	$52,581,409
Total Capital Costs ($/gallon)	$0.86
Operating cost Conversion ($)	$13,228,870
Cost of Feedstock ($/gallon)	$0.32
Operating Cost of Transportation ($)	$666,396
Operating Cost of Hydro-Processing	$36,011,556
Brownfield or other operating Credit ($)	($1,293,184)
Total Operating Cost	$48,613,638
Total Operating Costs ($/gallon)	$1.12
Total Capital and Operating Costs ($/gallon)	$1.97

Table 36: Summary of plant-gate costs in the conversion of pennycress oil to green-naphtha

Pennycress fed Green Naphtha Production Pathway	
	Costs - full integrated
Facility demand (pounds of seed)	14,100,975
Price ($/pound)	$0.15
Yield (pounds/acre)	1,600
Production (Gallons)	4,463,966
Name Plate Capacity (gallons)	61,320,000
Capital Feedstock to bio-oil ($)	$42,656,808
Capital Transportation ($)	$118,915
Capital Cost of Hydroprocessing ($)	$9,805,686
Total Capital Costs ($)	$52,581,409
Total Capital Costs ($/gallon)	$0.86
Operating cost Conversion ($)	$13,228,870
Cost of Feedstock ($/gallon)	$0.47
Operating Cost of Transportation ($)	$666,396
Operating Cost of Hydro-Processing	$36,011,556
Brownfield or other operating Credit ($)	($1,293,184)
Total Operating Cost	$48,613,638
Total Operating Costs ($/gallon)	$1.27
Total Capital and Operating Costs ($/gallon)	$2.12

Table 37: Pennycress enterprise budget assuming 1600lbs per acre

2017 Field Pennycress (Thlaspi Arvense)					
		Unit	Quantity	Price	Total
Revenue	Pennycress [1]	lbs	1600	$0.15	$240.00
				Total Revenue	**$240.00**
Variable Expenses	Seed [2]	lbs	5	$2.50	$12.50
	Fertilizer & Lime (Table 1.)	Acre	1	$26.50	$26.50
	Chemical (Table 2.) [3]	Acre	1	$0.00	$0.00
	Repair & Maintenance (Table 3.) [4]	Acre	1	$6.52	$6.52
	Fuel, Oil & Filter (Table 3.) [4]	Acre	1	$6.93	$6.93
	Operator Labor (Table 3.) [4]	Acre	1	$3.05	$3.05
	Machinery Rental [8]	Acre	1	$10.00	$10.00
	Crop Insurance [5]	Acre	1	$0.00	$0.00
	Operating Interest [6]	%	$65.50	6.00%	$1.97
	Other Variable Costs	Acre	1	$0.00	$0.00
				Total Variable Expenses	**$67.47**
Fixed Expenses	Machinery [4]				
	Capital Recovery (Table 3.)	Acre	1	$18.24	$18.24
	Other Fixed Machinery Costs	Acre	1	$0.00	$0.00
	Property Taxes	Acre	1	$0.00	$0.00
	Insurance (Non-Machinery)	Acre	1	$0.00	$0.00
	Other Fixed Costs [7]	Acre	1	$0.00	$0.00
				Total Fixed Expenses	**$18.24**
				Return Above All Specified Expenses	**$154.29**

Table 38: Changes from base case to alternate scenario with pennycress price of $0.15 per pound

	Harvested Acres (Mil Ac)			Total Supply (Mils)			PRICE ($/unit)			Net Returns		
	Base case	Alternate	% Change	Base case	Alternate	% Change	Base case	Alternate	% Change	Base case	Alternate	% Change
Corn (bu)	79.58	82.12	3.2%	14884	15252	2.5%	3.7	3.4	-8.1%	16145	20230	25.3%
Grain Sorghum (bu)	5.11	4.86	-5.0%	337.0	320.0	-5.0%	3.6	4.0	9.9%	798.0	266.8	-67%
Oats (bu)	0.90	0.86	-5.1%	62.3	59.3	-4.7%	2.2	2.1	-3.5%	-185.0	-222.2	20.1%
Barley (bu)	2.60	2.49	-4.2%	198.0	191.2	-3.4%	5.2	4.9	-5.2%	362.0	240.9	-33%
Wheat (bu)	44.11	40.19	-8.9%	2178.8	2044.1	-6.2%	5.0	5.5	11.0%	2742.0	3577.2	30.5%
Soybeans (bu)	80.06	84.07	5.0%	4044.6	4129.2	2.1%	8.9	8.8	-0.7%	19106.0	21078.9	10.3%
Cotton (bale)	8.40	7.30	-13%	14.8	13.0	-12.0%	0.6	0.7	21.4%	-1106.0	-1210.0	-9.4%
Rice (cwt)	2.89	2.74	-5.0%	234.0	223.3	-4.6%	13.8	16.8	21.9%	933.0	1663.2	78.3%
Hay (Tons)	56.77	56.73	-0.1%	136.0	135.5	-0.4%	144.7	154.5	6.8%	-	-	-
Pennycress (lbs)	0.00	18.62	-	0.0	23610.8	-	0	0.15	-	0.0	2219.0	-
Total All Crops	280.4	281.36	0.3%							38795.0	47843.8	23.3%

Note: Some figures rounded to reduce space requirements

Table 39: Changes from base case to alternate scenario with pennycress price of $0.50 per pound

	Harvested Acres (Mil Ac)			Total Supply (Mils)			PRICE ($/unit)			Net Returns		
	Base case	Alternate	% Change	Base case	Alternate	% Change	Base case	Alternate	% Change	Base case	Alternate	% Change
Corn (bu)	79.58	84.45	6.1%	14884	15497	4.1%	3.7	3.2	-12.8%	16145	19936	23.5%
Grain Sorghum (bu)	5.11	4.63	-9.4%	337.0	301.3	-11%	3.6	4.5	25.6%	798.0	405.6	-49%
Oats (bu)	0.90	0.83	-8.6%	62.3	57.7	-7.3%	2.2	2.1	-4.8%	-185.0	-209.4	13.2%
Barley (bu)	2.60	2.37	-8.8%	198.0	182.5	-7.8%	5.2	5.4	3.8%	362.0	309.5	-15%
Wheat (bu)	44.11	36.90	-16%	2178.8	1871.9	-14%	5.0	6.4	27.5%	2742.0	4732.5	72.6%
Soybeans (bu)	80.06	86.86	8.5%	4044.6	4152.5	2.7%	8.9	8.8	-1.6%	19106.0	21285.8	11.4%
Cotton (bale)	8.40	6.91	-18%	14.8	12.5	-16%	0.6	0.7	23.9%	-1106.0	-1029.3	6.9%
Rice (cwt)	2.89	2.51	-13%	234.0	205.5	-12%	13.8	18.5	34.3%	933.0	1864.5	99.8%
Hay (Tons)	56.77	56.52	-0.4%	136.0	135.3	-0.5%	144.7	154.5	6.8%	0.0	0.0	0.0%
Pennycress (lbs)	0.00	24.51	-	0.0	30486.8	-	0.0	0.5	-	0.0	13502.5	0.0%
Total All Crops	280.4	281.98	0.6%							38795.0	60797.1	56.7%

