2009

# Three essays concerning agriculture and energy

Mindy Lyn Baker
*Iowa State University*

**Three essays concerning agriculture and energy**

by

**Mindy Lyn Baker**

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Dermot Hayes, Major Professor
Bruce Babcock
Arne Hallam
Chad Hart
Sergio Lence

Iowa State University

Ames, Iowa

2009

# TABLE OF CONTENTS

## ACKNOWLEGEMENTS

I am grateful for the support of many people in my pursuit of this degree and in the completion of this dissertation. Dr. Dermot Hayes, my major professor, was an inexhaustible reserve of encouragement, enthusiasm, and patience; without healthy doses of each of these, I could not have endured. As an economist, I admire his intuition about important problems; as a friend, I admire his loyalty, generosity, and humor. I hope to emulate him in both of these planes.

My committee members: Dr. Bruce Babcock, Dr. Arne Hallam, Dr. Chad Hart, and Dr. Sergio Lence provided invaluable guidance in the writing of this dissertation. Their recommendations not only enhanced this document, but my professional development as well.

Friends who are liberal with encouragement, as well as red ink, are rare gems. I extend my thanks to Subhra Bhattacharjee, Jerome Dumortier, Carola Grebitus, Lihong McPhail, and Ofir Rubin in particular for reading and commenting on drafts of my work at a moment's notice.

My parents, Carol and Linda Mallory, whose advice I seek more and more, gave me the greatest gifts parents can bestow on a child: curiosity, motivation, and most of all love that is not conditioned on success or failure.

Lastly, I thank my husband Casey Baker. The patience and devotion he displayed throughout this process were nothing short of remarkable. As much as I sacrificed for this degree, he sacrificed more.

## ABSTRACT

This dissertation consists of three papers, each regarding a particular aspect of the relationship between energy and agriculture. The objective of the first paper is to create a model that will enhance informed policy decisions regarding the bioeconomy. A forward-looking stochastic model captures the effect of uncertainty in crude oil prices and commodity yields on biofuel industry development. Acreage limitations on feedstocks such as corn, soybeans, and switchgrass are shown to create competition for acreage among the crops. Investors in the model are rational in the sense that they engage in biofuel production only if returns exceed what they expect to earn from alternative investments.

The Energy Independence and Security Act of 2007 mandates the use of 36 billion gallons of biofuels by 2022 with significant requirements for cellulosic biofuel and biodiesel production. In the model, the price wedge created by mandated biofuel production at these levels is \$2.50 per gallon for biodiesel and \$1.07 per gallon for cellulosic biofuel. Long-run commodity prices were high in our simulation, with corn at \$7.38 per bushel and soybeans at \$19.57 per bushel. Intense competition for planted acreage drives the high commodity prices.

The second paper develops a model of the corn, soybean, and wheat markets to calculate welfare effects of increased biofuel production in the United States. Demand is disaggregated into livestock feed, food, energy. Uncertain crop yields permit the valuation of farm deficiency payments as options. Incorporating soybean and wheat markets capture indirect welfare effects through equilibrium price increases. Net welfare loss ranges from \$200 million to \$750 million depending on the size of biofuel increase. Consumers make a

sizable transfer to farmers. The sign of the net costs to taxpayers depends on the size of the biofuel industry.

In the third paper, the nature of the relationship between corn and ethanol prices is explored. Economic fundamentals should require that the price of corn and ethanol maintain a long run equilibrium relationship. The relationship is driven by a long run condition that says entry and exit in the industry will occur maintaining no sustained profits or losses for the industry. Both ethanol producers and traditional users of corn have a stake in the behavior of these markets, and their profitability will rely on their ability to determine accurately this relationship. I test for cointegration of these price series and find evidence that corn and ethanol prices are indeed maintain an equilibrium relationship.

Statistical cointegration tests are known to have problems in small samples. This is a potential issue in interpreting the results mentioned above because ethanol production only recently constituted a significant portion of the corn crop. With only a few years of the most recent data for which we suspect that an equilibrium relationship existed the small sample properties of cointegration tests are important particularly in this application. A Monte Carlo study tailored to mimic our actual data set is conducted. I find that the corn and ethanol price series are not long enough to rely on the asymptotic properties of the cointegration statistics, and therefore one should use small sample critical values in this kind of analysis.

# CHAPTER 1: GENERAL INTRODUCTION

## Introduction

It is clear that agricultural markets are increasingly interdependent with energy markets. U.S. agriculture has long been a consumer of energy as an input, but with increased biofuel production stimulated by high transportation fuel prices and governmental mandates agriculture is becoming a major producer of energy as well. This new facet in the relationship between agriculture and energy markets presents challenges and opportunities for individual producers of agricultural commodities and policy makers. This dissertation will explore the nature of these new relationships, what they mean to with regard to existing policies and to traditional players in agricultural markets.

The U.S. government's biofuel policy has been a mix of per unit tax credits and annual usage mandates. Increasing biofuel production from commodities grown on a (nearly) fixed amount of land has profound implications for both the traditional users of the commodities and for the cost to the taxpayer to maintain these programs.

On one hand, a production mandate requires a certain quantity of corn, soybeans, or cellulosic feedstock be devoted to biofuel production. With a (nearly) fixed quantity of land in agricultural production, a binding biofuel mandate causes prices of all agricultural commodities to rise. This means that the cost of producing biofuels increases since the largest cost is in purchasing feedstock, which causes the level of support required in the form of tax credits and subsidies to increase to keep biofuel producers in production.

It has been argued that increased commodity prices caused by biofuel production have decreased taxpayer's liability in farm programs in the form of loan deficiency payments and counter cyclical payments. The analysis in this dissertation finds that in fact, increasing biofuel production has a nonlinear effect on farm program liabilities. A little bit of biofuel production increases commodity prices out of the range where they are likely to trigger farm program payments, but once commodity prices are high enough further increase makes almost no difference in the taxpayer's expected payment to the farmer because this value is already almost zero.

Further, previous research into the welfare effects of U.S. biofuel policy has focused primarily on the welfare effects driven by action in the energy markets. This dissertation focuses on welfare transfers occurring in the agricultural markets, disaggregating demands to show the relative size of transfer among different commodity using sectors. Further, in this dissertation indirect welfare effects are considered in the soybean and wheat markets whereas previous literature only discusses welfare effects in the corn market.

This dissertation partially fills a gap left by previous research concerning the evolving relationship of energy and agriculture. It attempts to provide insight into the way biofuel policy affects commodity markets, existing farm programs, and the welfare effects felt by economic agents in these markets.

**Dissertation Organization**

This dissertation is organized into five chapters. The present chapter contains a general introduction into the subject at hand, including motivation as to why the joint study of energy and agriculture is needed, and a describes the general organization of the dissertation. Chapters two, three, and four are each concerned with a specific question regarding the relationship of energy and agriculture. Each of these chapters can be read independently as they are largely self contained.

In chapter 2 the objective is to create a model that will enhance informed policy decisions regarding the bioeconomy. A forward-looking stochastic model captures the effect of uncertainty in crude oil prices and commodity yields on biofuel industry development. Acreage limitations on feedstocks such as corn, soybeans, and switchgrass create competition for acreage among the crops, which has to be taken into account. Investors in the model are rational in the sense that they engage in biofuel production only if returns exceed what they expect to earn from alternative investments. These two insights drive the result that competition for acreage drives up commodity prices as well as increases the cost of supporting biofuel industries which rely on land intensive crops.

Chapter 3 develops a model of the corn, soybean, and wheat markets to calculate welfare effects of increased biofuel production in the United States. Demand is disaggregated into livestock feed, food, energy. Uncertain crop yields permit the valuation of farm deficiency payments as options. Incorporating soybean and wheat markets capture indirect welfare effects through equilibrium price increases. Net welfare loss ranges from $200 million to $750 million depending on the size of biofuel increase. Consumers make a sizable transfer to farmers. The sign of the net costs to taxpayers depends on the size of the biofuel industry.

In chapter 4 the nature of the relationship between corn and ethanol prices is explored. Economic fundamentals should require that the price of corn and ethanol maintain a long run equilibrium relationship. The relationship is driven by a long run condition that says entry and exit in the industry will occur maintaining no sustained profits or losses for the industry. Both ethanol producers and traditional users of corn have a stake in the behavior of these markets, and their profitability will rely on their ability to determine accurately this relationship. I test for cointegration of these price series and find evidence that corn and ethanol prices are indeed maintain an equilibrium relationship.

Statistical cointegration tests are known to have problems in small samples. This is a potential issue in interpreting the results mentioned above because ethanol production only recently constituted a significant portion of the corn crop. With only a few years of the most recent data for which we suspect that an equilibrium relationship existed the small sample properties of cointegration tests are important particularly in this application. A Monte Carlo study tailored to mimic our actual data set is conducted. I find that the corn and ethanol price series are not likely long enough to rely on the asymptotic properties of the cointegration statistics, and therefore one should use small sample critical values in this kind of analysis.

The final chapter contains general conclusions of the dissertation, and summarizes the most interesting findings.

# CHAPTER 2: CROP-BASED BIOFUEL PRODUCTION WITH ACREAGE CONSTRAINTS AND UNCERTAINTY

## Introduction

Congress signed into law the Energy Independence and Security Act (EISA) in December 2007.[1] The Renewable Fuel Standard, or RFS, in the EISA mandated the use of 36 billion gallons of biofuels by 2022, of which 15 billion gallons can come from corn-based ethanol and 21 billion must come from advanced biofuels — including 16 billion of cellulosic biofuels and 1 billion from biomass based diesel. The text of the act provides specific year-by-year ramp up targets, even though growth in the production of corn-based ethanol already was strong. In the year 2000, corn-based ethanol production was 1.63 billion gallons, and by the end of 2007 production reached 6.50 billion gallons.[2] The increase in corn-based ethanol production led to record high nominal corn prices in 2008, and competition for acreage transferred the demand pressure in corn markets to other crops — soybeans and hay, for example — causing the prices of these to increase as well. This paper examines the incentives required to encourage production of the mandated quantities of biofuels and explore the impact of these on U.S. agriculture.

We present a model based on the assumption that one can predict biofuel production levels if one understands the factors influencing the decisions made by agents in the economy. Farmers make planting decisions based on expected market prices. Further,

---

[1] HR 6, The Energy Independence and Security Act of 2007, available at http://thomas.loc.gov/cgi-bin/bdquery/z?d110:h.r.00006:

[2] Industry reported levels found at http://www.ethanolrfa.org/industry/statistics/

farmers recognize that land devoted to biofuel feedstock production has an opportunity cost, they will only grow biofuel feedstock such as switchgrass for cellulosic ethanol production if it is profitable to do so. Investors who build biofuel plants do so only if they expect a risk-adjusted return on par with or superior to investments made elsewhere in the economy. Existing plants operate only if the marginal cost of production is less than the value of output. Those who blend and use biofuels do so only if the market price of ethanol is less than the prices of alternatives.

Taking each decision just described, as well as parameters and data from the literature, we model the decisions of relevant agents in the economy and the market forces guiding them. We combine the resulting sub-models within a stochastic simulation model of U.S. crop and biofuel markets and calibrate it to reflect actual market conditions as of the spring of 2008. We evaluate the response of market participants to changes in incentives using this model.

The expansion of biofuel production happened quickly, and over a short number of growing seasons. The short amount of time for which data is available creates challenges in estimating an econometric model, but policymakers still need a tool with which to look forward. In order to help fill this gap we develop a model that simulates the decisions of important players in the economy; providing insight about the implications of many policy choices.

**Previous Literature**

Elobeid et al. (2007) provided the first comprehensive model of the bioeconomy. Later, Tokgoz et al. (2007) expand this work, strengthening some of its elements. They include equilibrium relationships for prices of biofuel co-products, most notably distillers grains. Both use the world agricultural model from the Food and Agricultural Policy Research Institute (FAPRI) to determine the potential size of the corn-based ethanol sector and to describe how it affects crop and livestock markets. First, they assume biofuel investments will continue until expected profit is zero, and they calculate the break-even corn price that drives margins on new corn-based ethanol plants to zero. Second, they assume this corn price clears the market, and they calculate the size of the biofuel sector required to bring about this price. Finally, once they determine the break even corn price, they evaluate its impact on U.S. and world agriculture. They ignore biofuels from cellulose and biodiesel because their results suggest these are not economically viable. They also ignore risks associated with investments in biofuel plants.

Our model enhances this work by incorporating awareness of risk into the decision problem of the biofuel investor. Returns to biofuel production primarily are a function of energy and feedstock prices, which are uncertain. Crude oil prices are the main determinant of the prices of transportation fuels such as gasoline, ethanol, and diesel, and crop yields affect feedstock costs through its effects on equilibrium commodity prices.

Incorporating uncertainty in crude oil prices and crop yields allows us to compare the risk-adjusted return in the production of each type of biofuel, determining which types are attractive to investors. Basing the investor's decision on risk-adjusted returns is more realistic than using a zero profit condition, which implies a risk-neutral investor. A stochastic model that delivers probability distributions over commodity prices and returns in the biofuel industry allows us to build a model in which the investor cares not only about the mean of returns but also about the variance.

**The Model Economy**

The agricultural economy has three goods: corn, soybeans, and switchgrass. [3] Switchgrass is a perennial grass native to the U.S. tall-grass prairie, and scientists consider it a good candidate feedstock for cellulosic ethanol production (Schmer, et al., 2008). The three major players in the economy are farmers, consumers (who are processors of cereal grains, oilseeds, and livestock), and investors. The consumers buy agricultural commodities and use them as input in producing either food or energy; investors can choose to build a corn ethanol plant, a biodiesel plant, or a switchgrass ethanol plant. Alternatively, they can simply choose to invest in an alternative we call the market portfolio.

We introduce uncertainty through agricultural commodity yields and crude oil prices and assume the two random variables are independent with joint probability distribution $f(\zeta, \varepsilon) = g(\zeta)h(\varepsilon)$ where $\zeta$ is a vector of yield realizations and $\varepsilon$ is the realization of crude oil prices. By specifying commodity yields and crude oil prices as independent, we implicitly assume that shocks to domestic biofuel production do not influence world crude oil price shocks.

Agents form expectations about crude oil prices and future crop yields. Farmers allocate acreage among corn, soybeans, and switchgrass, and investors plan long-run capacity in the biofuel sectors. Acreage allocation is carried out implicitly through time, and we

[3] We consider only corn, soybeans, and switchgrass because we focus on the decision of a farmer who must allocate crop ground. Other cellulosic feedstocks such as woodchips are not well suited to crop ground (Lewandowski et al., 2003).

assume that farmers choose the proportion of land in annual crops (an equilibrium corn-soybean rotation) and land in the perennial, switchgrass. With this we avoid dealing with the difference in the timing of cost and payoff between the annual and perennial crops; what we lose in specificity in the farmer's yearly decision is not pertinent to the results and permits us to work with a cleaner model where we can focus on the long run results.

Supply is determined by acreage allocations and by crop yield realizations. Demand for the commodities comes from the livestock feed, human food and export sectors but is represented in the model by an aggregate demand function. Demand from the biofuel sectors is determined by production capacity, which is determined by a long run equilibrium condition. We describe these components of the model in more detail below.

*Commodity Supply*

There exists a single representative and competitive farmer with an endowment of land, who takes both output prices and his cost function as given. The farmer allocates his land to three different crops: corn, soybeans, and switchgrass. The crops are indexed as follows: corn, *i* = 1; soybeans, *i* = 2; and switchgrass, *i* = 3. The endowed land is representative of total U.S. cropland devoted to these commodities.

The farmer's per acre profit is given by

$$w = \sum_{i=1}^{3} p_i \cdot \zeta_i \cdot \pi_i - c_i\left(\pi_i \,;\, \Theta_i\right) \tag{1}$$

where $p_i$ is crop *i*'s output price, $\zeta_i$ is the realized per acre yield of crop *i*, and $\pi_i$ is the proportion of crop land allocated to crop *i*. The cost function for crop *i* is $c_i\left(\pi_i; \Theta_i\right)$, where $\Theta_i$ is a vector of parameters defining each crop's cost function. The set of parameters of the cost functions of each crop are defined by $\boldsymbol{\Theta} = \{\Theta_1, \Theta_2, \Theta_3\}$. Aggregate (national) profit is then calculated by multiplying the per acre profit times the amount of acres in production.

The farmer is risk neutral in profit; he wishes to maximize expected profit subject to land constraints. To this end, he chooses a land allocation vector, $[\pi_1 \quad \pi_2 \quad \pi_3]'$, to solve the problem:

(2) $$\max_{\pi_1, \pi_2, \pi_3 \geq 0} E[w] \;\; s.t. \; \sum_{i=1}^{3} \pi_i \leq 1$$

If we denote for crop *i* the expected price by $\bar{p}_i$, the expected yield for crop *i* by $\mu_i$, and the shadow value of land by $\lambda^*$, the Kuhn-Tucker conditions for this problem are:

$$\bar{p}_i \mu_i - \frac{\partial c_i(\pi_i^*; \Theta_i)}{\partial \pi_i} - \lambda^* \leq 0, \quad \pi_i^* \geq 0, \quad \pi_i^* \left[ \bar{p}_i \mu_i - \frac{\partial c_i(\pi_i^*; \Theta_i)}{\partial \pi_i} - \lambda^* \right] = 0 \quad \text{for } i = 1, 2, 3$$

and

$$1 - \pi_1^* - \pi_2^* - \pi_3^* \leq 0, \quad \lambda^* \geq 0, \quad \lambda^* \left[ 1 - \pi_1^* - \pi_2^* - \pi_3^* \right] = 0 .$$

Assuming an interior solution, the first order conditions are:

(3) $$\bar{p}_i \mu_i - \frac{\partial c_i}{\partial \pi_i} = 0 \quad \text{for } i = 1, 2, 3$$

The farmer's price expectations are assumed to be such that, when combined with demand, cause the ex-ante (but post planting) price distribution to have a mean of $\bar{\mathbf{p}} = [\bar{p}_1 \quad \bar{p}_2 \quad \bar{p}_3]'$.

The optimal acreage decisions combined with yield realizations, $\zeta_i$, give the supply function for each crop:

(4) $$Q_i^s(\bar{\mathbf{p}}; \boldsymbol{\mu}, \boldsymbol{\Theta}, \zeta_i) = \zeta_i \; \pi_i^*(\bar{\mathbf{p}}; \boldsymbol{\mu}, \boldsymbol{\Theta})$$

Notice that both the expected output price and production cost of the other crops, enter each crop's supply function.

*Commodity Demand*

Demand for commodity *i* is denoted by $Q_i^d\left(\mathbf{p}\ ,n_i\ ;\Omega_i\right)$. The demands are a function of commodity prices, $\mathbf{p}=\left[p_1 \quad p_2 \quad p_3\right]'$, and the number of biofuel plants in operation, $n_i$ . The set of parameters defining the demand function is denoted by $\Omega_i$, and the set of all three demand parameters is defined by $\mathbf{\Omega}=\left\{\Omega_1,\Omega_2,\Omega_3\right\}$. Later, we implement the model and specify functional forms for the demand equations.

*The Investors*

In each period, investors can choose among four different options: a corn-based ethanol plant, a biodiesel plant, a switchgrass ethanol plant, or an alternative investment we call the market portfolio — as in the capital asset pricing model (CAPM) of Sharpe (1964). The market portfolio functions as an option if none of the biofuel investments are attractive.

We assume investors seek the largest risk-adjusted return on investment possible and we assume that there exists a riskless asset in the economy returning the risk-free rate, *RFR*. Investors use the CAPM to evaluate investment alternatives; they calculate the security market line to measure the required rate of return for an asset, *a*.

(5) $Required\ \ Return_a = RFR + \beta_a\left(R_M - RFR\right)$

where *M* is the market portfolio and $R_M$ is the expected return of the market portfolio. $\beta_a$ is defined by

(6) $$\beta_a = \frac{Cov(R_a, R_M)}{\sigma_M^2}$$

where the variance of returns on the market portfolio is $\sigma_M^2$ and $R_a$ is the return of asset *a*.

The investors calculate the difference in the expected return and required return of asset *a* as calculated with the CAPM. The investors choose the project with the highest excess returns over the required return. However, if the difference is negative for each of the biofuel plants, an investor will choose to invest in the market portfolio. As long as the biofuel industry earns excess returns over the required return, it will continue to attract investment, and thus continue to expand.

*Returns to Biofuel Production*

Input costs in each sector are determined by feedstock, production, and other capital costs. We do not consider technological advancement in the production of biofuels; non-feedstock production costs and capital costs are exogenous in the model. Feedstock costs are the most important input cost to biofuel production, and these are determined by market equilibrium.

The annualized per gallon rate of return to producing biofuel of type *a* is $R_a = \frac{q_a(\varepsilon)}{k_a\left(p_i^{pergal}\right)}$.

The effective price received by the plant for its product is $q_a(\varepsilon)$; that is, the market price of the biofuel plus any incentive offered by the government such as the blenders tax credit.

(7) $$q_a(\varepsilon) = p_a(\varepsilon) + tax\ credit_a$$

The market price of biofuel, $p_a(\varepsilon)$, is a function of the crude oil price realization, $\varepsilon$. The per gallon cost of producing biofuel of type *a* is $k_a\left(p_i^{pergal}\right)$,

(8) $$k_a\left(p_i^{pergal}\right) = p_i^{pergal} + NFcost_a$$

where the per gallon feedstock cost is denoted by $p_i^{pergal}$. That is, in the case of corn-based ethanol production, $p_1^{pergal}$ is the equilibrium price of corn transformed into a cost per gallon of ethanol, in the case of biodiesel production, $p_2^{pergal}$ is the equilibrium price of soybeans transformed into a cost per gallon of biodiesel, and in the case of cellulosic ethanol, $p_3^{pergal}$ is the equilibrium price of switchgrass transformed into a per cost per gallon of ethanol. The non-feedstock cost of producing biofuel of type *a* is $NFcost_a$; this includes capital cost which is expressed per gallon and on an annual basis.

This relationship essentially measures the annual return on capital invested over input costs of the plant. Notice that the production cost depends on the realization of the commodity prices. For example, in a year when crop yields are relatively poor production costs will be higher than expected due to high equilibrium commodity (feedstock) prices. Likewise, high energy prices imply high returns in the biofuel sectors because the output price of biofuel will be high.

*Long Run Competitive Equilibrium*

In our agricultural economy, a long-run competitive equilibrium is defined by

pricing functions $p_i\left(\zeta, \varepsilon, \mathbf{n}, \mathbf{\Omega}, \mathbf{\Theta}\right)$ for $i = 1,\ 2,\ 3$

crop demand functions $Q_i^d\left(\mathbf{p}; n_i, \mathbf{\Omega}\right)$ for $i = 1,\ 2,\ 3$

crop supply functions $Q_i^s\left(\bar{\mathbf{p}}; \mathbf{\Theta}, \zeta\right)$ for $i = 1,\ 2,\ 3$

investment functions $n_i\left(\bar{\mathbf{p}},\bar{\varepsilon}\right)$ for $i=1,\ 2,\ 3$

Given the pricing functions, biofuel plants in operation, crop yield realizations, and crude oil prices, commodity markets clear. That is,

(9) $Q_i^s\left(\bar{\mathbf{p}};\boldsymbol{\Theta},\zeta\right)=Q_i^d\left(\mathbf{p};n_i,\boldsymbol{\Omega}\right)$ for $i=1,\ 2,\ 3$

The long-run equilibrium condition requires that, at the margin, the returns of each project equal the required risk-adjusted returns as determined by the CAPM:

$$R_{corn\ ethanol}\left(\mathbf{p}^*,n_1^*\right)=RR_{corn\ ethanol}$$
$$R_{biodiesel}\left(\mathbf{p}^*,n_2^*\right)=RR_{biodiesel}$$
$$R_{switch\ ethanol}\left(\mathbf{p}^*,n_3^*\right)=RR_{switch\ ethanol}$$

where $RR$ is the required return to biofuel prodiction as determined by the CAPM.

The zero excess return conditions ensure we have investment in each sector until the prices of feedstock (corn, soybeans, and switchgrass) are bid up to the point at which an investor is indifferent between any of the biofuel plants and the market portfolio. If the return to biofuel production is less than the required return for all industry sizes, then investment equals zero in this biofuel sector.

**Implementing the Model**

Our question is empirical in nature. The incentives present for the biofuel industry to expand or contract depend on many factors, some of which are the price of crude oil, demand for corn and soybeans for food uses, and weather variability. Exploring more than the most basic results of this model requires us to specify functional forms and evaluate the results numerically via the Monte Carlo method.[4] We calibrated the model to spring 2008, when producers of corn, soybeans, and switchgrass (hay) were making acreage decisions.

*Algorithm for Simulating the Model*

Our strategy for simulating the economy is to specify functional forms for both crop supply and demand and to calibrate the distribution of crude oil prices and crop yields. A joint draw from these distributions implies an equilibrium price for corn, soybeans, and switchgrass and thus implies return levels in each biofuel industry. The simulation algorithm is as follows:

1) Form crop yield and crude oil price distributions and make joint draws. Calculate biofuel prices from the crude oil price draws
2) Solve for the equilibrium crop prices for each draw using the market clearing conditions on crop supply and demand for a given level of biofuel capacity
3) Calculate the implied distribution of returns to biofuel production using the biofuel prices from (1) and equilibrium crop (feedstock) prices from (2)
4) Determine the tax credit or subsidy level on each biofuel type required in order for the long run (zero excess return) condition to be met

For example, setting the levels of biofuel production high (as in the EISA RFS) causes crop prices to increase, due to the increase in demand for these commodities. This increase in crop (feedstock) prices causes returns to biofuel production to shrink and ordinarily would cause the industry to contract. If this is the case, we calculate the subsidy on the sale of biofuel output required to keep the plants just indifferent between operating and not. The following subsections describe the details of implementing this algorithm further.

*Commodity Supply*

We assume the nominal cost functions of the crops are quadratic, given by

[4] All simulations were conducted in Matlab.

$c_i\left(\pi_i\right) = a_i \pi_i + \kappa_i\left(\pi_i\right)^2 \quad \forall \; i = 1, 2, 3 \quad \forall \; i = 1, 2, 3$. The proportion of land allocated to each crop, $\pi_i$ , is the farmer's choice variable. This specification works well because we can separate out increasing and constant marginal production costs, and use the Economic Research Service (ERS) commodity cost and return budgets as estimates for corn and soybeans.[5] ERS does not keep data on cost and returns to producing grass hay; instead, we use a production budget from Ohio State University Extension in 2003.[6]

The yield realizations, $\zeta$ , are drawn from the joint beta distribution of yields using the algorithm developed by Magnussen (2004).

---

[5] Available at http://www.ers.usda.gov/Data/CostsAndReturns/. The parameters are calibrated to the most recent estimates available (2006 crop year), and inflated by the expected percent increase in crude oil prices from 2006 to 2012. The price of crude oil in 2006 was $60 per barrel; we run several different long-run crude oil price scenarios. From the ERS budgets we use fertilizer cost as a proxy for the increasing marginal cost portion, $\kappa_i$ , and all other operating costs as the constant marginal cost portion, $a_i$ .

[6] Available at http://aede.osu.edu/Programs/FarmManagement/Budgets/crops-2003/grass.pdf. These costs from 2003 are inflated to 2012 levels in the same manner as for corn and soybeans. As with corn and soybeans, we assume fertilizer is the only element of the increasing marginal cost portion, $\kappa_3$ , and all others contribute to the constant marginal cost portion, $a_3$ .

$$\zeta \sim \beta\left(\begin{bmatrix}\mu_{corn}\\ \mu_{soybean}\\ \mu_{switchgrass}\end{bmatrix}, \Sigma^{-1}, \mathbf{q}_{max}, \mathbf{q}_{min}\right)$$

$$\Sigma^{-1} = \begin{bmatrix}213.84 & 36.18 & 1.49\\ 36.18 & 10.29 & 0.31\\ 1.49 & 0.31 & 0.031\end{bmatrix},$$

$$\mathbf{q}^{t}_{\max} = \begin{bmatrix}\mu_{corn} + 2\sigma_{corn}\\ \mu_{soybean} + 2\sigma_{soybean}\\ \mu_{switchgrass} + 2\sigma_{switchgrass}\end{bmatrix},$$

$$\mathbf{q}^{t}_{\min} = \begin{bmatrix}\mu_{corn} - 3\sigma_{corn}\\ \mu_{soybean} - 3\sigma_{soybean}\\ \mu_{switchgrass} - 3\sigma_{switchgrass}\end{bmatrix}$$

The mean of the beta distribution is $\boldsymbol{\mu} = \begin{bmatrix}160.05 & 52.72 & 3.36\end{bmatrix}'$, with corn and soybeans in bushels per acre and switchgrass in tons per acre. Since trend yields are important (especially for corn), we need to set the mean at a particular year's trend level. Recall that the long run, in our model, is the length of time approximately necessary for the biofuel sectors to reach maturity given current technology. We set the trend at year 2012.[7, 8] The matrix $\Sigma^{-1}$ is the

[7] We assume yields follow a linear trend, which we estimated from historical yield data for the years 1980 through 2006, maintained by the National Agricultural Statistics Service (http://www.nass.usda.gov/):

$$\mu^{t}_{corn} = -3843.83 + 1.99t, \quad \mu^{t}_{soybean} = -993.52 + .52t, \quad \mu^{t}_{switchgrass} = -13.94 + .0086t$$

variance-covariance matrix for the yields of the three crops, and the σ are the standard deviations of each crop found in $\Sigma^{-1}$.

*Commodity Demand*

We specify a constant elasticity, reduced-form demand function for each crop; we use the intermediate term own- and cross-price demand elasticities for beef from the ERS/Penn State Trade Model[9] as our estimates. The price distribution of crude oil influences commodity demands indirectly through the number of biofuel plants of each type (corn ethanol, biodiesel, and cellulosic ethanol). In our simulation, crude oil prices are lognormal and calibrated to match current conditions in the futures market:[10, 11]

(10) $$Q_i^d\left(\mathbf{p}; n_i, \mathbf{\Omega}\right) = \alpha_0^i p_1^{\alpha_1^i} p_2^{\alpha_2^i} p_3^{\alpha_3^i} n_i^{\alpha_4^i} \quad \text{for } i = 1,\ 2, 3$$

One of the equilibrium conditions requires the number of biofuel plants in each industry to be such that there are no excess returns over the required return. The parameter

---

[8] This is given in per harvested acre. We use alfalfa as a proxy for switchgrass yields, since the tonnage per acre is approximately equivalent to the switchgrass yields projected in the literature.

[9] Model documentation of the ERS/Penn State Trade Model can be found at http://trade.aers.psu.edu/pdf/ERS_Penn_State_Trade_Model_Documentation.pdf.

[10] The prices of other fuels (e.g., gasoline and biodiesel) are based on their relationship to crude oil prices.

[11] Implied volatility in crude oil prices is estimated from 2007 option data. We vary the mean of the crude oil price distribution in different scenarios: \$100, \$150, and \$200 per barrel.

$\alpha_4^i$ is an elasticity measuring the percentage change in quantity demanded of the crop over the percentage change in biofuel capacity, which depends only on the conversion factor of feedstock to biofuel.

*Accounting for Cellulosic Ethanol from Corn Stover and Wood Chips*

Biomass sources that do not compete directly for acres with high-value crops, such as corn and soybeans, would not have large implicit land costs. Since corn stover and woody biomass do not compete for crop acreage, it seems reasonable to assume the RFS for cellulosic ethanol of 16 billion gallons per year will only be met with a contribution of feedstock from these sources — if more land-intensive biomass like switchgrass is profitable, stover and woody biomass will be profitable also. Because this production occurs outside the framework of our model, we need to make assumptions about how much ethanol will be produced from these sources. We assume ethanol from both corn stover and woody biomass is produced when the economy produces a nonzero amount of switchgrass ethanol. It remains unclear exactly how much cellulosic ethanol will come from the sources not competing for land with traditional crops, so we present several scenarios varying the amount of cellulosic ethanol that must come from switchgrass to meet the mandate.

*Calculating Returns to Biofuel Production*

Returns to biofuel production are affected most by feedstock costs and governmental policy, with feedstock costs determined endogenously in the model. Ethanol and cellulosic ethanol plants use corn and switchgrass as feedstock, but biodiesel uses soy oil (not soybeans directly) as feedstock. Our model produces equilibrium soybean prices but not soy oil prices.

We estimate a simple linear relationship between the price of soybeans and the price of soy oil using recent data:[12]

$$\textit{Soy Oil Price} = 0.044\ \textit{Soybean Price} - .009 \quad R^2 = 0.878$$

Each biofuel production process generates a co-product, which creates value and offsets some feedstock cost. Corn ethanol produces dried distillers grains, dried distillers grains with solubles (DDGS), or wet distillers grains, which are used in beef, pork, and poultry rations. Distillers grains have approximately the same digestible energy content as corn, so we give credit to corn ethanol plants for DDGS consistent with its ability to substitute for corn (Shurson et al., 2003). The biodiesel production process yields glycerin, fatty acids, and filter cakes. We credit 8¢ per gallon to the biodiesel producer based on recent market value for these co-products (Paulson and Ginder, 2007).

Production of ethanol from switchgrass produces lignin, which is combustible and used to generate electricity within the facility, or sold back to the electrical grid (Aden et al., 2002). We credit switchgrass ethanol with 10¢ per gallon as suggested in Aden et al. The per gallon non-feedstock costs of producing corn-based ethanol and cellulosic ethanol are 76¢ per gallon and 97¢ per gallon, respectively, while the non-feedstock cost of producing biodiesel is 55¢ per gallon (Paulson and Ginder, 2007; Tokgoz et al., 2007).

A biofuel plant's revenue relates directly to crude oil prices through the relationship between crude oil, ethanol, and diesel. We estimate the price of ethanol and diesel as deterministic linear functions of the price of crude oil, using monthly spot prices from

[12] We estimated the relationship from the daily nearest cash prices on the CBOT from October 17, 2005, to September 14, 2007.

January 1994 through August 2007 of the Cushing Oklahoma crude oil, New York Harbor conventional gasoline, and U.S. No. 2 wholesale/resale markets: [13]

$$\textit{Wholesale Gasoline Price} = 0.21 + 2.84\textit{Crude Oil Price} \qquad R^2 = 0.97$$

$$\textit{Wholesale Diesel Price} = -4.00 + 3.13\textit{Crude Oil Price} \qquad R^2 = 0.98$$

E10 (10% ethanol, 90% gasoline) is used for its ability to oxygenate gasoline, which enhances combustion and reduces emissions (NSTC, 1997). Following Tokgoz et al. (2007), we assume, based on the demand-side model, that when annual production is greater than 15 billion gallons per year, the E10 market becomes saturated, causing ethanol to be priced at the margin according to its energy value (about two-thirds) compared to gasoline (Shapouri et al., 1995). When production is below this threshold, we assume ethanol is priced at a premium to gasoline and valued for its properties as an additive (Hurt et al., 2006). We interpolate between the additive and energy value to preserve a continuous pricing rule as follows:

$$P_{ethanol} = \begin{cases} 1.05 * P_{gasoline} & \text{if } \textit{ethanol production} < 14 \textit{ bil gal} \\ \left(1.05\lambda + .667(1-\lambda)\right) * P_{gasoline} & \text{if } 14 \textit{ bil gal} < \textit{ethanol production} < 16 \textit{ bil gal} \\ .667 * P_{gasoline} & \text{if } \textit{ethanol production} > 16 \textit{ bil gal} \end{cases}$$

where $\lambda = \dfrac{\textit{ethnaol production} - 14}{16 - 14}$ .

[13] Historical data maintained at http://tonto.eia.doe.gov/dnav/pet/pet_pri_spt_s1_d.htm.