Note: Some figures rounded to reduce space requirements

3. A.2 Figures

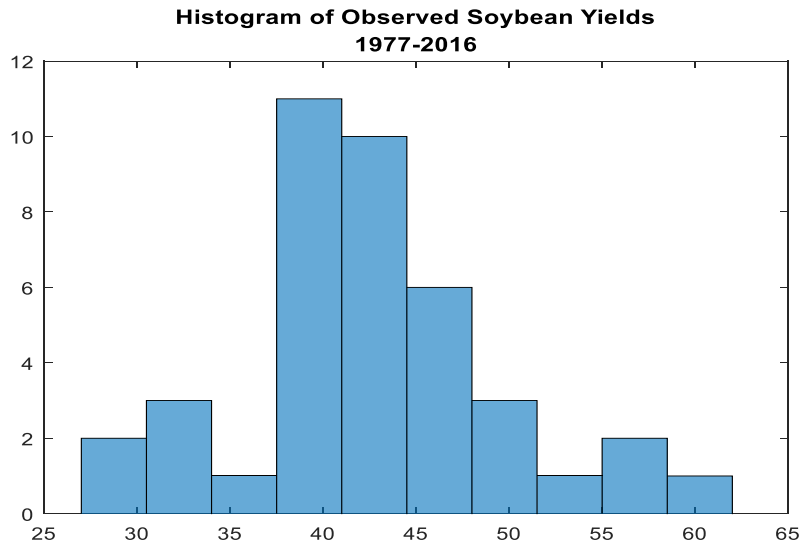

Figure 31: Distribution of Illinois soybean yields

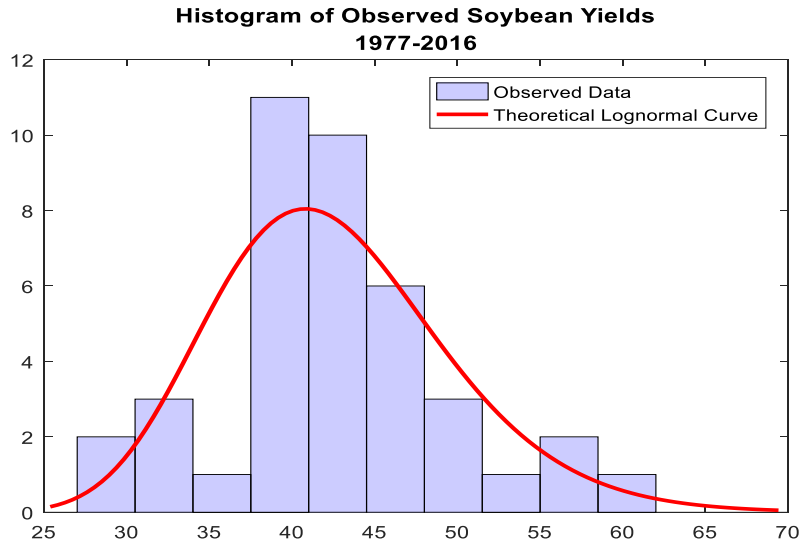

Figure 32: Observed data overlaid by best of fit theoretical distribution

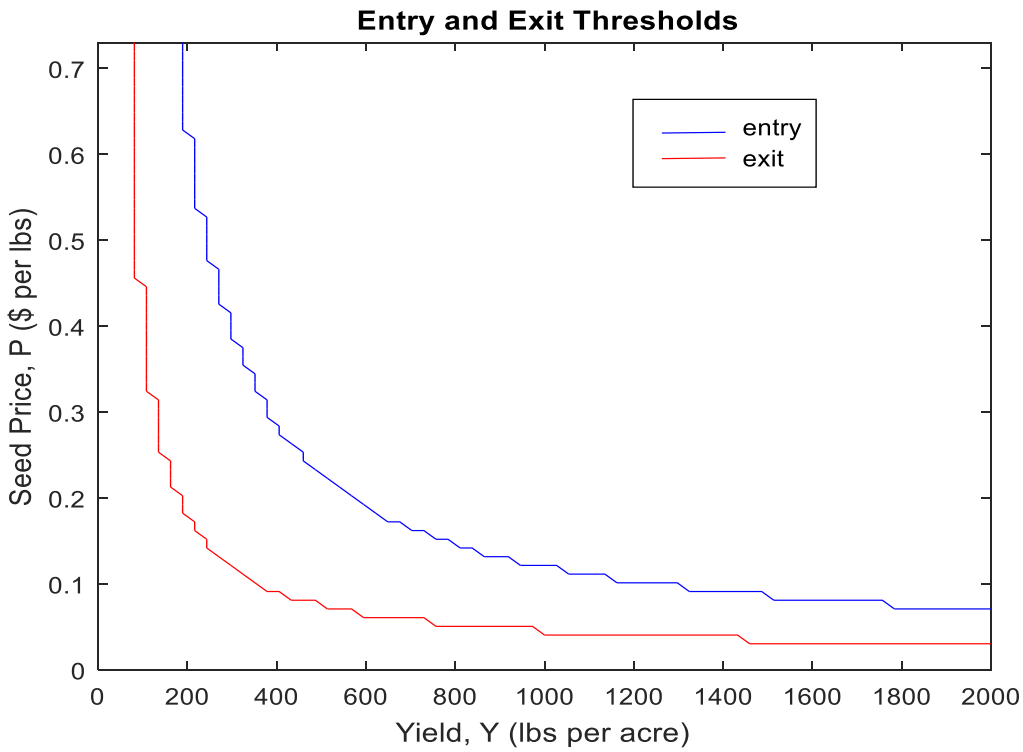

Figure 33: Entry and Exit Thresholds for all combinations of price and yield

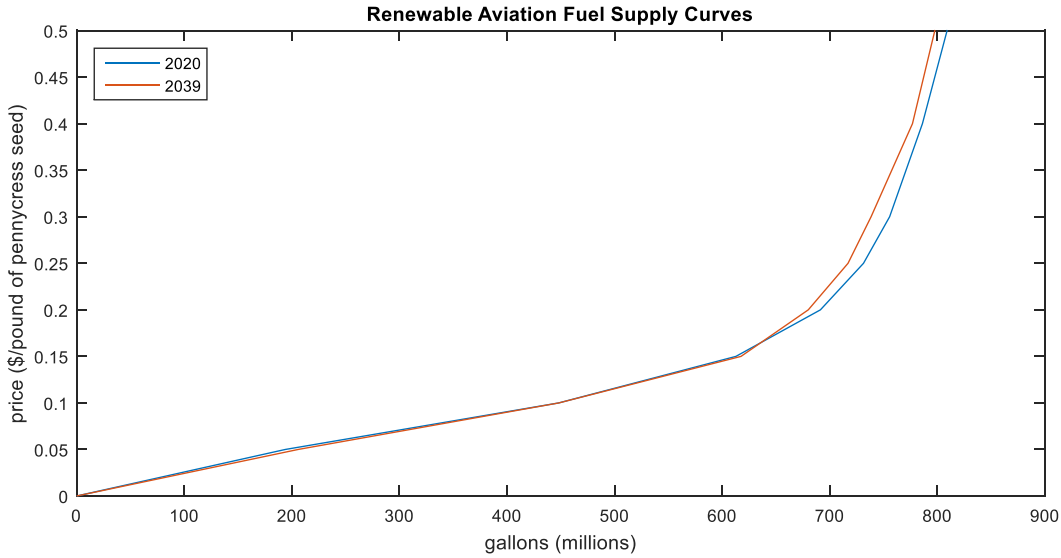

Figure 34: Supply curves under the assumption of 1600 pounds per acre yield

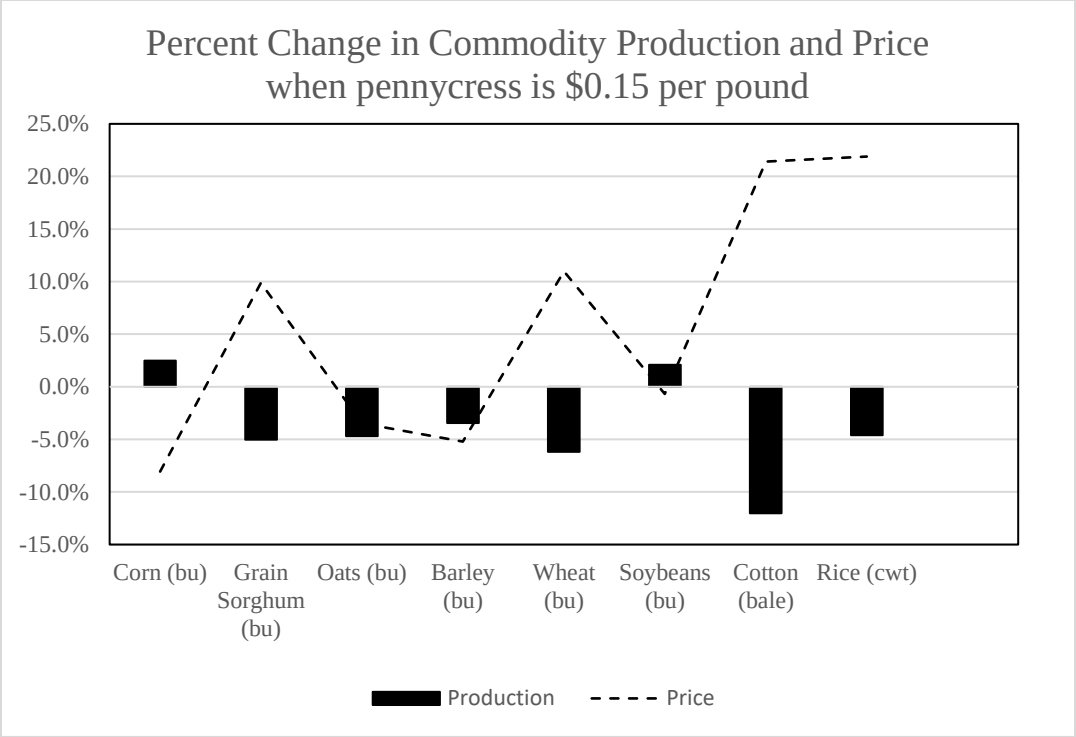

Figure 35: Pennycress disruption effects on commodity production and price