**Results**

We impose biofuel production at levels set in the RFS of the EISA of 2007 and consider the bioeconomy's equilibrium outcomes for three different long-run crude oil price scenarios. After imposing the biofuel production levels, our model allows us to solve for the level of subsidy (tax credit) required to maintain the zero-excess-return condition, in addition to delivering long-run equilibrium crop price distributions and acreage allocations. For a crude oil price of $150 per barrel corn has a mean long-run equilibrium price of $7.38 per bushel; soybeans, $19.57 per bushel; and switchgrass, $203.76 per ton. Long-run equilibrium acreage allocations are 48% of acres in corn, 27% in soybeans, and 22% in switchgrass or hay. This is equivalent to 105 million acres of corn, 59 million acres of soybeans, and 48 million acres of switchgrass or hay.[14]

We compare the level of tax credit required to maintain the no-excess-return conditions across different crude oil price scenarios. Table 2 displays the results. We could not say, a priori, whether high crude oil prices imply higher or lower tax credit levels required to achieve long-run equilibrium at the mandated biofuel quantities. Crude oil prices act on biofuel returns in two ways: high crude prices imply high biofuel prices, positively affecting returns, but in addition, high crude prices have a negative effect on returns through upward pressure on commodity prices. Without simulating the model, we cannot determine if the positive or negative effect is stronger; table 1 illustrates, however, that as crude oil prices

[14] Compare this with 93 million acres, 63.6 million acres, and 61.5 million acres in corn, soybeans, and hay, respectively, in 2007.

rise, the tax credit required to maintain the production targets decreases.

Table 1: Long-run results under different tax credits, RFS mandate, and crude oil price scenarios

| | New RFS Low Crude | | New RFS Mid Crude | | New RFS High Crude | |
|---|---|---|---|---|---|---|
| $E[p_{crude}]$ ($/barrel) | $100 | | $150 | | $200 | |
| $E[p_{corn}]$ ($/bu) | $7.40 | | $7.38 | | $7.33 | |
| $E[p_{sb}]$ ($/bu) | $19.64 | | $19.57 | | $19.38 | |
| $E[p_{sg}]$ ($/ton) | $204.77 | | $203.76 | | 201.17 | |
| Land Allocations | (0.48 0.27 0.22) | | (0.48 0.27 0.22) | | (0.48 0.27 0.22) | |
| Production | | | | | | |
| Corn (mil bu) | 16,760 | | 16,760 | | 16,760 | |
| Soybeans (mil bu) | 2,908 | | 2,908 | | 2,908 | |
| Switchgrass (mil tons) | 166 | | 166 | | 166 | |
| Usage | Biofuel | Non-biofuel | Biofuel | Non-biofuel | Biofuel | Non-biofuel |
| Corn (mil bu) | 5,357 | 11,402 | 5,357 | 11,402 | 5,357 | 11,402 |
| Soybeans (mil bu) | 682 | 2226 | 682 | 2226 | 682 | 2226 |
| Switchgrass (mil tons) | -- | 166 | -- | 166 | -- | 166 |
| Corn ethanol production (million gallons) | 15,000 | | 15,000 | | 15,000 | |
| Biodiesel production[a] (million gallons) | 1,000 | | 1,000 | | 1,000 | |
| Switchgrass ethanol production (million gallons) | 0 | | 0 | | 0 | |
| Tax credit Corn ethanol ($/gal) | $0.80 | | $0 | | $0 | |
| Tax credit–biodiesel ($/gal) | $4.10 | | $2.50 | | $0.88 | |
| Tax credit–cellulosic ethanol ($/gal) | $2.04 | | $1.07 | | $0.10 | |

Land allocations, under the EISA RFS, shift toward crops whose associated fuels are mandated at a high level. Incentives provided to the greener fuels diffuse through the economy and cause a shift in land use patterns, and the increased usage of all the crops for biofuels cause higher commodity prices than in the baseline. If the cellulosic mandates are designed to avoid the feed-versus-fuel trade-off, our results suggest it will actually exacerbate the problem by inducing higher feedstuff costs than under the policy in which only corn ethanol is produced. With a fixed amount of land, it is impossible to increase the amount of *each* crop devoted to energy *and* maintain the same level of consumption of each commodity for food. With mandates reducing the price sensitivity of biofuel producers, the other users of the crops, namely food producers, adjust most.

*Sensitivity of Results to Required Levels of Switchgrass Production*

The amount of cellulosic ethanol than can be feasibly produced from corn stover and woody biomass is uncertain. This amount will be a significant factor in determining long-run commodity prices and land use patterns because the amount of cellulosic ethanol not covered by corn stover and woody biomass must be made up with land-intensive biomass crops such as switchgrass. The more land-intensive biomass needed to meet the cellulosic ethanol requirements, the greater the intensity of competition for acreage.

Table 2 presents the results of several scenarios increasing the amount of switchgrass ethanol needed to meet the new RFS. In the first scenario, we consider the case in which the new standard for cellulosic ethanol is met exclusively with corn stover and woody residue — with no land intensive biomass needed. Note that we calculate that the subsidy given to cellulosic ethanol (including corn stover and wood chip ethanol) can be as high as $1.07 per gallon before the switchgrass ethanol sector begins to expand. The final scenario requiring 750 million gallons of switchgrass ethanol per year required tax credits of \$0.46, \$5.26, and \$2.69 for corn ethanol, biodiesel and switchgrass ethanol, respectively. With increasing requirements on land-intensive switchgrass ethanol, we see higher commodity prices and higher subsidy levels needed to maintain the mandated biofuel quantities.

Table 2: Sensitivity of crop prices and required subsidy levels to increasing switchgrass ethanol levels

| | New RFS Mid Crude | | New RFS Mid Crude | | New RFS Mid Crude | | New RFS Mid Crude | |
|---|---|---|---|---|---|---|---|---|
| $E\left[p_{crude}\right]$ ($/barrel) | $150 | | $150 | | $150 | | $150 | |
| $E\left[p_{corn}\right]$ ($/bu) | $7.38 | | $9.12 | | $9.57 | | $9.66 | |
| $E\left[p_{sb}\right]$ ($/bu) | $19.57 | | $25.53 | | $27.09 | | $27.42 | |
| $E\left[p_{sg}\right]$ ($/ton) | $203.76 | | $284.84 | | $305.40 | | $309.92 | |
| Land Allocations $(\pi_1 \ \ \pi_2 \ \ \pi_3)$ | (0.48 0.27 0.22) | | (0.46 0.27 0.27) | | (0.45 0.27 0.28) | | (0.45 0.27 0.29) | |
| Production | | | | | | | | |
| Corn (mil bu) | 16,760 | | 15,846 | | 15,723 | | 15,648 | |
| Soybeans (mil bu) | 2,908 | | 2,656 | | 2,633 | | 2,613 | |
| Switchgrass (mil tons) | 166 | | 204 | | 210 | | 215 | |
| Usage | Biofuel | Non-biofuel | Biofuel | Non-biofuel | Biofuel | Non-biofuel | Biofuel | Non-biofuel |
| Corn (mil bu) | 5,357 | 11,402 | 5,357 | 10,489 | 5,357 | 10,366 | 5,357 | 10,291 |
| Soybeans (mil bu) | 682 | 2226 | 682 | 1,982 | 682 | 1,951 | 682 | 1,931 |
| Switchgrass (mil tons) | -- | 166 | 4 | 200 | 7 | 203 | 11 | 204 |
| Corn ethanol production (million gallons) | 15,000 | | 15,000 | | 15,000 | | 15,000 | |
| Biodiesel production[a] (million gallons) | 1,000 | | 1,000 | | 1,000 | | 1,000 | |
| Switchgrass ethanol production (million gallons) | **0** | | **250** | | **500** | | **750** | |
| Tax credit–corn ethanol ($/gal) | $0 | | $0.30 | | $0.41 | | $0.46 | |
| Tax credit–biodiesel ($/gal) | $2.50 | | $4.57 | | $5.10 | | $5.26 | |
| Tax credit–cellulosic ethanol ($/gal) | $1.07 | | $2.29 | | $2.59 | | $2.69 | |

**Conclusions**

Competition for land ensures that policy providing an incentive to even one type of biofuel — and indirectly to the crop it uses as feedstock — will increase the equilibrium prices of all crops. This means in order for switchgrass ethanol to be commercially viable, it must receive a differential subsidy over that awarded to corn-based ethanol. Homogeneous subsidy levels for both types of ethanol cannot entice expansion of switchgrass ethanol. Since switchgrass competes for the same acres as corn, and corn-based ethanol is less expensive to produce, corn-based ethanol will always have a comparative advantage over switchgrass ethanol in the absence of a differential subsidy.

Corn and soybeans also compete for the same acreage, so when energy prices are such that corn-based ethanol expands, the price of soybeans also increases, ensuring that a farmer allocates some land to soybeans. This increase in soybean prices reduces the profitability of biodiesel even in scenarios in which energy prices are high; and under pre-EISA subsidy levels, the soy oil biodiesel sector is not viable under any energy price considered. If the EISA mandates are to be met in a voluntary fashion, the biodiesel sector requires a higher subsidy, relative to corn-based ethanol, than it has today.

We calculate the subsidies required to stimulate biofuel production to levels mandated in the EISA RFS and find that subsidy levels need to be in the range of \$0 to \$0.80 per gallon for corn ethanol, and \$0.88 to \$4.10 per gallon for biodiesel, even assuming the cellulosic ethanol standard can be met solely with corn stover and woody biomass. Future crude oil prices largely determine the subsidy levels required to maintain industry sizes required by the EISA RFS.

Meeting the RFS causes much higher commodity prices than in the baseline, suggesting the cellulosic mandate in the EISA — which appear designed to avoid the feed-versus-fuel trade-off — may actually exacerbate the situation relative to the scenario with corn-based ethanol only. Cellulosic ethanol is more expensive to produce, and switchgrass-based ethanol is more land intensive than corn-based ethanol.

We calculate that the tax credits required to maintain industry sizes prescribed in the 2007 RFS are between \$0 to \$0.46, \$2.56 to \$5.26, and \$1.07 to \$2.69 for corn ethanol, biodiesel and switchgrass ethanol, respectively. The severity of upward pressure on commodity prices caused by the new RFS will be determined largely by the ability to produce cellulosic ethanol from biomass that is not land intensive: policies that expand cellulosic ethanol beyond levels that can be supplied by corn stover and woody biomass are more expensive in terms of the subsidy required and the resulting increase in food and feed prices.

**References**

Aden, A. , Ruth, M., Ibsen, K., Jechura, J., Neeves, K., Sheehan, J., Wallace, B., Montague, L., Slayton, A., Lukas, J. 2002. Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. NREL/TP-510-32438. National Renewable Energy Laboratory, Golden, CO.

Elobeid, A., Tokgoz, S., Hayes, D.J., Babcock, B.A., Hart, C.E. 2007. The long-run impact of corn-based ethanol on the grain, oilseed, and livestock sectors: a preliminary assessment. *AgBioForum* 10 (1), 11-18.

Hurt, C., Tyner, W., Doering, O. 2006. Economics of ethanol. (ID-339). Purdue University Cooperative Extension. http://www.ces.purdue.edu/extmedia/ID /ID-339.pdf (accessed Feb. 2008).

Lewandowski, I., Scurlock, J., Lindvall, E., Christou, M. 2003. The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. *Biomass and Bioenergy* 25, 335-361.

Magnussen, S. 2004. An algorithm for generating positively correlated beta-distributed random variables with known marginal distributions and a specified correlation. *Computational Statistics & Data Analysis* 46 (2), 397-406.

NSTC. National Science and Technology Council, Committee on Environment and Natural Resources. 1997. *Interagency assessment of oxygenated fuels.* Report to the Office of the President of the United States, Washington, DC.

Paulson, N.D., Ginder, R.G. 2007. The growth and direction of the biodiesel industry in the United States. CARD Working Paper 07-WP 448, Center for Agricultural and Rural Development, Iowa State University.

Schmer, M. R., et al. "Net energy of cellulosic ethanol from switchgrass." *Proceedings of the National Academy of Sciences* 105, no. 2(2008): 464.

Shapouri, H., Duffield, J.A., Graboski, M.S. 1995. Estimating the net energy balance of corn ethanol. Agricultural Economic Report No. 721. U.S. Department of Agriculture, Economic Research Service, Office of Energy and New Uses.

Sharpe, W.F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19 (3), 425-442.

Shurson, J., Spiehs, M., Wilson, J., Whitney, M. 2003. Value and use of 'new generation' distiller's dried grains with solubles in swine diets. *Alltech's 19th International Feed Industry Proceedings,* May 12-14, Lexington, KY.

Tokgoz, S., Elobeid, A., Fabiosa, J.F., Hayes, D.J., Babcock, B.A., Yu, T-H., Dong, F. F., Hart, C.E., Beghin, J.C. 2007. Emerging biofuels: Outlook of effects on U.S. grain, oilseed, and livestock markets. CARD Staff Report 07-SR 101, Center for Agricultural and Rural Development, Iowa State University.

**Appendix: Supplemental Table**

The following table contains a summary of the parameter values used to implement the simulation. The first two columns contain supply parameters and the remaining columns contain demand parameters for each commodity under consideration.

Table 3: Parameters used in Monte Carlo simulation

| | $\kappa_i$ | $a_i$ | $\alpha_0^i$ | $\alpha_1^i$ | $\alpha_2^i$ | $\alpha_3^i$ | $\alpha_4^i$ |
|---|---|---|---|---|---|---|---|
| Corn | 200.43 | 125.81 | 81.35 | - 0.258 | 0.002 | 0 | 0.155 |
| Soybeans | 27.70 | 85.78 | 23.99 | 0.081 | - 0.379 | 0 | 0.538 |
| Switchgrass | 9.94 | 46.43 | 2.04 | 0 | 0 | - 0.16 | 0.058 |

# CHAPTER 3: WELFARE CHANGES FROM INCREASED BIOFUEL PRODUCTION ON U.S. AGRICULTURE: THE ROLE OF UNCERTAINTY AND INTERLINKED COMMODITY MARKETS

## Introduction

Ethanol production increased more than five hundred percent since the turn of the century[15]. Stimulated by a combination of tax incentives, a renewable fuel standard[16], and strong energy prices, the industry has experienced seemingly exponential growth (see figure 1). Ethanol currently is produced almost exclusively from corn, meaning that each gallon of ethanol produced diverts some corn away from its alternative uses, which are food for humans and feed for livestock. Increased ethanol production amounts to a shift outward in the demand for corn, and it is important to understand the distributional impact of this shift on other agricultural sectors. In this paper we hope to illuminate some pertinent issues in quantifying these welfare effects and contribute to the discussion that has evolved regarding their nature and magnitude.

---

[15] According the Renewable Fuels Association industry statistics http://www.ethanolrfa.org .

[16] The American Jobs Creation act of 2004 (H.R. 4520) provided a $0.51 per gallon Volumetric Ethanol Excise Tax Credit to blender's ethanol which was reduced to $0.45 per gallon as of January 1, 2009, and the Energy Security and Independence Act of 2007 (H.R.6) gradually increases the Renewable Fuel Standard to 36 billion gallons per year by 2022.

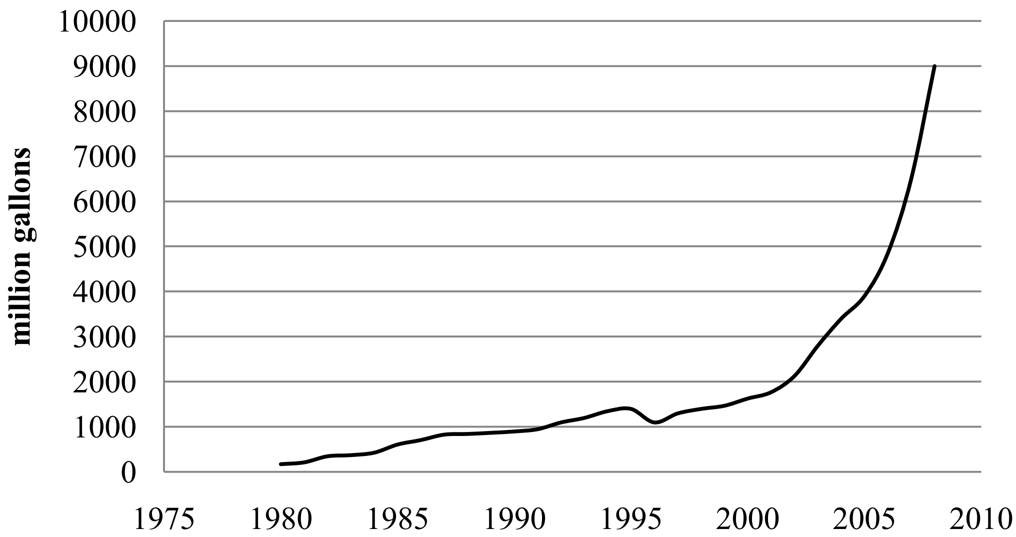



Figure 1: Annual Ethanol Production in the United States
Source: Renewable Fuels Association

A number of researchers have approached the problem using different modeling assumptions. There are several features present in agricultural markets that, as a set, are overlooked or generalized in previous studies. These include the valuation of farm deficiency programs as options, a farmer's ability to grow alternative crops, and demand disaggregated to understand the relative size of welfare loss in different consumption sectors.

Government spending shifts from price deficiency programs to biofuel tax credits. The shift could theoretically reduce taxpayers' liability if increased biofuel production raises agricultural commodity prices to a level such that there is saving in farm program payments larger than expenditure on the tax credit. However, in order to include the change in farm program liabilities in a welfare analysis, price uncertainty should be included in the model. Measuring a change in taxpayers' liability using a deterministic model does not allow one to place a value on the option-like nature of these programs. Therefore, we include uncertainty in crop yields, which translates into uncertain crop prices.

We model the farmer's choice to allocate acres among many crops. If two or more crops can be planted on the same acreage, the supply functions must include the price expectations of all crops under consideration. This means biofuel production will affect the corn market as well as related crop markets. These indirect impacts should be included in a welfare analysis of increased biofuel production. The crops we will consider are corn, soybeans and wheat.

Using an aggregate demand function does not allow one to quantify the relative size of welfare gain or loss in different sectors. A welfare analysis is used to describe who is benefitting and who is losing due to a particular phenomena. Disaggregated demands allow for a more rich discussion of the welfare transfers involved. We model demand for corn, soybeans and wheat from the livestock feed, human food, and export sectors.

We calculate welfare effects based on the following scenarios:

- Baseline:

  Biofuel production and crop yields are set at 2004 levels (3.4 billion gallons of ethanol and 25 million gallons of biodiesel) [17].

- +3 Scenario:

[17] Annual ethanol production estimates from http://www.ethanolrfa.org/industry/statistics/, and annual biodiesel production estimates from http://www.biodiesel.org/pdf_files/fuelfactsheets/Production_graph_slide.pdf.

Biofuel production levels increased by 3 billion gallons of ethanol and 150 million gallons of biodiesel.

- + 6 Scenario:

  Biofuel production levels increased by 6 billion gallons of ethanol and 300 million gallons of biodiesel.