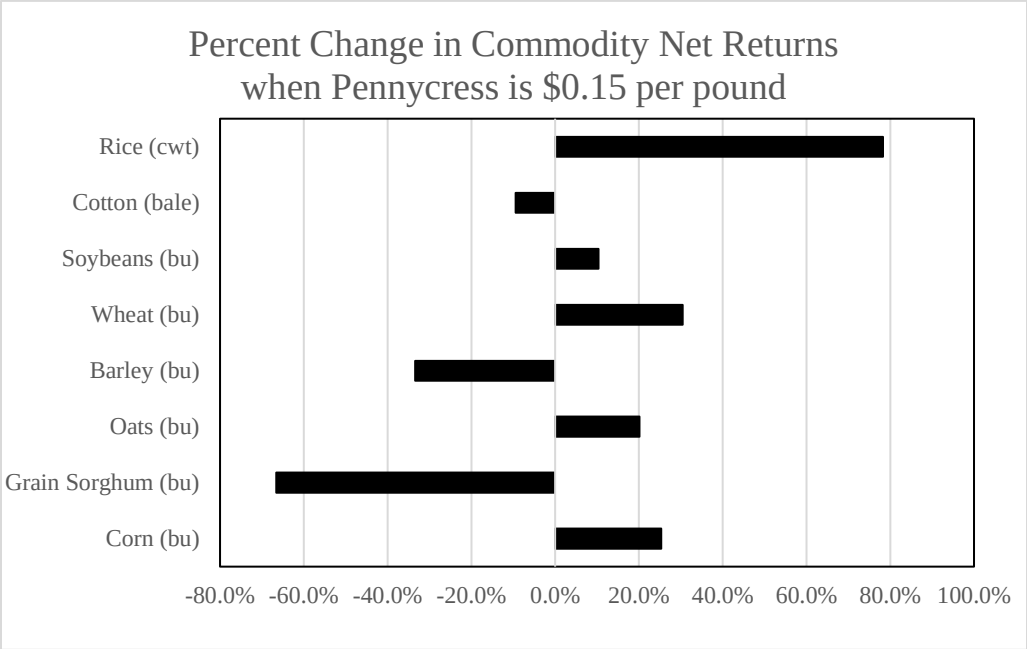

Figure 36: Pennycress disruption effects on commodity net returns

Figure 37: Simulated top pennycress producing county

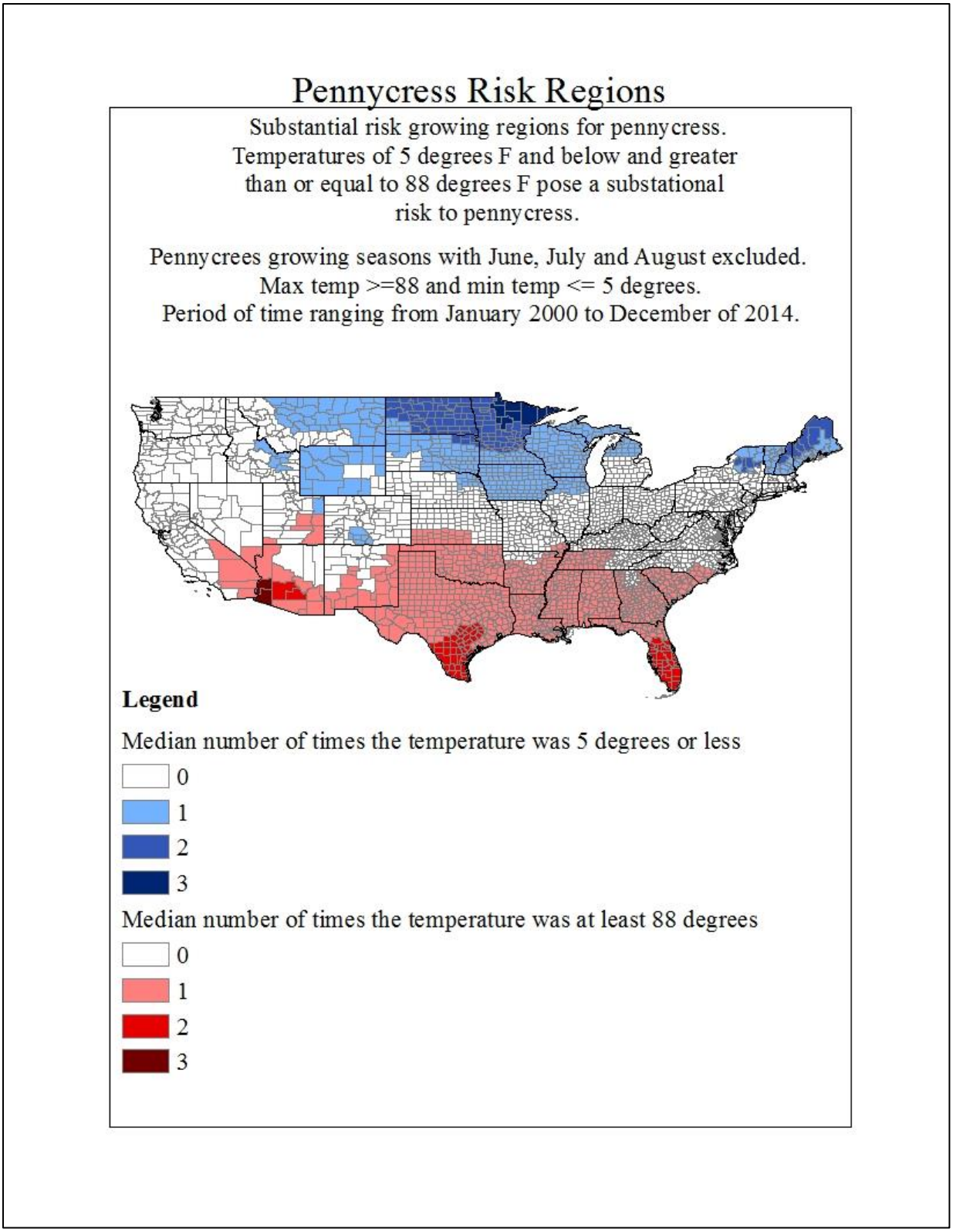

Figure 38: Regions that pose a risk to pennycress based on historical climate data