- + 9 Scenario:

  Biofuel production levels increased by 9 billion gallons of ethanol and 450 million gallons of biodiesel.

Using 2004 as the baseline scenario gives us a point of reference before ethanol and biodiesel production levels seemed to make a sizeable impact in the agricultural markets. In the scenarios, we increase biofuel production levels and calculate the associated welfare changes.

The remainder of article proceeds as follows. In the next section we describe the economic model used to conduct the welfare analysis in more detail; a section describes the way in which we calculate welfare changes. A section describing the model's results comes next, followed by a section discussing our model's sensitivity to some parameter assumptions. We provide a comparison of our findings with other recent studies in the literature, and the last section concludes.

**The Model Economy**

The model economy has three goods: corn, soybeans and wheat. Consumers buy agricultural commodities and use them as an input in producing livestock feed, human food, energy, or for export. We introduce uncertainty in the model through agricultural commodity yields. In the first period, agents form expectations about future prices and crop yields, and farmers allocate acreage among corn, soybeans and wheat. In the second period, crop supply is determined by acreage allocations and crop yield realizations.

Period 1

- Expectations formed about prices and crop yields
- Farmers allocate acres
- Demand functions for crops determined

Period 2

- Crop yield and therefore supply is realized
- Equilibrium prices emerge
- Markets clear

*Commodity Supply*

There exists a single representative and competitive farmer with an endowment of land, who takes both output prices and his cost function as given. While output prices and yields are uncertain, the farmer and consumers know the joint distribution. The crops are indexed as follows: corn, $i$ = 1; soybeans, $i$ = 2; and wheat, $i$ = 3. In order to incorporate an acreage response in the model we assume the farmer can increase the quantity of available land, but that it is costly to do so. He chooses the optimal level of land according to the problem

$$\max_{L} p^I L - c_{cl}(L) \ ,$$

where $L$ is the quantity of land, and $c_{cl}(\cdot)$ can be thought of as the cost of preparing the land to bring into production, and $p^I$ is an index of the prices $p^1$, $p^2$, $p^3$, which are the prices of corn soybeans and wheat respectively. The index is defined by

$$p^I = \frac{p^1\bar{\pi}_1 + p^2\bar{\pi}_2 + p^3\bar{\pi}_3}{p^1_{04}\bar{\pi}_1 + p^2_{04}\bar{\pi}_2 + p^3_{04}\bar{\pi}_3},$$

where $p_{04}^{i}$ is the price of commodity $i$ in 2004 and $\bar{\pi}_i$ is the proportion of land allocated to crop $i$ in 2004. To be consistent with our choice of baseline scenario we use the price level in 2004 as the reference point of the price index.

We assume the optimal acreage function $L^*\left(p^I\right)$ is of the form $L^*\left(p^I\right)=\bar{A}*\left(p^I\right)^{\eta}$ where $\bar{A}$ represents acres from the baseline year, and $\eta$ is the elasticity measuring aggregate acreage's responsiveness to the price level. The actual harvested acreage for 2004 and 2007 implies a value of $\eta = 0.03$, 2004 to 2008 implies a value of $\eta = 0.10$, and 2004 to 2009 implies a value of $\eta = 0.06$. For our analysis we use $\eta = 0.03$, since this seems to be more reasonable as a long run acreage response than the levels we saw in 2008 and 2009. Table 1 contains the actual planted and harvested acres from 2004 to 2009.

Table 1: Actual Planted and Harvested Acreage from 2004 – 2007 (Sum of Corn, Soybeans and Wheat)

| | PA | % Change from 2004 | HA | % Change from 2004 |
|---|---|---|---|---|
| 2004 | 215,781 | | 197,558 | |
| 2005 | 211,025 | -0.02 | 196,472 | -0.01 |
| 2006 | 211,183 | -0.02 | 192,040 | -0.03 |
| 2007 | 218,728 | 0.01 | 201,665 | 0.02 |
| 2008 | 224,847 | 0.04 | 208,966 | 0.06 |
| 2009 | 219,648 | 0.02 | 207,109* | 0.05 |

Soucre: NASS
Units: 1,000 acres
*Projected

We do not account for spatial heterogeneity in the productivity of land and therefore the total acreage decision can be made independently from the acreage allocation decision.

The farmer allocates his land in period one to three different crops: corn, soybeans, and wheat. The endowed land is representative of total U.S. cropland devoted to these commodities. Cellulosic ethanol was not commercially viable in the period of 2004 and thus is not included in this analysis[18].

The producer's per acre profit is given by

(11) $$w = \sum_{i=1}^{3} p_i \cdot \zeta_i \cdot \pi_i - c_i\left(\pi_i ; \Theta_i\right)$$

where $p_i$ is crop $i$'s output price, $\zeta_i$ is the realized per acre yield of crop $i$, and $\pi_i$ is the proportion of crop land allocated to crop $i$. The cost function for crop $i$ is $c_i\left(\pi_i; \Theta_i\right)$, where $\Theta_i$ is a vector of parameters defining each crop's cost function. Aggregate (national) profit is then calculated by $L^* w$.

The producer is risk neutral in profit; he wishes to maximize his expected profit subject to land constraints. To this end, he chooses a land allocation vector, $\left[\pi_1 \quad \pi_2 \quad \pi_3\right]'$, to solve the problem:

(12) $$\max_{\pi_1,\pi_2,\pi_3 \geq 0} E\left[w\right] \quad s.t. \ \sum_{i=1}^{3} \pi_i \leq 1$$

If we denote for crop $i$ the expected price by $\bar{p}_i$, the expected yield for crop $i$ by $\mu_i$, and the shadow value of land by $\lambda^*$, the Kuhn-Tucker conditions for this problem are:

$$\bar{p}_i \mu_i - \frac{\partial c_i\left(\pi_i^*; \Theta_i\right)}{\partial \pi_i} - \lambda^* \leq 0, \quad \pi_i^* \geq 0, \quad \pi_i^*\left[\bar{p}_i \mu_i - \frac{\partial c_i\left(\pi_i^*; \Theta_i\right)}{\partial \pi_i} - \lambda^*\right] = 0 \quad \text{for } i = 1, 2, 3$$

---

[18] A report on the status (as of February 2007) of biofuel production in the U.S. written by the Energy Information Administration can be found at: http://www.eia.doe.gov/oiaf/analysispaper/biomass.html.

and

$1-\pi_1^*-\pi_2^*-\pi_3^*\le 0,\quad \lambda^*\ge 0,\quad \lambda^*\left[1-\pi_1^*-\pi_2^*-\pi_3^*\right]=0$.

Assuming an interior solution, the first order conditions are:

(13) $\bar{p}_i\mu_i-\dfrac{\partial c_i}{\partial \pi_i}=0$ for $i$ = 1, 2, 3

The farmer's price expectations above are assumed to be such that, combined with demand, cause the ex-ante (but post planting) price distribution to have a mean of $\bar{\mathbf{p}}=\left[\bar{p}_1 \quad \bar{p}_2 \quad \bar{p}_3\right]'$.

The optimal acreage decisions combined with yield realizations, $\zeta_i$, give the supply function for each crop:

(14) $S^i\left(\bar{\mathbf{p}};\boldsymbol{\mu},\Theta,\zeta_i\right)=\zeta_i\;\pi_i^*\left(\bar{\mathbf{p}};\boldsymbol{\mu},\boldsymbol{\Theta}\right)L^*\left(\bar{p}^I\right)$

Notice that both the expected output price and production cost of the other crops, enter each crop's supply function.

*Specification of Cost Function and Crop Yield Distributions*

Each crop's cost function is assumed to be quadratic; i.e., $c_i\left(\pi_i\right)=a_i\,\pi_i+\kappa_i\left(\pi_i\right)^2$ for $i=1,\,2,\,3$. The proportion of land allocated to each crop, $\pi_i$, is the farmer's choice variable and $a_i$ and $\kappa_i$ are parameters. Notice that the same solution would prevail if the cost function were re-parameterized in the more traditional way as a function of expected output, $c_i\left(q_i\right)=\tilde{a}_i\,q_i+\tilde{\kappa}_i\left(q_i\right)^2$ for $i=1,\,2,\,3$ where $q_i=\mu_i\pi_i$. Parameterizing our model as a function of land allocation allows for the more intuitive interpretation of the cost function as variable cost per acre.[19] The cost function parameters are calibrated such that the expected price levels of the crops in 2004 brings about the land allocations we saw in 2004 and can be found in table 2.

Table 2: Parameters Used in Monte Carlo Simulation

| | | Corn | Soybeans | Wheat |
|---|---|---|---|---|
| Demand | $\alpha_0^i$ | 8.815 | 640.132 | -- |
| | $\alpha_1^i$ | -0.550* | -0.379 | -- |
| | $\alpha_2^i$ | 0.002 | 0.081 | -- |
| | $\beta_0^i$ | 1.645 | 16.922 | 1.216 |
| | $\beta_1^i$ | -0.059 | -0.150 | -0.050 |
| | $\beta_2^i$ | 0.002 | 0.003 | 0.004 |
| | $\gamma_0^i$ | 3.228 | 3.462 | 2.378 |
| | $\gamma_1^i$ | -0.840 | -0.840* | -0.670* |
| | $\delta_i$ | $\left(1-17/56\right)/2.8$ | 7.7/11 | -- |
| Supply | $a_i$ | -68.25 | 27.95 | 190.31 |
| | $\kappa_i$ | 2388.50 | 1293.35 | 1415.28 |
| | $\eta$ | 0.03 | | |

Units of demand equations are billion bushels
Demand elasticities are from the ERS trade model except where denoted by an *.
These values were calibrated to ensure price volatilities in the baseline that were typical of price volatility in 2004.

A cost function that displays an increasing marginal cost of production is necessary in

[19] This is also closer to the way farmers actually think about their allocation problem. They typically think about per acre expected revenue and costs when making planting decisions each year, not per bushel revenue and cost.

this model because if faced with constant marginal cost, a price-taking farmer would choose to allocate all land to one crop – the crop yielding the highest expected profit. Using a cost function characterized by increasing marginal cost in land allocation is one way to ensure the farmer plants a mix of crops in the model. This is appropriate because we observe farmers planting a mix of crops in reality.

The yield realizations, $\zeta$, are drawn from the joint beta distribution of yields using the algorithm developed by Magnussen (2004).

$$\zeta \sim \beta\left(\begin{bmatrix}\mu_1\\ \mu_2\\ \mu_3\end{bmatrix}, \Sigma^{-1}, \mathbf{q}_{max}, \mathbf{q}_{min}\right)$$

$$\mathbf{q}^{t}_{\max} = \begin{bmatrix}\mu_1 + 2\sigma_1\\ \mu_2 + 2\sigma_2\\ \mu_3 + 2\sigma_3\end{bmatrix}$$

$$\mathbf{q}^{t}_{\min} = \begin{bmatrix}\mu_1 - 3\sigma_1\\ \mu_2 - 3\sigma_2\\ \mu_3 - 3\sigma_3\end{bmatrix}$$

Where $\boldsymbol{\mu} = \begin{bmatrix}\mu_1 & \mu_2 & \mu_2\end{bmatrix}'$ is 2004 trend yield.[20] The matrix $\Sigma^{-1}$ is the variance-covariance

[20] We assume yields follow a linear trend. We estimate the trend from yield data (per harvested acre) for the years 1980 through 2006. Available at http://www.nass.usda.gov/ The estimated yield trends used in the model are : (note that $t$ is the actual year)

$\mu^{t}_{corn} = -3843.83 + 1.990t$, $\mu^{t}_{soybean} = -993.52 + .516t$, $\mu^{t}_{wheat} = -460.50 + 0.250t$. So the 2004 trend yields are $\boldsymbol{\mu}^{2004} = \begin{bmatrix}144.98 & 42.77 & 40.50\end{bmatrix}'$

matrix for yields per harvested acre of the three crops, and σ are the standard deviations of each crop found in $\Sigma^{-1}$. We set the maximum and minimum values of the draws from the beta distribution to be two and three standard deviations away from the mean respectively.

*Commodity Demand*

Demand for each commodity comes from four different subsectors, livestock feed (*l*), human food consumption (*f*), exports (*x*), and energy (*e*). All demands are constant elasticity functions of prices. Table 2 contains parameter estimates of the demand equations used in the Monte Carlo simulation. Most of the price elasticities are taken from the ERS/Penn State trade model[21] for livestock feed and food/consumer demand. We use one of the demand elasticities of each crop to calibrate the model's price volatility to levels typically found in the market in the baseline year. Having realistic price volatility in our baseline scenario is important since we are going to estimate an ex-ante value of the loan deficiency and countercyclical payments. The constant terms are calibrated to 2004 consumption levels, and we provide a sensitivity analysis on some of these parameter values later in the article.

*Corn Demand Equations*

More specifically, the corn demand equations are given by:

$$D_l^1\left(p_1, p_2^{sbm}; \Omega_1\right) = \alpha_0^1\left(p_1\right)^{\alpha_1^1}\left(p_2^{sbm}\right)^{\alpha_2^1} \tag{15}$$

[21] The ERS/Penn State trade model documentation can be found at http://trade.aers.psu.edu/pdf/ERS_Penn_State_Trade_Model_Documentation.pdf.

$$D_f^1\left(p_1, p_2^{oil}; \Omega_1\right) = \beta_0^1 \left(p_1\right)^{\beta_1^1} \left(p_2^{oil}\right)^{\beta_2^1}$$

$$D_x^1\left(p_1; \Omega_1\right) = \gamma_0^1 \left(p_1\right)^{\gamma_1^1}$$

$$D_e^1 = \delta_1 \, n_1$$

$$D^1\left(\mathbf{p};\ \Omega_1\right) = D_e^1 + D_x^1 + D_f^1 + D_l^1$$

where $p_1$ is the price of corn, $p_2^{sbm}$ is the price of soybean meal, and $p_2^{oil}$ is the price of soybean oil (in the description of soybean demand we describe the way soybean oil and soybean meal prices relate to soybean prices in the model). The size of the corn-based ethanol industry in billion gallons per year is $n_1$. Therefore, $\delta_1$ transforms the amount of ethanol produced into the amount of corn required. In the ethanol production process distiller's grains are produced as a co-product with ethanol. This substance is valuable as livestock feed and thus is added back in the model on a corn equivalent basis (Shurson, et al., 2003). Ethanol has a yield of 2.8 gallons per bushel of corn and for every bushel (56 lbs) of corn 17lbs of distillers grains are produced; therefore $\delta_1 = \left(1 - {}^{17}\!/_{56}\right)/2.8$ is the amount of corn (net of distiller's grains) required to produce $n_1$ billion gallons of ethanol (Shapouri and Gallagher, 2005). We denote the parameters of the corn demand function (elasticities and constant terms) by $\Omega_1$.

*The Soybean Complex*

Soybeans typically are not consumed in their natural state but instead are processed (crushed) into soybean meal and soybean oil. The meal is used as animal feed and the oil is used mostly for human consumption and biofuel production. Therefore, soybean demand is comprised of crush demand and exports. For simplicity we assume a constant elasticity of demand for all subsectors of the soybean market. Denoting the parameters of the soybean demand functions by $\Omega_2$, soybean meal demand is given by equation 6 in bushels of soybeans required, and the amount of soybeans required to meet demand for soybean oil is given by equation 7, with both equations including a term for the cross price effect of corn.

$$(16)\qquad D^2_{sb-meal}\left(p_2^{sbm}, p_1; \Omega_2\right) = \frac{1}{47}\alpha_0^2\left(p_2^{sbm}\right)^{\alpha_1^2}\left(p_1\right)^{\alpha_2^2}$$

$$(17)\qquad D^2_{sb-oil}\left(p_2^{sboil}, p_1; \Omega_2, n_2\right) = \frac{1}{11}\beta_0^2\left(p_2^{sboil}\right)^{\beta_1^2}\left(p_1\right)^{\beta_2^2} + \delta_2\, n_2$$

One bushel of soybeans yields 47 lbs of meal and 11 lbs of oil when crushed[22]. Therefore, including the constants (1/47) and (1/11) in equations (6) and (7) means, for example, that we calibrate the units of $\alpha_0^2\left(p_{sbmeal}\right)^{\alpha_2^2}\left(p_1\right)^{\alpha_2^2}$ to be in lbs of meal and the (1/47) converts pounds of meal into a soybean bushel equivalent. The size of the biodiesel industry in billion gallons per year is $n_2$ and $\delta_2$ is the conversion factor that calculates the bushels of soybeans needed to produce enough soybean oil to produce $n_2$ gallons of biodiesel; so, the expression $\delta_2\, n_2$ in equation (7) is the amount of soybean oil required to meet the biodiesel production level. Soybean oil weighs 7.3 pounds per gallon and one pound of soybean oil can produce 0.973 pounds of biodiesel, so 7.5 pounds of soybean oil are required to produce one gallon of biodiesel (AltIn, et al., 2001). Converting to a soybean bushel equivalent we have $\delta_2 = 7.5/11$.

[22] Rounded to the nearest lb/bu for 2007 processing yields as reported by the National Oilseed Processors Association. http://www.nopa.org/content/stats/stats.html

Soybean crush demand is determined by the amount of soybeans required to satisfy the demand for meal and oil and equations 8 and 9 define soybean crush demand.

(18) $D^2_{sb-oil}\left(p_2^{sboil}, p_1; \Omega_2, n_2\right) = D^2_{sb-meal}\left(p_2^{sbm}, p_1; \Omega_2\right)$

(19) $D^2_{crush}\left(p_2^{sboil}, p_2^{sbm}, p_1; \Omega_2, n_2\right) = D^2_{sb-oil}$

Since soybean meal and soybean oil are produced in fixed proportions in the crush process, the amount of soybeans required to meet soybean meal demand must equal the amount of soybeans required to meet soybean oil demand in equilibrium. This is because we assume no storage or exports of soybean meal or oil.

Crush demand depends upon the price of soy oil and soybean meal, not the price of soybeans directly. However, the prices of these link tightly to the price of soybeans, and for simplicity, we estimate a simple deterministic linear relationship of each with the price of soybeans using recent data.[23] This allows the demand for soybean meal and soybean oil to be written as a function of soybean prices.

Soybean export demand is written in equation 10 and the total demand for soybeans is the sum of soybean export and crush demand (equation 11).

(20) $D_x^2\left(p_2; \Omega_2\right) = \gamma_0^2\left(p_2\right)^{\gamma_1^2}$

[23] We estimate the relationship between the price of soy oil and soybeans from daily nearest contract prices on the CBOT from Oct. 17, 2005 to Sept. 14, 2007. For soybean meal we use daily nearest contract prices of soybean meal and soybeans on the CBOT from April 30, 2007 to March 3, 2008. $\textit{Soy Oil Price}^t = 0.044\ p_{sb}^t - 0.009$, $R^2 = 0.878$. $\textit{Soybean Meal Price}^t = 33.53 + 0.23\left(p_{sb}^t \cdot 100\right)$, $R^2 = 0.929$.

(21) $$D^2\left(p_2, p_1; \Omega_2, n_2\right) = D^2_{crush} + D^2_x$$

*Wheat Demand Equations*

Only demand from the food and exports sectors is included in the wheat demand functions, because only a small amount is used for livestock feeding and none is used for biofuel production. Denoting the parameters of the wheat demand functions by $\Omega_3$, the wheat demand equations are given by:

$$D^3_f\left(p_3, p_1; \Omega_3\right) = \beta^3_0\left(p_3\right)^{\beta^3_1}\left(p_1\right)^{\beta^3_2}$$

$$D^3_x\left(p_3; \Omega_3\right) = \gamma^3_0\left(p_3\right)^{\gamma^3_1}$$

$$D^3\left(p_3, p_1; \Omega_3\right) = D^3_f + D^3_x$$

*Competitive Equilibrium*

In our economy, a competitive equilibrium is defined by

pricing functions $p_i\left(\boldsymbol{\zeta}, \boldsymbol{\Omega}, \boldsymbol{\Theta}\right)$ for $i = 1,\ 2,\ 3$,

crop demand functions $D^i\left(\mathbf{p}, \Omega_i\right)$ for $i = 1,\ 2,\ 3$,

crop supply functions $S^i\left(\mathbf{p}; \boldsymbol{\mu}, \boldsymbol{\Theta}, \zeta_i\right)$ for $i = 1,\ 2,\ 3$,

where the set of demand parameters is defined by $\boldsymbol{\Omega}$; i.e., $\boldsymbol{\Omega} = \left\{\Omega_1, \Omega_2, \Omega_3\right\}$, so that given the pricing functions, biofuel capacity, and crop yield realizations, the commodity markets clear. That is, $S^i\left(\mathbf{p}; \boldsymbol{\mu}, \boldsymbol{\Theta}, \zeta_i\right) = D^i\left(\mathbf{p}, \Omega_i\right)$ for $i = 1, 2, 3$ .

**Calculating Welfare Changes**

This modeling framework allows the calculation of welfare changes of five different groups of stakeholders: farmers, livestock producers, processors of cereal grains and oilseeds for food use, gasoline consumers or blenders, and taxpayers. Equilibrium price and quantity outcomes define the change in welfare resulting from the increase in biofuel production levels. We detail how the change in welfare is calculated in the subsections below.

*Farmer Surplus*

The farmer's profit function, $w$, is given in equation 1, so the change in farmer surplus is the change in profit resulting from a shift from the baseline scenario to the tax credit scenario, or

(22) $$\Delta PS_{farmer} = \Delta w = w_{scenario} - w_{baseline}$$

*Consumer surplus*

Consumer surplus is calculated as the path integral from $\left[p_1^0, p_2^0, p_3^0\right]'$, the equilibrium prices in the baseline, to $\left[p_1^1, p_2^1, p_3^1\right]'$, the equilibrium prices in the scenario (Bullock, 1993, Larson, et al., 2002):

(23) $$\Delta CS = \int_{p_2^0}^{p_2^1} D^2\left(p_1^0, p_2, p_3^0, \Omega_2\right)dp_2 + \int_{p_3^0}^{p_3^1} D^3\left(p_1^0, p_2^1, p_3, \Omega_3\right)dp_3 + \int_{p_1^0}^{p_1^1} D^1\left(p_1, p_2^1, p_3^1, \Omega_1\right)dp_1$$

So that the particular path on which we integrate from $\left[p_1^0, p_2^0, p_3^0\right]'$ to $\left[p_1^1, p_2^1, p_3^1\right]'$ is the line from $\left[p_1^0, p_2^0, p_3^0\right]'$ to $\left[p_1^0, p_2^1, p_3^0\right]'$, $\left[p_1^0, p_2^1, p_3^0\right]'$ to $\left[p_1^0, p_2^1, p_3^1\right]'$, and from $\left[p_1^0, p_2^1, p_3^1\right]'$ to $\left[p_1^1, p_2^1, p_3^1\right]'$.

*Taxpayer Costs*

Taxpayers have two potential expenditures: the blender's credit and deficiency payments to farmers in the form of loan deficiency payments (LDPs) and countercyclical payments (CCPs). Since the scenarios we consider define specific amounts of biofuel to be produced in each case, the change in the taxpayer cost from the blender's credit due to the increase in biofuel production is a straightforward calculation. The change in the taxpayer's liability in the price deficiency programs is not as straightforward; we describe in some detail how we calculate this change below.

LDPs are made when the market price of the commodity is below the loan rate (see table 3 for a summary of relevant information on the farm LDP and CCP programs)[24]. Taxpayers make a payment to the farmer equivalent to the difference between the loan rate and the posted county price times the amount of commodity owned on the date, *T*, the farmer chooses. The LDP is equal to

[24] Information about the loan deficiency payment program is available at http://www.ers.usda.gov/Briefing/FarmPolicy/malp.htm; information about the countercyclical payment program is available at http://www.ers.usda.gov/Briefing/FarmPolicy/CounterCyclicalPay.htm; information about the direct payment program is available at http://www.ers.usda.gov/BRIEFING/FarmPolicy/directpayments.htm.

(24) $LDP_i^T\left(pcp_i^T, LR_i, T\right) = \mathbf{max}\left(LR_i - pcp_i^T, 0\right),$

where $LR_i$ is the loan rate for commodity *i* and $pcp_i^T$ is the posted county price for crop *i* at time *T*.

Table 3: Farm Program Details

| | Corn | Soybeans | Wheat |
|---|---|---|---|
| *adjustment* | \$0.08 | \$0.05 | \$0.13 |
| Base acres (million) | 73.83 | 52.79 | 75.53 |
| Payment Yield (bu) | 114 | 34.10 | 36.10 |
| Target price | \$2.63 | \$5.80 | \$3.92 |
| Loan Rate | \$1.95 | \$5.00 | \$2.75 |
| Direct Payment Rate | \$0.28 | \$0.44 | \$0.52 |

Countercyclical payments are similar to loan deficiency payments; however, they are determined not by the current season's production but by historical production. Further, they are not based on the harvest time price, but on the season average price or the loan rate. Payments are made when the target price is less than the effective price, which is defined by:

(25) $Effective\ Price_i = Direct\ Payment\ Rate_i + \mathbf{max}\left(p_i^{SA}, LR_i\right).$

Here the direct payment rate is a fixed quantity set by Congress and $p_i^{SA}$ is the season average price of crop *i*. At time *T* the *CCP* made to the farmer is,

$$CCP_i^T\left(p_i^{SA}, LR_i, Target\ Price_i,\ Direct\ Payment\ Rate_i,\ T\right) = \mathbf{max}\left[\left(Target\ Price_i - EffectivePrice_i\right),\ 0\right] \times 0.85 \times BA_i \times HY_i \tag{26}$$

where $BA_i$ is the base acreage for crop $i$, and $HY_i$ is the historical yield of crop $i$. See figure 2 for an example of the time $T$ net payoffs from the *LDP* and *CCP* programs.

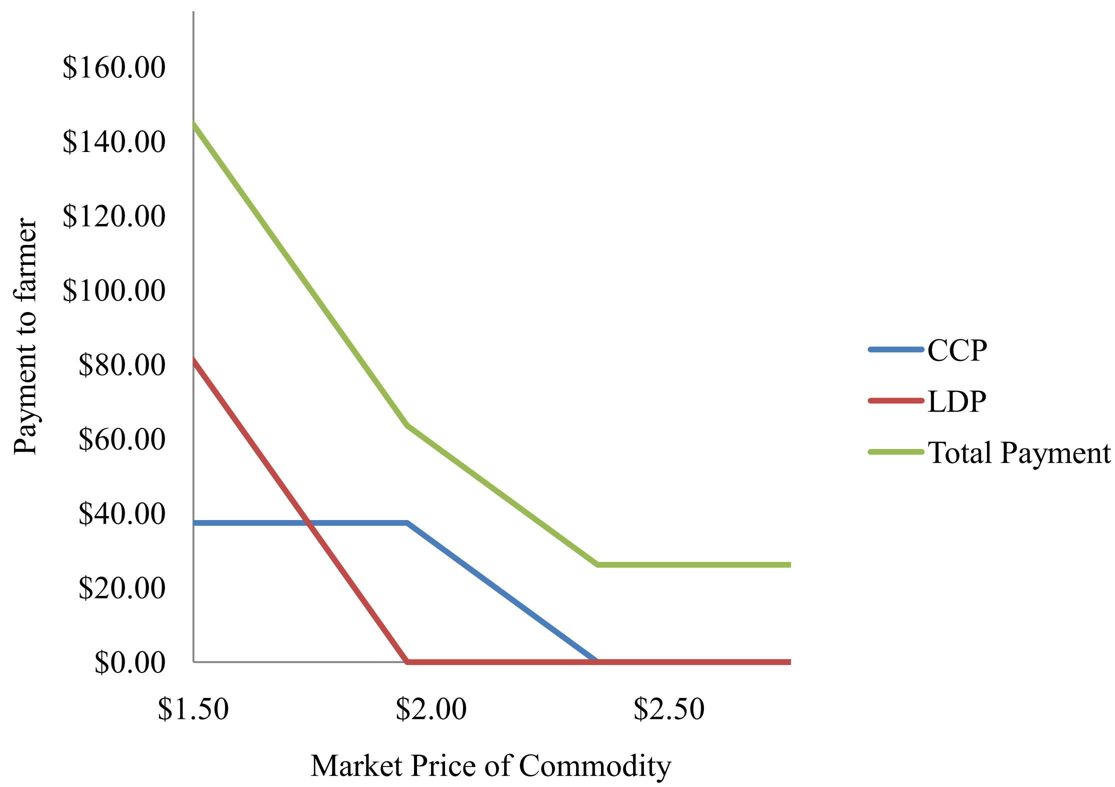



Figure 2: Deficiency payments to farmer as a function of market prices

Note: Corn market is depicted with a $1.95 loan rate, $2.63 target trice, and $0.28 direct payment. The CCP is made on 85% of base acres, and LDP is made on actual production. To illustrate, this is the per acre payment of the three programs and we assume the payment is based on 110 bu/acre historical yield and 170 bu/acre realized yield.

Looking at equation (14) one can see that the payoff of the *LDP* is of the same form as a put option with the $LR_i$ as the strike price (Gardner, 1977, Hull, 2002, Marcus and Modest, 1986). The counter cyclical payment has the same payoff as an Asian option when the season average price is higher than the loan rate, but the payment is capped by the amount, $MAXCCP_i = Target\ Price_i - Direct\ Payment\ Rate_i\ - LR_i$, when the season average price is below the loan rate.

Because of these option-like characteristics of the price deficiency programs, welfare studies that quantify the change in taxpayers' obligation should value the liabilities of these programs in a similar way as financial options with price uncertainty. The value of the LDP 'option' at time $t < T$ that is written by the taxpayer and owned by the farmer has a value of

$$LDP_i^t\left(p_i^t, LR_i, T\right) = e^{r(T-t)} E\left[\mathbf{max}\left(LR_i - p_i^T, 0\right)\right] \tag{27}$$

for crop *i*, where *r* is the risk free interest rate (assumed constant)[25], and *T* is the payment date of the LDP.

The farmer chooses the date on which his *LDP* payment will be calculated. Since our model does not have a mechanism for the evolution of price changes through the marketing year, the *LDP* payment in our analysis is calculated with respect to harvest time price.

Similarly, the value of the CCP 'option' at time $t < T$ that is written by the taxpayer and owned by the farmer has a value of,

---

[25] The Fed Funds rate was approximately 1.5% throughout most of 2004, we use this value to discount the LDP and CCP options.

http://www.newyorkfed.org/markets/omo/dmm/fedfundsdata.cfm

$$(28)\quad CCP_i^t\left(p_i^{SA}, LR_i, Target\ Price_i\,,\ Direct\ Payment\ Rate_i, T\right) \\ = e^{r(T-t)}E\Big[\mathbf{max}\Big(Target\ Price_i\ \text{-}\ Effecitve_i\,,\ 0\Big)\Big]\times 0.85\times BA_i\ \times HY_i$$

Since we do not have a way to calculate season average prices directly within the model we cannot fully account for the Asian option characteristics of the CCP payment. Asian options are less valuable than their European or American counterparts because the distribution of average prices have lower variance than the distribution of price levels (Hull, 2001). We recognize this and acknowledge that as a result our estimates of the option value of the CCP will be overstated.

We adjust the harvest-time price by the average of the difference between the season average price and the harvest-time price for each crop from 1960 through 2007 to adjust for the difference in the harvest time price level from the season average price level.

$$(29)\quad adj_i = \sum_{t=1960}^{2007} p_{it}^{SA} - p_{it}^{harvest}$$

which is \$0.08, \$0.05, and \$0.13 for corn soybeans and wheat respectively. *CCP* payments are therefore calculated using an approximated season average price defined by

$p_i^{SA} = p_i^{harvest} + adj_i$.

For the purpose of this analysis we consider the value of the option at planting time for each crop.

**Results**

We simulate the model by making 5,000 draws from the crop yield probability distribution and solving for the equilibrium price and quantity outcomes, the deficiency program payments, and welfare effects. Table 4 contains the modeled market outcomes, and table 5 is a modeled disappearance table of the commodity usage in each sector.

Table 4: Model Market Outcomes

| | | | Corn | Soybean | Wheat |
|---|---|---|---|---|---|
| Baseline | | | | | |
| | Price | | Corn | Soybean | Wheat |
| | | Level | $2.32 | $5.62 | $4.40 |
| | | volatility | 0.23 | 0.22 | 0.22 |
| | Land Use | | | | |
| | | Proportions | 0.37 | 0.37 | 0.26 |
| | | (mil acres) | 197 | | |
| Baseline + 3 billion gallons ethanol and 150 mil gals biodiesel | | | | | |
| | Price | | Corn | Soybean | Wheat |
| | | Level | $2.80 | $6.17 | $4.43 |
| | | volatility | 0.25 | 0.23 | 0.22 |
| | | Level Change | $0.48 | $0.55 | $0.03 |
| | Land Use | | | | |
| | | Proportions | 0.37 | 0.37 | 0.26 |
| | | (mil acres) | 198 | | |
| Baseline + 6 billion gallons ethanol and 300 mil gals biodiesel | | | | | |
| | Price | | Corn | Soybean | Wheat |
| | | Level | $3.45 | $6.83 | $4.46 |
| | | volatility | 0.29 | 0.25 | 0.22 |
| | | Level Change | $1.13 | $1.21 | $0.06 |
| | Land Use | | | | |
| | | Proportions | 0.37 | 0.37 | 0.26 |
| | | (mil acres) | 198 | | |
| Baseline + 9 billion gallons ethanol and 450 mil gals biodiesel | | | | | |
| | Price | | Corn | Soybean | Wheat |
| | | Level | $4.29 | $7.61 | $4.66 |
| | | volatility | 0.34 | 0.27 | 0.22 |
| | | Level Change | $1.97 | $1.99 | $0.26 |
| | Land Use | | | | |
| | | Proportions | 0.38 | 0.37 | 0.26 |
| | | (mil acres) | 199 | | |

*Baseline contains 3.4 billion gallons ethanol and 25 million gallons biodiesel production (2004 levels).

Table 5: Model Disappearance Table – (million bushels)

| | Feed | Food | Export | Biofuel | Total | Supply |
|---|---|---|---|---|---|---|
| **Baseline*** | | | | | | |
| Corn | 5,970 | 1,653 | 1,920 | 972 | 10,515 | 10,515 |
| Soybean | 1,987 | 1,987 | 974 | 17 | 2,978 | 2,978 |
| Wheat | - | 1,130 | 954 | - | 2,088 | 2,088 |
| **Baseline + 3 billion gallons ethanol and 150 mil gals biodiesel** | | | | | | |
| Corn | 5,416 | 1,635 | 1,657 | 1,831 | 10,539 | 10,539 |
| Soybean | 1,960 | 1,960 | 905 | 119 | 2,985 | 2,985 |
| Wheat | - | 1,129 | 951 | - | 2,084 | 2,084 |
| **Baseline + 6 billion gallons ethanol and 300 mil gals biodiesel** | | | | | | |
| Corn | 4,872 | 1,617 | 1,412 | 2,689 | 10,590 | 10,590 |
| Soybean | 1,933 | 1,933 | 837 | 222 | 2,991 | 2,991 |
| Wheat | - | 1,129 | 947 | - | 2,081 | 2,081 |
| **Baseline + 9 billion gallons ethanol and 450 mil gals biodiesel** | | | | | | |
| Corn | 4,377 | 1,598 | 1,202 | 3,547 | 10,724 | 10,724 |
| Soybean | 1,903 | 1,903 | 770 | 324 | 2,997 | 2,997 |
| Wheat | - | 1,126 | 920 | - | 2,053 | 2,053 |

*Baseline contains 3.4 billion gallons ethanol and 25 million gallons biodiesel production (2004 levels).
**Recall that the food and feed demand for soybeans represents the same bushels; e.g., 1,903 + 770 + 324 = 2,997 million bushels.

In the baseline scenario market prices were $2.32, $5.62, and $4.40 for corn, soybeans, and wheat respectively. In the scenarios, the mean of the equilibrium price distributions for corn soybeans and wheat ranged from $2.80, $6.17, and $4.43 in the +3 scenario to $4.29, $7.61, and $4.66 in the +9 scenario.

Not only did the level of commodity prices increase, but also the volatility. Corn price volatility went from 23% in baseline to 34% in the +9 scenario, while soybean price volatility increased from 22% to 27% and wheat price volatility remained constant at 22%. The probability distributions of the commodity prices in each scenario are depicted in figures 3 through 5.