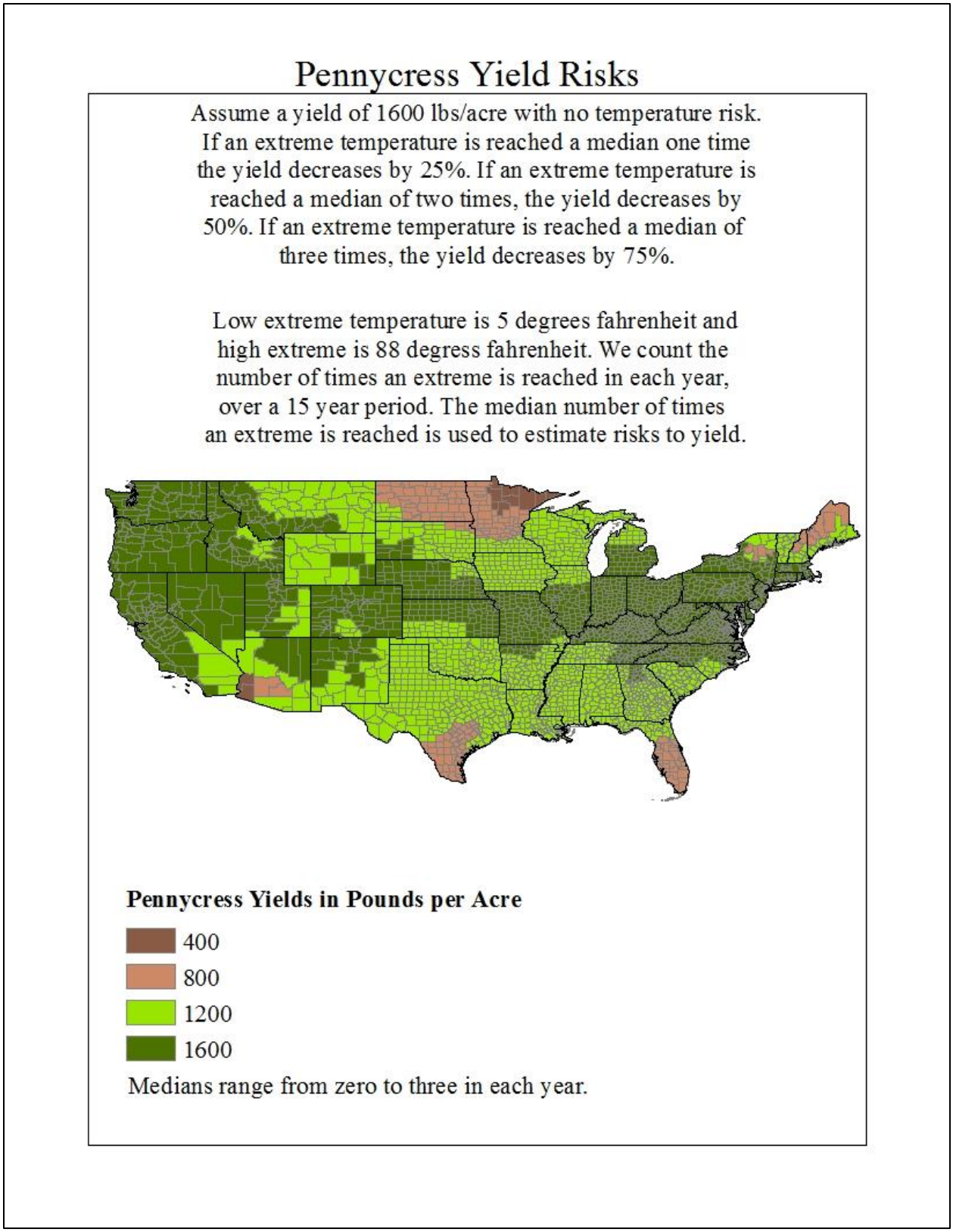

Figure 39: Regions and their expected yields based on climate risk regions

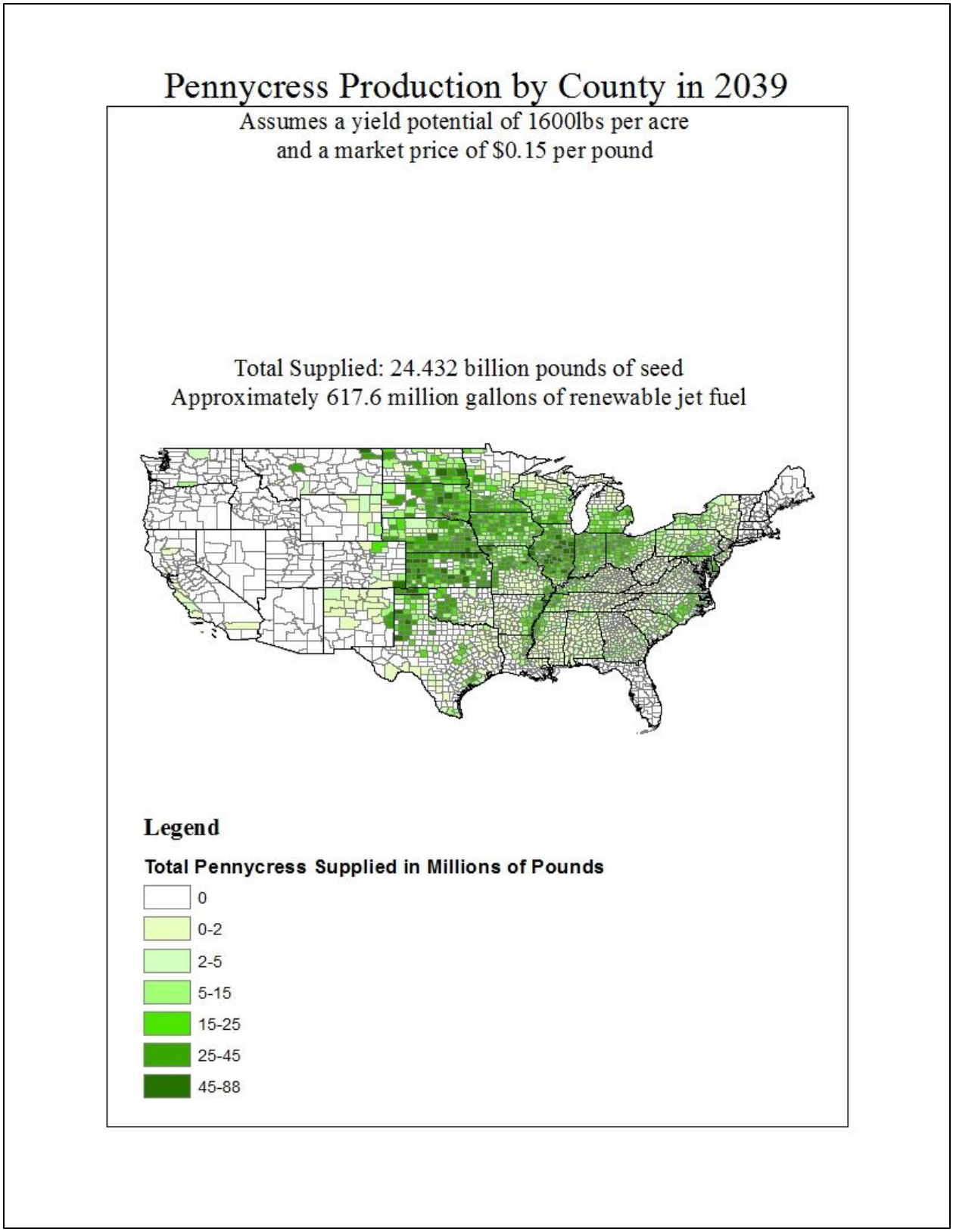

Figure 40: Simulated pennycress production in year 2039

Conclusion

This research is aligned with identifying barriers throughout the alternative jet-fuel supply chain. Barriers include complex markets for tradable credits, policy uncertainty and insufficient supplies of feedstock and fuel.

Production of second-generation biofuel such as alternative jet-fuel is supported by public policy efforts, such as the program for a tradable credit known as the renewable identification number (RIN). The RIN market is complex and uncertain. This research aids in disembroiling these complexities by examining past RIN price behavior and drivers of RIN price regime change. The renewable fuel standard is often uncertain year to year, forcing market participants, obligated parties and other agents to adapt their behaviors to the changing regulations. This leads to volatile RIN price which is found to delay optimal investment decisions for the producer for alternative jet-fuels and other second-generation biofuels. Producer tax incentives work to counteract the effects of policy uncertainty, but these incentives would be more effective if coupled with reductions in policy uncertainty.

To identify opportunities of supply, an experimental crop known as pennycress is evaluated for its potential to support a renewable aviation industry. To do so, an enterprise budget is developed and partial equilibrium simulations are carried out. Furthermore, this analysis identifies the optimal price, which triggers investment into pennycress production while accounting for the yield and price uncertainty, as well as irreversible sunk costs. Yields ranging from 800 to 1600 pounds per acre coincide with an entry trigger price of $0.15 to $0.08 per pound. Based on the estimates it appears that pennycress has the potential to supply approximately 600 to 800 million gallons to an alternative aviation fuel industry.

References

Agricultural Marketing Resource Center (AGMRC). (2015). "Pennycress." Retrieved Februrary 20, 2016, from http://www.agmrc.org/commodities-products/grains-oilseeds/pennycress/.