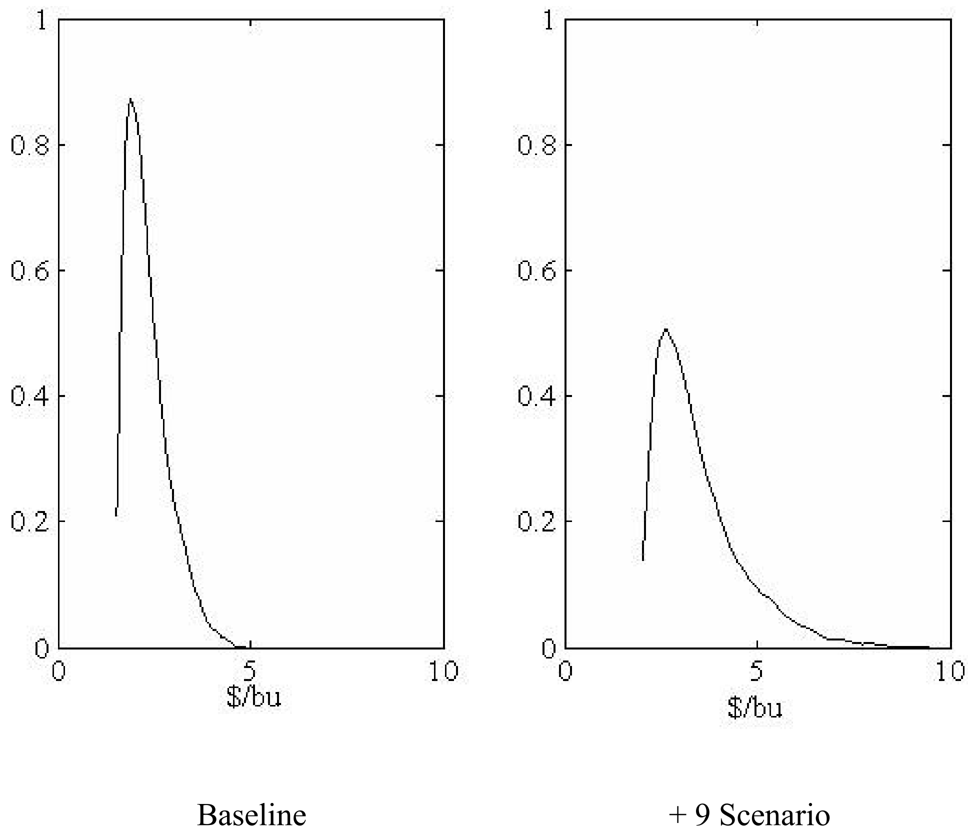
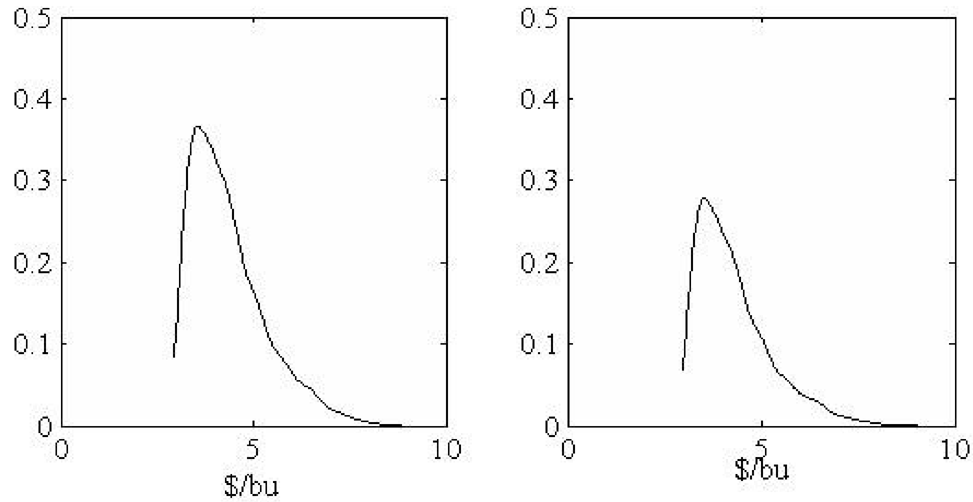



Figure 3: Probability Distribution of Corn Price




Baseline + 9 Scenario

Figure 4: Probability Distribution of Soybean Price

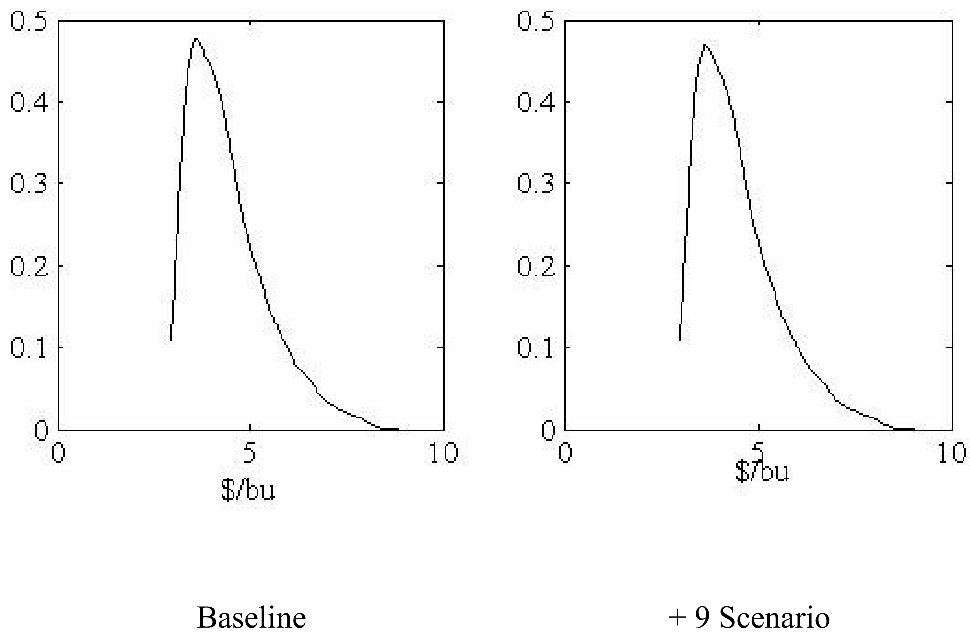



Figure 5: Probability Distribution of Wheat Price

Crop acreage increased from just over 197 million acres in the baseline to 199 million acres in the +9 scenario, and land allocation shifted slightly in favor of corn production and away from wheat as more ethanol production increased across the scenarios.

Increased demand for corn and soybean oil in biofuel causes equilibrium prices to be higher, resulting in a reduction in the quantity demanded of corn, soybeans, and wheat for livestock feed, food, and export. Exports of each commodity, being the most elastic, experience the largest reduction as is shown in table 5.

The simulated welfare effects of the three scenarios are contained in tables 6-8. In each case, the groups with net welfare gains are farmers and consumers of transportation fuel. Farmers experience a welfare loss on the order of $2 billion because the increase in crop prices pushes the LDP and CCP options out of the money, but they benefit much more from the increase in market prices of the commodities. They gain from $6.5 to $14.4 to $23.4 billion in surplus across the three scenarios.

Table 6: Welfare effects (millions of dollars) –Baseline + 3 billion gallons

| | | Corn | Soybean | Wheat | Total |
|---|---|---|---|---|---|
| *Change in PS* | | | | | |
| Market | (+) | $4,841 | $1,592 | $83 | $6,516 |
| Lost LDP | (-) | ($561) | ($354) | $0 | ($915) |
| Lost CCP | (-) | ($704) | ($76) | ($3) | ($783) |
| Total | | $3,576 | $1,162 | $80 | $4,818 |
| Change in CS | | | | | |
| Feed | (-) | ($2,585) | ($945) | -- | ($3,530) |
| Food | (-) | ($786) | ($1,075) | ($31) | ($1,892) |
| Export | (-) | ($789) | ($480) | ($25) | ($1,294) |
| Total | | ($4,160) | ($2,500) | ($56) | ($6,716) |
| | | | | Total Domestic | ($5,422) |
| Consumers of fuel | (+) | $1,530 | $150 | -- | $1,680 |
| Change in Taxpayer Outlays | | | | | |
| LDP | (+) | $561 | $354 | $0 | $915 |
| CCP | (+) | $704 | $76 | $3 | $783 |
| Blender's Credit | (-) | ($1,530) | ($150) | -- | ($1,680) |
| Total | | ($265) | $280 | $3 | $18 |
| | | | Net | Domestic | Total |
| | | | Mean | $1,093 | ($201) |
| | | | St. Dev. | $332 | $212 |

Table 7: Welfare effects (millions of dollars) –Baseline + 6 billion gallons

| | | Corn | Soybean | Wheat | Total |
|---|---|---|---|---|---|
| *Change in PS* | | | | | |
| Market | (+) | $10,747 | $3,470 | $190 | $14,408 |
| Lost LDP | (-) | ($633) | ($581) | $0 | ($1,214) |
| Lost CCP | (-) | ($1,084) | ($146) | ($6) | ($1,236) |
| Total | | $9,030 | $2,743 | $184 | $11,957 |
| Change in CS | | | | | |
| Feed | (-) | ($5,687) | ($2,025) | -- | ($7,712) |
| Food | (-) | ($1,835) | ($2,327) | ($65) | ($4,227) |
| Export | (-) | ($1,679) | ($1,001) | ($52) | ($2,733) |
| Total | | ($9,201) | ($5,354) | ($117) | ($14,672) |
| | | | | Total Domestic | ($11,939) |
| Consumers of fuel | (+) | $3,060 | $300 | -- | $3,360 |
| Change in Taxpayer Outlays | | | | | |
| LDP | (+) | $633 | $581 | $0 | $1,214 |
| CCP | (+) | $1,084 | $146 | $6 | $1,236 |
| Blender's Credit | (-) | ($3,060) | ($300) | -- | ($3,360) |
| Total | | ($1,343) | $427 | $6 | ($910) |
| | | | Net | Domestic | Total |
| | | | Mean | $2,469 | ($264) |
| | | | St. Dev. | $1,258 | $814 |

Table 8: Welfare effects (millions of dollars) –Baseline + 9 billion gallons

| | | Corn | Soybean | Wheat | Total |
|---|---|---|---|---|---|
| *Change in PS* | | | | | |
| Market | (+) | $16,897 | $5,717 | $804 | $23,417 |
| Lost LDP | (-) | ($633) | ($681) | $0 | ($1,314) |
| Lost CCP | (-) | ($1,134) | ($204) | ($20) | ($1,358) |
| Total | | $15,130 | $4,832 | $784 | $20,745 |
| Change in CS | | | | | |
| Feed | (-) | ($9,200) | ($3,272) | -- | ($12,472) |
| Food | (-) | ($3,174) | ($3,803) | ($291) | ($7,269) |
| Export | (-) | ($2,623) | ($1,570) | ($231) | ($4,424) |
| Total | | ($14,997) | ($8,646) | ($522) | ($24,165) |
| | | | | Total Domestic | ($19,740) |
| Consumers of fuel | (+) | $4,590 | $450 | -- | $5,040 |
| Change in Taxpayer Outlays | | | | | |
| LDP | (+) | $633 | $681 | $0 | $1,314 |
| CCP | (+) | $1,134 | $204 | $20 | $1,358 |
| Blender's Credit | (-) | ($4,590) | ($450) | -- | ($5,040) |
| Total | | ($2,823) | $435 | $20 | ($2,368) |
| | | | Net | Domestic | Total |
| | | | Mean | $3,677 | ($748) |
| | | | St. Dev. | $3,482 | $2,629 |

Consumers or blenders of transportation fuel receive welfare gains from biofuel tax credits in an amount equal to taxpayer liability from the blender's credit. The model does not distinguish how much of this welfare gain is passed through to retail gasoline consumers. However, Du and Hayes (2009) examine the effect of the ethanol industry on retail gasoline prices and find that ethanol production has reduced retail gasoline prices by \$0.29 to \$0.40 per gallon depending on the region. This suggests that blenders pass at least some of the welfare gains on to retail gasoline consumers.

Interestingly, taxpayers are not net losers in the first scenario. Increasing ethanol production by 3 billion gallons causes the LDP and CCP payment options to move out of the money reducing their value by more than taxpayers spent in tax credits to biofuel. As biofuel expands further this effect is overturned. Once biofuel production expands far enough, the LDP and CCP options are so far out of the money that they are effectively worthless and thus taxpayer savings are capped.

It is the consumer, though, that is the biggest loser with welfare losses ranging from \$6.7 to \$24.1 billion across the scenarios. The animal feeding industry experiences the largest losses, since they are the largest user of the corn and soybeans. The export sector, with the most elastic demand and small usage share, experiences the smallest welfare loss.

Aggregate welfare loss ranges from approximately \$260 million to \$750 million. It is interesting to note, however, that there is a net positive *domestic* welfare effect on the order of \$2 billion; i.e., ignoring the welfare of the export market, the remaining welfare effects are net positive. Assuming policymakers have an objective to increase the production of biofuels, and if policymakers place equal weight on domestic and foreign players in the agricultural markets, then the biofuel policy has a net negative welfare effect. If, however, policymakers disregard welfare losses in the export market, then there is a net positive welfare effect from increased biofuel production from their perspective. Table 9 contains a summary of the welfare effects across scenarios in both levels and per billion gallons of ethanol added to the scenario.

Table 9: Summary of Transfers

| | Scenario | *Change in PS* | *Change in CS* | *Taxpayer Outlays* | *Net* |
|---|---|---|---|---|---|
| Levels | Baseline + 3* | $4,805 | ($6,716) | $31 | ($201) |
| | Baseline + 6 | $11,940 | ($14,672) | ($892) | ($264) |
| | Baseline + 9 | $20,725 | ($24,165) | ($2,348) | ($748) |
| Per bil gals ethanol added | Baseline + 3 | $1,602 | ($2,239) | $10 | ($67) |
| | Baseline + 6 | $1,990 | ($2,445) | ($149) | ($44) |
| | Baseline + 9 | $2,303 | ($2,685) | ($261) | ($83) |

We should note that these results abstract from any distortions occurring in the fuel markets; so these aggregate results should not be viewed as aggregate in terms of the economy at large but aggregate in the agricultural sector only. Market distortions causing more than the economically efficient amount of ethanol or gasoline to be produced and consumed would make the welfare losses larger than are reported here.

**Sensitivity Analysis**

Our model relies on the use of many parameters borrowed from outside sources. We perform a sensitivity analysis on two parameters that are likely to affect the outcomes, the own price elasticity of corn feed demand and the own price elasticity of soybean export demand. The welfare change between the baseline and the +6 scenario is used to conduct the sensitivity analysis. We vary the absolute value of the parameter in question and report an abbreviated version of the welfare results. The sensitivity analysis is found in table 10 and the results from the previous analysis are in the first row for comparison.

Table 10: Sensitivity analysis

| Parameter scenario | *Change in PS* | *Change in CS* | *Taxpayer Outlays* | *Net* |
|---|---|---|---|---|
| Original | \$11,940 | (\$14,672) | (\$892) | (\$264) |
| $\alpha_1^1$ **+5%** | \$11,842 | (\$14,404) | (\$1,134) | (\$336) |
| $\alpha_1^1$ **−5%** | \$12,777 | (\$15,240) | (\$1,024) | (\$127) |
| $\gamma_2^1$ **+5%** | \$13,779 | (\$15,933) | (\$1,582) | (\$376) |
| $\gamma_2^1$ **−5%** | \$13,375 | (\$15,367) | (\$1,584) | (\$216) |

Sensitivity analysis done with respect to the +6 scenario

**Comparison with the Existing Literature**

Several researchers have recently conducted analyses on the welfare effects of biofuel production. A side-by-side comparison is provided in table 11of the market structure and important modeling assumptions of these studies. The numerical welfare results from each study outlined are contained in table 12. Table 13 contains the same information as table 12, but is presented on a per billion gallons of ethanol basis. The estimates of welfare change range from -\$3.1 billion to \$1.3 billion, compared to our results that are on the order of -\$200 to -\$750 million depending on the scenario.

Table 11: Components included in recent welfare analyses

| | This study | Babcock 2008 | de Gorter & Just 2009 | Du et al. 2009 | Gardner 2007 | McPhail Babcock 2008 | Schmitz et al. 2007 |
|---|---|---|---|---|---|---|---|
| *Markets considered* | | | | | | | |
| Corn | X | X | X | X | X | X | X |
| Soybeans | X | | | | | | |
| Wheat | X | | | | | | |
| Ethanol | | X | X | X | X | X | X |
| *Demand* | | | | | | | |
| Aggregate | X | X | X | X | X* | X | X |
| Feed | X | X | | | | X | X |
| Food | X | X | | | | X | X |
| Export | X | X | | | | X | X |
| *Farm programs* | | | | | | | |
| LDP | X | | X | | X[**] | | X |
| CCP | X | | X | | | | X |
| Stochastic | X | | | | | X | |

[*]Distinguishes ethanol and non-ethanol demand for corn
[**]Not the traditional deficiency payment program. Gardner calculates the benefit of a transfer directly to corn growers instead of the ethanol subsidy

A paper by de Gorter and Just (2009) calculates the welfare effects of the U.S. ethanol tax credit and they note that the tax credit causes significant rectangular welfare losses. Their modeling framework includes the corn market with demand disaggregated into foreign and domestic producers, with the excess supply of corn absorbed by the ethanol industry. They then model the fuel market to calculate the effect of ethanol in the domestic fuel market. They find net welfare losses from the ethanol tax credit.

Schmitz et al. (2007) calculate the welfare costs and benefits of U.S. ethanol production, and find a positive aggregate welfare gain. The driver of this result is that while tax revenue is decreased because of the blender's tax credit, taxpayer liabilities are reduced by more than this amount because subsidies to corn farmers are reduced. They conduct their analysis using a deterministic model. Welfare effects to non-ethanol users of corn are separated among food, alcohol, and industrial use verses feed and residual use.

Du et al. (2009) perform a welfare analysis focusing on the fuel markets for both gasoline and ethanol on a energy equivalent basis, and the market for corn. They find the net welfare effect of ethanol production to be negative. Babcock (2008) uses a deterministic model of the ethanol and corn markets, distinguishing demand for corn as coming from feed, food, and exports. A paper written by McPhail and Babcock (2008) incorporates uncertainty in corn yields, corn demand, and ethanol demand, but does not include the cross market effects with soybeans and wheat.