Alich, J. A. and R. Inman (1976). "Energy from agriculture: The most economic method of large scale solar energy conversion." Energy **1**(1): 53-61.

Appell, Miller and Wender (1970). Conversion of urban refuse to oil. Washington, D.C., U.S. Dept. of the Interior, Bureau of Mines.

Appell, H., I. Wender and R. Miller (1969). Solubilisation of low rank coal with carbon monocide and water, SOC Chemical Industry 14 Belgrave Square, London, England.

Appell, H. R., Y. C. Fu, S. Friedman, P. M. Yavorsky, I. Wender and C. Pittsburgh Energy Research (1971). Converting organic wastes to oil: a replenishable energy source. Washington, D.C., U.S. Bureau of Mines.

Arvens Technology Inc. (2010). "Pennycress: America's Next Advanced Biofuel Source." Retrieved Februrary 20, 2016, from http://www.growpennycress.com/farming.html.

Babcock, B. A. (2015). "Breaking the link between food and biofuels." Iowa Ag Review **14**(3): 1.

Bai, J. and P. Perron (1998). "Estimating and testing linear models with multiple structural changes." Econometrica: 47-78.

Bassham, J. A. (1976). Covered energy farms for solar energy conversion, California Univ., Berkeley (USA). Lawrence Berkeley Lab.

Bracmort, K. (2015). "The Renewable Fuel Standard (RFS): Cellulosic Biofuels." Congressional Research Service **31**.

Brekke, K. A. and B. Øksendal (1994). "Optimal switching in an economic activity under uncertainty." SIAM Journal on Control and Optimization **32**(4): 1021-1036.

Brooks, C. (1997). "Linear and non-linear (non-) forecastability of high-frequency exchange rates." Journal of forecasting **16**(2): 125-145.

Brown, T. R. and R. C. Brown (2013). "A review of cellulosic biofuel commercial-scale projects in the United States." Biofuels, bioproducts and biorefining **7**(3): 235-245.

Brown, T. R. and R. C. Brown (2013). "Techno-economics of advanced biofuels pathways." RSC Advances **3**(17): 5758-5764.

Center for Farm Financial Management. University of Minnesota. (2015). "Whole Farm Financial Summary: All States, Crop Farms, 2009-2015." Retrieved January 26, 2017, from https://finbin.umn.edu/FmSummOpts/Index.

Centre for Agriculture and Bioscience International (CABI). (2016). "Thlaspi arvense (field pennycress)." Retrieved February 20, 2016, from http://www.cabi.org/isc/datasheet/27595.

Clark, A. (2012). Managing Cover Crops Profitably. Sustainable Agriculture Research and Education (SARE) program, with funding from the National Institute of Food and Agriculture, U.S. Department of Agriculture.

Clark, S. F. (2009). "The profitability of transitioning to organic grain crops in Indiana." American journal of agricultural economics **91**(5): 1497-1504.

Clements, M. and D. Hendry (1998). Forecasting economic time series, Cambridge University Press.

Cleveland, W. S. and C. Loader (1996). Smoothing by local regression: Principles and methods. Statistical theory and computational aspects of smoothing, Springer**:** 10-49.

Cryer, J. and K. Chan (2008). Time Series Analysis with Applications in R, New York, Springer.

Davidson, C., E. Newes, A. Schwab and L. Vimmerstedt (2014). "An Overview of Aviation Fuel Markets for Biofuels Stakeholders." National Renewable Energy Laboratory, Retrieved September **5**: 2014.

De Gooijer, J. G. and K. Kumar (1992). "Some recent developments in non-linear time series modelling, testing, and forecasting." International Journal of Forecasting **8**(2): 135-156.

Dijk, D. v., T. Teräsvirta and P. H. Franses (2002). "Smooth transition autoregressive models—a survey of recent developments." Econometric reviews **21**(1): 1-47.

Dugas, D. J. (1973). "Fuel from Organic Matter." Paper P-5100 **RAND Corporation. Santa Monica CA,**.

EcoEngineers. (2012). "RIN Price Index Calculation Details." Retrieved 12/21/2015, 2015, from http://rinindex.ecoengineers.us/Views/Public/Products.aspx.

Eitrheim, Ø. and T. Teräsvirta (1996). "Testing the adequacy of smooth transition autoregressive models." Journal of Econometrics **74**(1): 59-75.

English, B., J. Menard and D. D. L. T. Ugarte (2000). "Using Corn Stover for Ethanol Production: A Look at the Regional Economic Impacts for Selected Midwestern States." Department of Agricultural Economics, University of Tennessee, Knoxville, TN. Available at http://web. utk. edu/~ aimag/pubimpact. html.

English, B., J. Menard, Yu Edward T. and K. Jensen (2016). Analysis of the Impacts of a Cost-Based Surface for Biofuels Investment. D. o. A. a. R. E. Bio-Based Energy Analysis Group, The University of Tennessee Institute of Agriculture.

Epplin, F. M. (1996). "Cost to produce and deliver switchgrass biomass to an ethanol-conversion facility in the southern plains of the United States." Biomass and Bioenergy **11**(6): 459-467.

Fackler, P. L. (2008). Solving optimal switching models. http://www4.ncsu.edu/unity/users/p/pfackler/www/ECG766/switch.pdf.

Fan, J., D. R. Shonnard, T. N. Kalnes, P. B. Johnsen and S. Rao (2013). "A life cycle assessment of pennycress (Thlaspi arvense L.)-derived jet fuel and diesel." biomass and bioenergy **55**: 87-100.

Food and Agriculture Organization of the United Nations (1976). The State of Food and Agriculture. Agriculture Series No. 4. Italy, United Nations.

Franses, P. H. and D. Van Dijk (2000). Non-linear time series models in empirical finance, Cambridge University Press.

Good, D. and S. Irwin (2015). "The EPA's Proposed Ethanol Mandates for 2014, 2015, and 2016: Is There a'Push'or Not?" farmdoc daily **5**(5): 102).

Gowda, M. S. and R. Sznajder (1994). "The generalized order linear complementarity problem." SIAM Journal on Matrix Analysis and Applications **15**(3): 779-795.

Graham, R. L., B. C. English and C. E. Noon (2000). "A geographic information system-based modeling system for evaluating the cost of delivered energy crop feedstock." Biomass and bioenergy **18**(4): 309-329.

Granger, C. W. and T. Terasvirta (1993). "Modelling non-linear economic relationships." OUP Catalogue.

Grantham, J. B. and T. H. Ellis (1974). "Potentials of wood for producing energy." Journal of Forestry **72**(9): 552-556.

Guenther, G. (2015). The Section 179 and Bonus Depreciation Expensing Allowances: Current Law and Issues for the 114th Congress. CRS Report prepared for Members and Committees of Congress.

H.R. 110-627 (2008). House Report. 110-627 to accompany H.R. 2419. 110th Congress (2007-2008) at page 1048.

Irwin, S. (2014). "Understanding the Behavior of Biodiesel RINs Prices." farmdoc daily.

Irwin, S. (2016). "What's Up with RINs Prices?" farmdoc daily **6**(6): 188).

Irwin, S. and D. Good (2015). "Does it Matter Whether the EPA Targets Volumetric or Fractional RFS Standards?".

Johnson, G. A., M. B. Kantar, K. J. Betts and D. L. Wyse (2015). "Field pennycress production and weed control in a double crop system with soybean in Minnesota." Agronomy Journal **107**(2): 532-540.

Keenan, D. M. (1985). "A Tukey nonadditivity-type test for time series nonlinearity." Biometrika **72**(1): 39-44.

Kelly, J. T. and D. L. White (1974). Project Independence: Federal Energy Administration, Project Independence blueprint: final task force report. Washington, DC, For sale by the Supt. of Docs., U.S. G.P.O.

Lade, G. E., C.-Y. C. Lin and A. Smith (2015). Policy Shocks and Market-Based Regulations: Evidence from the Renewable Fuel Standard, Working Paper.

Lade, G. E., C.-Y. C. Lin and A. Smith (2015). "Policy shocks and market-based regulations: Evidence from the Renewable Fuel Standard." University of California–Davis Working Paper DEEP WP **8**.

Lailas, N. (1989). "Advancing biofuels technology acceptance in the USA: Part I. The department of energy's regional biomass energy program." Biomass **19**(3): 195-213.

Lawton, K. (2016). Soybean Digest – The 1970s. Corn and Soybean Digest.

Markel, E., C. D. Clark and D. M. Lambert (2014). Renewable Fuel Standard and Treatment of Woody Biomass. Southeastern Partnership for Integrated Biomass Supply Systems (IBSS).

Mason, B., M. E. Phippen and W. B. Phippen. (2013). "Field Pennycress (Thlaspi arvense L.) Response to Nitrogen and Sulfur Rates." from http://www.wiu.edu/pennycress/current-experiments/mason-poster-final.pdf.

McCarty, T. and J. Sesmero (2014). "Uncertainty, irreversibility, and investment in second-generation biofuels." BioEnergy Research **8**(2): 675-687.

McPhail, L., P. Westcott and H. Lutman (2011). "The renewable identification number system and US biofuel mandates." ERS Report BIO-03, USDA, November.

Meyer, S. and W. Thompson (2012). "How do biofuel use mandates cause uncertainty? United States environmental protection agency cellulosic waiver options." Applied Economic Perspectives and Policy: pps033.

Moser, B. R., G. Knothe, S. F. Vaughn and T. A. Isbell (2009). "Production and Evaluation of Biodiesel from Field Pennycress (Thlaspi arvense L.) Oil†." Energy & Fuels **23**(8): 4149-4155.

Oursbourn, C., R. D. Lacewell, W. Lepori and W. P. Patton (1978). "Energy potential from agricultural residues in Texas." Southern Journal of Agricultural Economics **10**(02): 73-80.

Parker, K. and W. B. Phippen (2012). Impact of Heat Stress on Field Pennycress Seed Yield and Pollen Viability.

Pearlson, M., C. Wollersheim and J. Hileman (2013). "A techno-economic review of hydroprocessed renewable esters and fatty acids for jet fuel production." Biofuels, Bioproducts and Biorefining **7**(1): 89-96.

Perlack, R., J. Ranney, W. Barron, J. Cushman and J. Trimble (1986). "Short-rotation intensive culture for the production of energy feedstocks in the US: a review of experimental results and remaining obstacles to commercialization." Biomass **9**(2): 145-159.

Phippen, W. B., B. John, M. Phippen and T. Isbell (2010). Planting date, herbicide, and soybean rotation studies with field pennycress (Thlaspi arvense L.). poster presentation at the 22nd Annual Meeting of the Association for the Advancement of Industrial Crops, Fort Collins, CO, USA.

Phippen, W. B. and M. E. Phippen (2012). "Soybean seed yield and quality as a response to field pennycress residue." Crop Science **52**(6): 2767-2773.

Ray, D., D. D. L. T. Ugarte, M. Dicks and K. Tiller (1998). "The POLYSYS modeling framework: A documentation." Agricultural Policy Analysis Center, University of Tennessee, Knoxville, Tennessee. Available at http://agpolicy. org/polysys. htm.

Reese, R. A., S. V. Aradhyula, J. F. Shogren and K. S. Tyson (1993). "Herbaceous biomass feedstock production: the economic potential and impacts on US agriculture." Energy policy **21**(7): 726-734.

Schmit, T. M., J. Luo and J. M. Conrad (2011). "Estimating the influence of US ethanol policy on plant investment decisions: A real options analysis with two stochastic variables." Energy Economics **33**(6): 1194-1205.

Schmit, T. M., J. Luo and L. W. Tauer (2009). "Ethanol plant investment using net present value and real options analyses." biomass and bioenergy **33**(10): 1442-1451.

Schnitkey, G. (2017). 2017 Illinois Crop Budgets. Farm Business Management. Department of Agricultural and Consumer Economics University of Illinois.

Sedbrook, J. C., W. B. Phippen and M. D. Marks (2014). "New approaches to facilitate rapid domestication of a wild plant to an oilseed crop: Example pennycress (Thlaspi arvense L.)." Plant Science **227**: 122-132.

Shapouri, H. and J. Duffield (1993). "The economics of producing energy crops." US Department of Agriculture, Office of Energy Washington, DC **20250**.

Sharma, N., D. Cram, T. Huebert, N. Zhou and I. A. Parkin (2007). "Exploiting the wild crucifer Thlaspi arvense to identify conserved and novel genes expressed during a plant's response to cold stress." Plant molecular biology **63**(2): 171-184.

Shumaker, G. A., J. C. McKissick, J. Woodruf and B. A. Doherty (2003). "Georgia oilseed initiative: report on the feasibility of an oilseed processing facility in Georgia."

Sindelar, A. J., M. R. Schmer, R. W. Gesch, F. Forcella, C. A. Eberle, M. D. Thom and D. W. Archer (2015). "Winter oilseed production for biofuel in the US Corn Belt: Opportunities and limitations." GCB Bioenergy.

Steffgen, F. W. (1974). "Energy from agricultural products." A New Look at Energy Sources(anewlookatenerg): 23-35.