Gardner (2007) uses a deterministic model of the ethanol market including corn producers, users of corn including ethanol, feed and exports, taxpayers, ethanol producers, and ethanol consumers.

The present study is the only one of the articles discussed to treat farm program payments as an option farmers own and taxpayers are obligated to honor. We therefore include the reduced farm program payment as a reduction in farmer surplus as well as a decrease in taxpayer cost. In an aggregate welfare analysis, the reduction of farmer surplus from the decreased value of deficiency payments should be included in calculating the farmer's welfare change.

This study is also the first to include the indirect welfare effects on the soybeans and wheat markets. These markets are important to consider in a welfare analysis since they are produced on the same type of land and compete for acreage. An equilibrium increase in one commodity's price will affect equilibrium prices in related markets as well.

Table 12: Comparison of Recent Welfare Studies (Millions of Dollars)

| | This study | Babcock 2008 | de Gorter & Just 2009 | Du et al. 2009 | Gardner 2007 | McPhail Babcock 2008 | Schmitz et al. 2007 |
|---|---|---|---|---|---|---|---|
| *Ag Producers* | | | | | | | |
| Market | $14,408 | $4,108 | $1,484 | $7,430 | $2,029 | $1,581 | $1,154 |
| Lost CCP | ($1,214) | -- | -- | -- | -- | -- | -- |
| Lost LDP | ($1,236) | -- | ($1,388) | -- | -- | -- | -- |
| Total | $11,957 | $4,108 | $96 | $7,430 | $2,029 | $1,581 | $1,154 |
| *Ag Consumers* | | | | | | | |
| Feed | ($7,712) | ($1,990) | ($1,484) | -- | -- | ($730) | ($1,008) |
| Food | ($4,227) | ($400) | | -- | -- | ($197) | ($3,094) |
| Export | ($2,733) | ($695) | $433 | -- | -- | ($276) | ($993) |
| Total | ($14,672) | ($3,085) | ($1,051) | ($10,440) | ($1,731) | ($1,203) | ($5,095) |
| *Consumers of Trans. Fuel* | $3,360 | $3,291 | $1,606 | $1,660 | $1,428 | $2,337 | $3,883 |
| *Taxpayers* | | | | | | | |
| LDP/CCP | $2,450 | -- | $1,388 | $3,450 | -- | -- | $4,084 |
| Blender's Credit | ($3,360) | ($7,450) | ($3,330) | ($2,990) | ($2,600) | ($1,391) | ($2,761) |
| Total | ($910) | ($7,450) | ($1,942) | ($890) | ($2,600) | ($1,391) | $1,323 |
| Net Change | ($264) | ($3,137) | ($1,291) | ($780) | ($665) | ($568) | $1,281 |

*Comparison uses the +6 scenario of this study

Table 13: Comparison of recent welfare studies – per billion gallons of ethanol (millions of dollars)

| | This study | Babcock 2008 | de Gorter & Just 2009 | Du et al. 2009 | Gardner 2007 | McPhail Babcock 2008 | Schmitz et al. 2007 |
|---|---|---|---|---|---|---|---|
| *Change in Ethanol Production* | 6.00 | 4.00 | 2.7 | 5.9 | 2.60 | 1.50 | 2.00 |
| *Ag Producers* | | | | | | | |
| Market | $2,401 | $1,027 | $550 | $1,259 | $780 | $1,450 | $577 |
| Lost CCP | ($202) | -- | -- | -- | -- | -- | -- |
| Lost LDP | ($206) | -- | ($514) | -- | -- | -- | -- |
| Total | $1,993 | $1,027 | ($36) | $1,259 | $780 | $1,450 | $577 |
| *Ag Consumers* | | | | | | | |
| Feed | ($1,285) | ($498) | ($550) | -- | -- | ($670) | ($504) |
| Food | ($705) | ($100) | | -- | -- | ($181) | ($1,547) |
| Export | ($456) | ($174) | $160 | -- | -- | ($253) | ($497) |
| Total | ($2,445) | ($771) | ($389) | ($1,770) | ($666) | ($1,104) | ($2,548) |
| *Consumers of Trans. Fuel* | $560 | $823 | $595 | $281 | $549 | $2,144 | $1,942 |
| *Taxpayers* | | | | | | | |
| LDP/CCP | $408 | -- | $514 | $585 | -- | -- | $2,042 |
| Blender's Credit | ($560) | ($1,863) | ($1,233) | ($507) | ($1,000) | ($1,276) | ($1,381) |
| Total | ($152) | ($1,863) | $719 | ($151) | ($1,000) | ($1,276) | $662 |
| Net Change | ($44) | ($784) | ($478) | ($133) | ($256) | ($521) | $641 |

*Comparison uses the +6 scenario of this study

The range of estimates among these articles illustrates how sensitive welfare studies are to model assumptions – particularly to the model structure imposed. A model that assumes a market functioning in isolation will yield quantitatively different results than a model that includes market interactions; every piece left out distorts the picture of where transfers are going, the size of aggregate transfer, and deadweight loss. For example, in this analysis the size of transfer from consumers to producers is larger than in the other articles. This is because ignoring the soybean and wheat market effects understates the size of the transfers.

Further, the studies we reviewed each draw elasticity estimates from different sources compounding the issue that the models are constructed in very different ways. Since collectively the welfare estimates vary greatly an interpretation of the results is difficult, especially for policy recommendations.

**Conclusion**

We conduct a welfare analysis of the effect of an increase in the size of the biofuel industry, including some features that are not present in the existing literature on the subject. Our analysis includes uncertainty in crop yields, soybean and wheat markets, and disaggregated demands for the commodities.

Including uncertainty allowed us to value more appropriately the change in deficiency payments. LDP and CCPs are made based on the realization of uncertain crop prices and should be treated in the fashion of a financial option; therefore, only a model incorporating uncertainty is able to assess the change in value of these programs.

Including the soybean and wheat markets is important because these markets are linked to corn through competition for acreage suggesting that the expected price of each crop influences the supply decisions of all the crops. The indirect welfare effects in the soybean and wheat markets can severely understate the size of transfer caused by an increase in biofuel production.

Our results suggest that increased biofuel production results in welfare losses on the order of $200 to $750 million depending on the size of increase in the biofuel industry and resulted in a transfer largely from consumers to producers. If the welfare of only domestic consumers is considered (ignoring the welfare of those in the export market), then the welfare affect is net positive. This analysis was focused on the agricultural sector and abstracts from potential distortions or exogenous events occurring in the (fossil) fuel markets.

**References**

AltIn, R., S. Çetinkaya, and H. S. Yücesu. "The potential of using vegetable oil fuels as fuel for diesel engines." *Energy Conversion and Management* 42, no. 5(2001): 529-538.

Babcock, B. A. "Distributional Implications of U.S. Ethanol Policy." *Review of Agricultural Economics* 30, no. 3(2008): 533-542.

Bullock, D. S. "Multimarket effects of technological change: Comment." *Review of Agricultural Economics* (1993): 603-608.

de Gorter, H., and D. R. Just. "The Welfare Economics of a Biofuel Tax Credit and the Interaction Effects with Price Contingent Farm Subsidies." *American Journal of Agricultural Economics* 91, no. 2(2009): 477-488.

Du, X., and D. J. Hayes. "The Impact of Ethanol Production on U.S. and Regional Gasoline Prices and on the Profitability of the U.S. Oil Refinery Industry." *Energy Policy* Forthcoming(2009).

Du, X., D. J. Hayes, and M. L. Baker. "A Welfare Analysis of the U.S. Ethanol Subsidy." *Review of Agricultural Economics* forthcoming(2009).

Gardner, B. "Fuel Ethanol Subsidies and Farm Price Support." *Journal of Agricultural and Food Industrial Organization* 5, no. 4(2007).

Gardner, B. L. "Commodity Options for Agriculture." *American Journal of Agricultural Economics* 59, no. 5(1977): 986-992.

Hull, J. *Fundamentals of futures and options markets*: Prentice Hall, 2002.

Hull, J. C. *Fundamentals of Futures and Options Markets*. 4th ed. Upper Saddle River, NJ: Prentice Hall, 2001.

Larson, R., R. Hosteltler, and B. Edwards. *Multivariable Calculus*. Seventh ed. Edited by J. Shira. Boston, MA: Houghton Mifflin Company, 2002.

Magnussen, S. "An algorithm for generating positively correlated Beta-distributed random variables with known marginal distributions and a specified correlation." *Computational Statistics & Data Analysis* 46, no. 2(2004): 397-406.

Marcus, A. J., and D. M. Modest. "The valuation of a random number of put options: An application to agricultural price supports." *Journal of Financial and Quantitative Analysis* (1986): 73-86.

McPhail, L. L., and B. A. Babcock. "Short-Run Price and Welfare Impacts of Federal Ethanol Policies." *CARD Working Paper Series* 08-WP 468(2008).

Schmitz, A., C. B. Moss, and T. G. Schmitz. "Ethanol: No Free Lunch." *Journal of Agricultural and Food Industrial Organization* 5, no. 3(2007).

Shapouri, H., and P. Gallagher. "USDA's 2002 Ethanol Cost-of-Production Survey." United States Department of Agriculture.

Shurson, G. C., et al. (2003) Value and use of "new generation" distiller's dried grains with solubles in swine diets. Lexington, KY.

# CHAPTER 4: INSIGHT INTO A POSSIBLE EQUILIBRIUM RELATIONSHIP BETWEEN THE PRICE OF CORN AND ETHANOL – ANALYSIS AND A SMALL SAMPLE MONTE CARLO STUDY

## Introduction

Energy markets, particularly oil and gasoline, seem to have an increasingly important influence on the corn market. It makes intuitive sense for there to be a link between corn and ethanol prices because the existence of a large ethanol sector makes corn an energy commodity in addition to an edible commodity. Corn's value as an energy commodity will be larger than its value as an edible commodity if energy prices are high enough. This implies that if corn is priced at the margin by its value as an input in producing energy, then it may respond to the fundamentals in the energy markets as much as it would respond to fundamentals in the agricultural markets. This paper explores the statistical evidence that there exists a link between the corn and energy sector.

Recent research attempted to pin down the relationship between energy and agriculture created by corn-based ethanol production. Tokgoz et al. (2007) use the model maintained by the Food and Agricultural Policy Research Institute (FAPRI) to make long run projections of the effect of biofuel production on commodity prices and production. Tokgoz et al. (2008) again uses the FAPRI model to simulate the effect of an exogenous event in one market on other markets; in particular they explore the effect of a spike in crude oil price and the effect of a significant drought coupled with a renewable fuels mandate. Kruse et al. (2007) use a long run relationship to analyze the effect of removing biofuel subsidies, and Thompson et al. (2009) in a similar analysis as Kruse et al. (2007) examine the covariance among corn, ethanol, and oil markets.

We develop a theory that says a long run relationship between two or more price series in futures markets can be transmitted to spot prices through intertemporal arbitrage. This explains why we may observe spot prices behaving in such a way that is consistent with long-run equilibrium, but not required for a short-run equilibrium. We then try to determine if there is statistical support for the hypothesis that corn and ethanol prices maintain an equilibrium relationship. If there is an equilibrium relationship between corn and ethanol prices, it has been in place for a relatively short time. Since the statistical methods available for testing this hypothesis, namely the cointegration tests developed by Johansen (1991), rely on asymptotic properties of the test statistic, a detailed discussion of their small sample properties is necessary.

The paper is organized as follows. In the following section we describe the long run equilibrium condition that requires the ethanol industry to earn zero economic profit in the long run; we posit that futures prices far from maturity should be related according to this breakeven relationship. The next section shows how the futures market could transmit this relationship to spot prices through intertemporal arbitrage. We then test this theory using the cointegration tests developed by Johansen (1991). The small sample properties of the Johansen statistics are discussed next, including a review of some previous Monte Carlo work regarding these statistics. We also tailor a Monte Carlo study with a data generating process that is intended to mimic our data in some important ways. A final section concludes by summarizing our findings.

**A Theory of the Link between the Corn and Ethanol Markets**

Tokgoz et al. (2007) first provided intuition for why the long run price of corn might be drawn to the level at which the ethanol industry breaks even. The logic is that if the price of corn is too low, ethanol plants enter the market and bid the corn price higher, and if the price of corn is too high, ethanol plants exit putting downward pressure on corn prices. This break-even condition should impose a long run equilibrium relationship between the price of corn and the price of ethanol if the ethanol industry is large enough. We expand on this theory and describe the mechanism by which this long run relationship between corn and ethanol prices can be transmitted to spot prices.

The long run zero economic profit or breakeven rule is that *Total Revenue – Total Cost* = 0. If the price of ethanol at time *t* is $p_t^{eth}$, the price of corn at time *t* is $p_t^c$, the ethanol yield per bushel of corn is 2.8 (2.8 gallons per bushel from (Shapouri and Gallagher, 2005)), the per gallon non-corn cost of producing ethanol is $C_{-corn}$, and the level of tax credit an ethanol producer receives is *TC*, then this breakeven rule is given by $p_t^{eth} + TC = p_t^c \left(1 - 17/56\right)/2.8 + C_{-corn}$.

The (1 – 17/56) in the expression above comes from the fact that the corn-based ethanol production process generates a co-product, distiller's grain, which is used for animal feed as a substitute for corn. For every bushel of corn (56 lbs) processed, an ethanol plant produces 17lbs of distillers grains. Distiller's grains contain approximately the same energy content as corn and thus are valuable as an alternative livestock feed (Shurson, et al., 2003). This means an ethanol plant's feedstock cost is not the full price of corn times the number of bushels of corn processed; e.g., if distiller's grains are valued at par with corn then for every bushel of corn processed ethanol plants only have to pay for (1 – 17/56) times the price of a bushel of corn. The remaining (17/56) comes back to them when they sell the distiller's grain.

Solving the breakeven rule for $p_t^c$, an expression for the energy value of corn (per gallon of ethanol) is:

(1) $$p_t^c = 2.8\left[ p_t^{eth} + TC - C_{-corn} \right] \Big/ \left(1 - {}^{17}\!\!/_{56}\right)$$

This theory of a breakeven relationship between corn and ethanol suggests that corn and ethanol prices should maintain a linear relationship with one another.

*Futures prices*

Assuming corn and energy futures markets are relatively efficient (in the sense of Fama (1970) and Malkiel (2003)) implies that deviations from the equilibrium relationship between corn and ethanol prices posited by equation (1) cannot be violated in the long run. The far to maturity futures contracts provide a signal to the ethanol industry to expand or contract. Speculators in futures markets can recognize this and take positions that allow them to gain when the relative price of corn and ethanol return to their equilibrium relationship.

For example, if corn is trading for less than its energy value for delivery two years from now, the market is sending a signal for additional ethanol plants to be built. This expansion will cause the price of corn to rise relative to the price of ethanol. Conversely, if corn is trading higher than its energy value, ethanol production will decrease and the price of corn relative to ethanol will fall.

If corn is selling below its energy value a speculator can potentially profit from buying corn and selling ethanol. Denote the time $t$ futures price of corn for delivery at time $\bar{T}$ by $F_{t,\bar{T}}^{c}$ and the time $t$ futures price of ethanol for delivery at time $\bar{T}$ by $F_{t,\bar{T}}^{eth}$. The presence of traders who take positions based on the spread between the corn and ethanol price means the relationship $F_{t,\bar{T}}^{c} = 2.8\left[F_{t,\bar{T}}^{eth} + TC - C_{-corn}\right] / \left(1 - 17/56\right)$ should hold in the futures markets whose time to delivery is long enough away that the size of the ethanol industry can respond to market incentives by expanding or contracting.

*Intertemporal Arbitrage and the Spot-Futures Price Relationship*

The theory above is only applicable to expectations about corn prices in the future. In the short-run the ethanol industry cannot quickly expand to take advantage of inexpensive spot prices of corn as there is some time lag involved in constructing a plant. Thus, there is no reason to expect a relationship between corn and ethanol spot prices on these grounds.

However, both corn and ethanol are storable commodities, and storage provides a mechanism by which a relationship that is maintained between two or more futures prices can be transmitted to the spot prices. A common model to explain the intertemporal price behavior of a commodity is the cost of carry model (Hull, 2002). If the spot price of ethanol at time $t$ is $p_t^{eth}$, and assuming the cost of storage is a constant proportion of the spot price, $u$, then the expected spot price of corn at time $\bar{T}$ should be $E_t\left(p_{\bar{T}}^{c}\right) = p_t^{c} e^{u(\bar{T}-t)}$. If for example the expected future spot price is higher than the spot price at time $t$ inflated by the cost of carry then one could buy ethanol at time *t*, store it and expect to make a profit on the spot market at time $\bar{T}$.

Assuming that the futures price of a commodity is an unbiased expectation of the future spot price, then $F_{t,\bar{T}}^{eth} = p_t^{eth} e^{u(\bar{T}-t)}$ , the spot price of ethanol at time *t* influences the futures price of ethanol for delivery at time $\bar{T}$. Then the breakeven relationship in ethanol production requires that $F_{t,\bar{T}}^{c} = 2.8\left[F_{t,\bar{T}}^{eth} + TC - C_{-corn}\right] / \left(1 - {}^{17}\!/_{56}\right)$ , and finally the intertemporal storage arbitrage in the corn market requires that $p_t^c = F_{t,\bar{T}}^{c} e^{-u(\bar{T}-t)}$ (see figure 1).

So the enforcement of a long run breakeven condition in the ethanol market causes the expected future prices of corn and gasoline to maintain an equilibrium relationship; then the intertemporal arbitrage between the futures and spot prices can cause this relationship to be transmitted to the spot market as well.

$$
\begin{tikzcd}
F^{eth}_{t,T} \arrow[r] & F^{corn}_{t,T} \arrow[d] \\
p^{eth}_{t} \arrow[u] & p^{corn}_{t}
\end{tikzcd}
$$

Figure 1: Intertemporal Arbitrage and the Spot-Futures Price Relationship

In the parlance of time series analysis, the theory above suggests the long run break-even condition in the ethanol market requires that the price of corn and the price of ethanol be cointegrated. Non-stationary time series are said to be cointegrated if there exists at least one linear combination of the variables that is itself stationary (Engle and Granger, 1987, Granger, 1981, Granger and Weiss, 2001). The following section describes the methods for detecting cointegration among time series; we employ these to determine if there is evidence of an equilibrium relationship between corn and ethanol prices.