Stock, J. H. (2015). "The Renewable Fuel Standard: A Path Forward." Columbia, SIPA Center on Global Energy Policy. April.

Stock, J. H. and M. W. Watson (1998). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series, National Bureau of Economic Research.

Strauss, C., P. Blankenhorn, T. Bowersox and S. Grado (1988). "Financial and energy costs of supplying woody biomass to conversion sites." Applied biochemistry and biotechnology **18**(1): 217-230.

Szego, G. C. and C. C. Kemp (1973). "Energy forests and fuel plantations." Chem. Technol.;(United States) **3**(5).

Teräsvirta, T. (2006). "Forecasting economic variables with nonlinear models." Handbook of economic forecasting **1**: 413-457.

The Tax Foundation (2013). U.S. Federal Individual Income Tax Rates History, 1862-2013 (Nominal and Inflation-Adjusted Brackets).

Thompson, W., S. Meyer and P. Westhoff (2009). "Renewable Identification Numbers are the Tracking Instrument and Bellwether of US Biofuel Mandates." EuroChoices **8**(3): 43-50.

Thompson, W., S. Meyer and P. Westhoff (2010). "The new markets for renewable identification numbers." Applied Economic Perspectives and Policy **32**(4): 588-603.

Thompson, W., S. Meyer and P. Westhoff (2011). "What to conclude about biofuel mandates from evolving prices for renewable identification numbers?" American Journal of Agricultural Economics: aaq120.

Tiao, G. C. and R. S. Tsay (1994). "Some advances in non-linear and adaptive modelling in time-series." Journal of forecasting **13**(2): 109-131.

Tsay, R. S. (1986). "Nonlinearity tests for time series." Biometrika **73**(2): 461-466.

U.S. Congress Office of Technology Assessment. USOAT (1979). Gasohol: A Technical Memorandum. Washington, Congress of the United States, Office of Technology Assessment : for sale by the Supt. of Docs. U.S. Govt. Print. Off.

U.S. Department of Agriculture, O. o. E. (1986). "Fuel Ethanol and Agriculture: An Economic Assessment." Agricultural Economic Report, Economic Report Number 562.

U.S. Department of Agriculture. Animal and Plant Health Inspection Service (2015). Weed Risk Assessment for Thlaspi arvense L. (Brassicaceae) - Field pennycress.

U.S. Department of Energy. (2016). "Energy Timeline: 1971 to 1980." Retrieved February 4, 2016, from http://web.archive.org/web/20110721035950/http://www.energy.gov/about/timeline1971-1980.htm.

U.S. Department of Energy, E. I. A. (2016). "Short-Term Energy Outlook: Real Prices Viewer." Retrieved February 9, 2016, from http://www.eia.gov/forecasts/steo/realprices/.

U.S. Department of Energy. Alternative Fuels Data Center. (2016). "Biodiesel Blends." from http://www.afdc.energy.gov/fuels/biodiesel_blends.html.

U.S. Department of Energy. Energy Information Administration. (2016). "Petroleum and other Liquids: Definitions, Sources and Explanatory Notes." from https://www.eia.gov/dnav/pet/TblDefs/pet_sum_sndw_tbldef2.asp.

U.S. Department of Transportation, Federal Aviation Administration (FAA),. (2015). "Aviation Greenhouse Gas Emissions Reduction Plan." from http://www.icao.int/environmental-protection/Lists/ActionPlan/Attachments/30/UnitedStates_Action_Plan-2015.pdf.

U.S. Department of Transportation BTS (2014). Air Carrier Financial Reports. Form 41, Schedule P-12(a) and P-6.

U.S. Environmental Protection Agency (2013). Environmental Protection Agency Finalizes 2013 Renewable Fuel Standards.

U.S. Environmental Protection Agency (2014). Renewable Fuel Standard : Obligated Parties and Exporters of Renewable Fuel. **2015**.

U.S. Environmental Protection Agency (2015). Notice of Opportunity to Comment on an Analysis of the Greenhouse Gas Emissions Attributable to Production and Transport of Pennycress (Thlaspi Arvense) Oil for Use in Biofuel Production[EPA-HQ-OAR-2015-0091; FRL-9924-65-OAR].

U.S. Environmental Protection Agency (2015). Renewable Fuel Standard Program: Standards for 2014, 2015, and 2016 and Biomass-Based Diesel Volume for 2017; Proposed Rule. 40 CFR Part 80. Washington D.C.

U.S. Environmental Protection Agency. (2016). "Renewable Fuel Standard Program: Fuel Nesting Scheme for Renewable Fuel Standard (RFS)." from https://www.epa.gov/renewable-fuel-standard-program/renewable-fuel-annual-standards.

U.S. Environmental Protection Agency (2016). Renewable Fuel Standard. Approved Pathways for Renewable Fuel. Table 1. Title 40, Chapter I, Subchapter C, Part 80, Subpart M. U.S. Environmental Protection Agency.

Ugarte, D. D. L. T., M. E. Walsh, H. Shapouri and S. P. Slinsky (2000). The economic impacts of bioenergy crop production on US agriculture. USDA.

Ugarte, D. G. D. L. T. and D. E. Ray (2000). "Biomass and bioenergy applications of the POLYSYS modeling framework." Biomass and Bioenergy **18**(4): 291-308.

University of Illinois Urbana Champagne College of Agricultural Consumer and Environmental Sciences. (2013). "Soybean Planting Dates." College News, from http://news.aces.illinois.edu/news/soybean-planting-dates.

UOP. (2016). "Honeywell Green Jet Fuel – Advanced Renewable Fuel Alternative to Traditional Jet Fuel." from http://www.uop.com/processing-solutions/renewables/green-jet-fuel/.

UOP, H. (2016). "Advanced Renewable Fuel Alternative to Traditional Diesel." Retrieved April 7 2016, 2016, from http://www.uop.com/processing-solutions/renewables/green-diesel/#biodiesel.

Walsh, M. E., G. Daniel, H. Shapouri and S. P. Slinsky (2003). "Bioenergy crop production in the United States: potential quantities, land use changes, and economic impacts on the agricultural sector." Environmental and Resource Economics **24**(4): 313-333.

Whistance, J. and W. Thompson (2014). "A Critical Assessment of RIN Price Behavior and the Implications for Corn, Ethanol, and Gasoline Price Relationships." Applied Economic Perspectives & Policy **36**(4).

Yacobucci, B. D. (2012). "Biofuels Incentives: A Summary of Federal Programs."

Zerbe, J. I. (1988). Biofuels: Production and potential. Forum for applied research and public policy.

Vita

The author is an applied economist with a passion for natural resources that support the functioning of our society.