**Testing for Cointegration**

A vector, $y_t$, of $n$ cointegrated time series are written in vector error correction form by,

(1) $$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{i=1}^{k-1} \Phi_i \Delta y_{t-i} + \mu_t + \varepsilon_t \quad \text{for } t = 1, \ldots, T,$$

where $\alpha$ is an $n \times r$ matrix of speed of adjustment coefficients, $\beta'$ is an $r \times n$ matrix of cointegrating vectors, the $\Phi_i$ are $n \times n$ matrices, and $\mu_t$ is a vector of intercept terms . This representation is the same as a traditional vector autoregression (VAR) in first differences, but with the additional error correction term ($\alpha\beta' y_{t-1}$) that ensures the series maintain the equilibrium relationships defined by $\beta' y_{t-1}$ with $\alpha$ a vector of speed of adjustment coefficients (Hamilton, 1994).

The first methods, developed by Engle and Granger (1987) and Phillips and Ouliaris (1990), test for the presence of an equilibrium relationship between two or more time series which are integrated of order one, $I(1)$, by examining the fitted residuals from a cointegrating regression. The basic idea behind these tests is to partition the vector $y_t$ from (1) into $y_t = \left[y_t^1 , y_t^2\right]$, where $y_t^1$ is a $T \times 1$ dimensional vector, and $y_t^2$ is a $T \times n-1$ dimensional vector and estimate the regression equation, $y_t^1 = \beta_0 + \beta_1 y_t^2 + \varepsilon_t$. If the fitted residuals, $\hat{\varepsilon}_t$, are stationary, then one can conclude the $1 \times n+1$ vector $\left[1 \quad -\beta_0 \quad -\beta_1\right]$ is a cointegrating vector for the series $y_t$. Phillips and Ouliaris (1990) note that the residual based tests are easy to apply and intuitively appealing, but also have the undesirable property that the results are sensitive to the choice of normalization (i.e., which series is chosen for $y^1$ ).

Johansen (1991, 1988), developed maximum likelihood techniques to determine the rank of and test for linear restrictions on the matrix $\alpha\beta'$. This is useful because the number of cointegrating relationships in the system can be estimated without specifying the nature causation, and it can test for the presence of multiple cointegrating vectors whereas the earlier residual based tests can only estimates one relationship at a time. In the VECM, the rank of the $n$ x $n$ matrix $\alpha\beta'$ in equation (1) determines the number of equilibrium relationships or cointegrating vectors in the system. The time series, $y$, are:

- stationary if $rank(\alpha\beta') = n$, or
- $I(1)$ and not cointegrated if $rank(\alpha\beta') = 0$; i.e., a traditional VAR model in first differences is the appropriate model, or
- cointegrated with $r$ equilibrium relationships if $rank(\alpha\beta') = r$, and $0 < r < n$.

The procedure for estimating the rank of the matrix $\alpha\beta'$ is detailed in Johansen (1991) and Hamilton (1994) as follows:

1) Perform auxiliary regressions.
    - Regress $\Delta y_t$ on $[1 \quad \Delta y_{t-1} \quad \Delta y_{t-2} \quad \cdots \quad \Delta y_{t-k+1}]$ and obtain residuals $\hat{u}_t$.
    - Regress $y_t$ on $[1 \quad \Delta y_{t-1} \quad \Delta y_{t-2} \quad \cdots \quad \Delta y_{t-k+1}]$ and obtain residuals $\hat{w}_t$.
2) Calculate the sample covariance matrices of the residuals
    - $\hat{\Sigma}_{\mathrm{VV}} = (1/T)\sum_{t=1}^{T}\hat{v}_t\hat{v}_t'$
    - $\hat{\Sigma}_{\mathrm{UU}} = (1/T)\sum_{t=1}^{T}\hat{u}_t\hat{u}_t'$
    - $\hat{\Sigma}_{\mathrm{UV}} = (1/T)\sum_{t=1}^{T}\hat{u}_t\hat{v}_t'$
3) Calculate the $n$ eigenvalues, $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_n$, of the matrix $M = \hat{\Sigma}_{VV}^{-1}\hat{\Sigma}_{UV}'\hat{\Sigma}_{UU}^{-1}\hat{\Sigma}_{VV}\hat{\Sigma}_{UV}$.
4) Perform hypothesis tests on the cointegrating rank, $r$, using Johansen's trace and/or maximum eigenvalue tests.

The trace and maximal eigenvalue statistics are defined by

(2) $$\lambda_{trace}(r) = -T\sum_{i=r+1}^{n}\ln\left(1-\hat{\lambda}_i\right)$$ , and

(3) $$\lambda_{\max}(r) = -T\ln\left(1-\hat{\lambda}_{r+1}\right)$$ ,

which test the null hypothesis of $r$ equilibrium relationships against the general alternative, and against the alternative of $r+1$ equilibrium relationships respectively. The $\lambda$ statistics are likelihood ratio statistics but are not asymptotically $\chi^2$. The asymptotic distributions of the statistics in (2) and (3) are the same as the distribution of the trace of matrices containing stochastic integrals of Brownian motions. As such, these distributions are non-standard and the percentiles have been tabulated by Monte Carlo simulation by Johansen and Juselius (1990), Osterwald-Lenum (1992), and MacKinnon et al. (1999). The distributions also are dependent upon the specification of the deterministic terms in (1).

**Data**

The data used are daily nearby settlement prices of the corn, ethanol, and soybeans contracts on the Chicago Mercantile Exchange from July 30, 2006 to April 7, 2009 archived at barchart.com. We choose this period as our sample because in 2006, popular interest in ethanol seemed to reach a new level. To illustrate, a search of the Google News[26] archives with the keyword 'ethanol' more than doubled the number of hits from 2005 to 2006 going from 18,000 to 41,000 articles.

[26] http://news.google.com/

Scholarly interest increased during approximately the same time period. A search of the term 'ethanol' in the IDEAS[27] database of published and unpublished scholarly articles in economics and finance returns only 47 articles in 2003, 53 articles in 2004, and 40 articles in 2005; but it returns 227 articles in 2006, 351 articles in 2007, and 427 articles in 2008. This suggests there was a change in the ethanol industry that caused the popular press and academic researchers to take notice in 2006.

The trace and maximal eigenvalue statistics are sensitive to the choice of sample in finite samples. To illustrate we provide a table that reports the statistics' value for different start dates of our data at fifteen-day intervals around the start date of July 30 used in the analysis below (see table 1). This table points to issues with the small sample properties of the statistics, but we postpone a more detailed look consideration of this to a later section.

Table 1: Sensitivity to Sample Analysis – trace and max statistic for ethanol, corn, and soybeans at 15 day intervals around the July 30, 2006 sample start date

| $H_2: r \le 0$ | Trace test | 5% c.v. | 1% c.v | Max test | 5% c.v. | 1% c.v |
|---|---|---|---|---|---|---|
| 6/15/06 | 35.67** | 29.68 | 35.65 | 25.58** | 20.97 | 25.52 |
| 6/30/06 | 40.21** | 29.68 | 35.65 | 30.30** | 20.97 | 25.52 |
| 7/15/06 | 38.88** | 29.68 | 35.65 | 28.74** | 20.97 | 25.52 |
| 7/30/06 | 32.24* | 29.68 | 35.65 | 22.94* | 20.97 | 25.52 |
| 8/16/06 | 34.30* | 29.68 | 35.65 | 25.30* | 20.97 | 25.52 |
| 8/30/06 | 22.62 | 29.68 | 35.65 | 14.96 | 20.97 | 25.52 |

Critical values from Osterwald-Lenum(1992)

No constant term in the cointegrating vector

[27] http://ideas.repec.org/

We are interested primarily in the relationship between corn and ethanol, if one exists, but one might expect the price of soybeans to be present in any equilibrium relationship that involves corn. The two commodities compete for the same acreage, have roughly the same growing season, are impacted by the same weather realizations, and are both used extensively as animal feed. Because of these reasons we should not exclude the price of soybeans in the analysis a priori.

We perform a preliminary analysis to determine whether soybeans should be present in the model. Testing whether all three series are cointegrated, both the trace and the max eigenvalue statistic reject the null hypothesis of no cointegration at the 5% asymptotic level, but cannot reject the null hypothesis that there is one cointegrating vector (note that there can be at most two linearly independent cointegrating vectors for three time series).

However, when the tests are run pair wise the only pair for which the null hypothesis can be rejected is ethanol and corn. See table 2 for a summary of these results. Further, when we estimate an error correction model of the form in equation (1) we find that the error correction term is not significant in the soybean price equation and the price of soybeans is not significant in the cointegrating vector. Causality tests suggest that soybeans do not Granger cause ethanol or corn in the short run sense. Ethanol and corn do not Granger cause soybeans in the short run sense. Therefore, we drop the price of soybeans from the subsequent analysis and focus solely on the price of ethanol and corn.

Table 2: Johansen tests– ethanol corn, and soybeans

| | $H_2$: | Trace test | 5% c.v. | 1% c.v | Max test | 5% c.v. | 1% c.v |
|---|---|---|---|---|---|---|---|
| Eth Corn & Soybeans | $r \leq 0$ | 32.24* | 29.68 | 35.65 | 22.94* | 20.97 | 25.52 |
| | $r \leq 1$ | 9.30 | 15.41 | 20.04 | 5.67 | 14.07 | 18.63 |
| | $r \leq 2$ | 3.63 | 3.76 | 6.65 | 3.63 | 3.76 | 6.65 |
| Corn & Soybeans | $r \leq 0$ | 9.58 | 15.41 | 20.04 | 6.92 | 14.07 | 18.63 |
| | $r \leq 1$ | 2.66 | 3.76 | 6.65 | 2.66 | 3.76 | 6.65 |
| Eth & Soybeans | $r \leq 0$ | 14.68 | 15.41 | 20.04 | 10.15 | 14.07 | 18.63 |
| | $r \leq 1$ | 4.54 | 3.76 | 6.65 | 4.53 | 3.76 | 6.65 |
| Eth & Corn | $r \leq 0$ | 23.41** | 15.41 | 20.04 | 18.36** | 14.07 | 18.63 |
| | $r \leq 1$ | 5.05 | 3.76 | 6.65 | 5.05 | 3.76 | 6.65 |

7/ 30/06 – 4/7/09

Critical values from Osterwald-Lenum(1992)

No constant term in the cointegrating vector

Ethanol and corn futures contracts are available for differing contract months so in order to make a conformable price series, we take daily settlement prices of the contract month corresponding to the contract offered least frequently, which is corn. Corn contracts are available for December, March, May, July, and September. As the nearby contract comes to maturity, the series is rolled forward to the daily settlement price of the next closest contract. We do this on the third business day prior to the 25$^{th}$ calendar day of the month preceding the delivery month. The data series contains $T$ = 1012 observations.

*Cost of carry and futures prices*

Some care is required when using the nearby futures contract as a proxy for the spot price. The futures price at time $t$ of a commodity with maturity $T$ includes a cost of storing the commodity; an unadjusted series of nearby futures prices contains jumps on the dates when the series rolls forward to the next contract. This is especially problematic when testing for cointegration because the spurious co-movement of the series on the days the contracts roll forward make it more likely we reject the null hypothesis when it is in fact true.

Assuming the cost of carry is proportional to the commodity's price, the spot price at time $t$ is related to the futures price (or the expected future spot price) of a contract with maturity $T$ by the equation

(4) $$F_t^T = S_t e^{u(T-t)}$$

This means that a series of nearby futures contract prices is an inaccurate proxy for the spot price by the amount

(5) $$F_t^T - S_t = S_t\left(e^{u(T-t)} - 1\right),$$

which is largest when the contract is far from expiration and goes to zero as the contract approaches maturity.

Since there is no direct way to measure the cost of carry, we use an estimate denoted by $\tilde{u}$. On days the series is rolled forward to a new contract we measure the implied cost of carry per day as the solution to the equation

(6) $$F_{T_0}^{T_1} - F_{T_0}^{T_0} = S_{T_0}\left(e^{u(T_1-T_0)} - 1\right),$$

which assumes $F_{T_0}^{T_0} = S_{T_0}$. The solution is then $\tilde{u} = \frac{1}{T_1 - T_0}\mathbf{log}\left(1 + \frac{F_{T_0}^{T_1} - F_{T_0}^{T_0}}{F_{T_0}^{T_0}}\right)$.

We deflate (or inflate if the market happens to be inverted) each futures price by the amount $e^{-\tilde{u}(T-t)}$, and obtain a series of synthetic spot prices defined by, $\tilde{S}_t = F_t^T e^{-\tilde{u}(T-t)}$, which have been removed of the spurious co-movement generated by the series' construction.

*Pretesting and lag length selection*

We select a lag length of $k = 2$ based on Sims (1980) likelihood ratio statistic, the Akaike information criterion (AIC), and Schwartz Bayesian information criterion (SBIC) (see table 3 ). The Phillips-Perron (PP) $Z_\alpha$, modified PP $MZ_\alpha$, $MZ_t$, and Elliot, Rothenberg and Stocks' $P_T$ statistics all fail to reject the null hypothesis of a unit root in each individual data series (Elliott, et al., 1996, Perron and Ng, 1996, Phillips, 1987, Phillips and Perron, 1988). So we conclude the corn and ethanol price series are non-stationary.

Table 3: Lag length selection

| Statistic | LR | df | p-value | AIC | SBIC |
|---|---|---|---|---|---|
| 1 | 64.29 | 9 | 0.00 | -15.34 | -15.17* |
| 2 | 16.18 | 9 | 0.06* | -15.36* | -15.05 |
| 3 | 17.86 | 9 | 0.04 | -15.32 | -14.89 |
| 4 | 15.66 | 9 | 0.07 | -15.26 | -14.72 |

*Seasonality in the data*

Agricultural commodity prices are expected to contain seasonality (Brennan, 1958). The spot price verses expected future spot prices must be such that sufficient amount of the crop is stored for consumption throughout the year.

To correct for seasonality in the data we include sinusoidal regressors with periods of 1 year, 6 months, and 3 months initially. We fit the data to a standard VAR to determine which of the seasonal regressors provide explanatory power to the data, and the *t*-probabilities from each regressor in each equation are included in table 4.

Table 4: Seasonal regressors

| Period ( fraction of yr) | 1 | 0.50 | 0.25 |
|---|---|---|---|
| Ethnaol | 0.04 | 0.05 | 0.46 |
| Corn | 0.13 | 0.35 | 0.89 |

**p*-values of the *t*-statistic on the sinusoidal regressors in the given price equatons of a VAR

The only sinusoidal regressor significant at the 5% level is that with period of 1 year in the corn price equation. We include only this one in the price equations in the model analyzed below.

**Preliminary (Asymptotic) Results**

We conduct a preliminary analysis of the prices of corn and ethanol. In table 2 observe that the null hypothesis of no cointegration is rejected at the 1% level by the trace statistic and at the 5% level by the maximal eigenvalue statistic. Table 5 contains the results of fitting the corn and ethanol data to a vector error correction model. Neither variable Granger causes the other in the short run, as illustrated by the short run Granger probabilities. The error correction term is significant in the ethanol price equation since the p-values on the $\alpha$ parameter in the ethanol equation is less than 0.01, but the error correction term is not significant in the corn price equation. This implies when the variables are shocked away from their equilibrium relationship, the price of ethanol adjusts two bring the two series back into equilibrium.

Table 5: Error correction results and Granger causality – corn and ethanol

| Short run | | *corn* | *ethanol* |
|---|---|---|---|
| Granger | *corn* | 0.11 | 0.80 |
| probabilities | *ethanol* | 0.13 | 0.63 |
| $\beta$ | | -0.19 | 0.98 |
| *p*-value | | 0.00 | 0.00 |
| $\alpha$ | | -0.02 | -0.03 |
| *p*-value | | 0.39 | 0.00 |

We caution against a strict interpretation of these analyses of the corn and ethanol price relationship. The Johansen tests we employ, while currently state of the art for detecting equilibrium relationships, have severe limitations in small samples; we explore these problems in more detail in the sections that follow.

**Small Sample Issues with Johansen's Cointegration tests**

Since Johansen's trace and maximum eigenvalue tests of cointegrating rank are *asymptotic* likelihood ratio tests they have undesirable properties in small samples. It is known that that the tests suffer from size distortion and low power in small samples, especially when the error correction model produces residuals are nearly I(1). Several Monte Carlo studies have been published outlining the severity of these issues.

Cheung and Lai (1993) determine the finite sample sizes of the Johansen tests and quantify the finite sample critical values using response surface analysis. They conclude the Johansen tests are biased toward rejecting a null of no cointegration too often in finite samples compared to the asymptotic distribution of the test statistics. Further, they conclude that the bias worsens as the dimension of the system or length of the lag structure increases.

Toda (1995) performs an independent study of the finite sample performance of the Johansen tests and determines that with 100 observations the simulated distribution of the asymptotic test statistic under the null is fairly good. However, 100 observations are not enough to determine the true cointegrating rank under the alternative if the one or more of the stationary roots of the process is nearly 1. Unlike Cheung and Lai, Toda asserts that this leads to underestimation of the cointegrating rank because of the nature of sequential testing inherent in the Johansen procedure. Further, he finds the test's performance is affected by initial values of the stationary component of the process. Toda concludes that one needs 300 observations for the test to perform well uniformly over the range of finite sample scenarios he considers.

Alternative to determining the critical values of the actual finite sample distribution, small sample corrections to the test statistics or critical values have been proposed. Johansen (2002) proposes a correction factor that depends on parameters of the error correction model as well as the sample size. However, the correction is fairly complicated to apply (the components of the correction which depend only on functionals of a random walk are simulated and described in (Johansen, et al., 2005)); it is not clear that one cannot obtain better estimates of the small sample critical values from simulating the small sample distributions directly. After all, the correction of the statistics developed by Johansen (2002) requires estimating the parameters of the data, just as is required to simulate the small sample distribution correctly.

Ahn and Reinsel (1990) and Reimers (1992) develop a correction that is a simple function of sample size, system dimension and lag order. However, as part of their Monte Carlo analysis Cheung and Lai (1993) conclude that the Ahn-Reinsel method does not yield unbiased estimates of the finite sample critical values.

Given these reservations about the propriety of the small sample corrections, it seems the better approach is to obtain small sample critical values from a simulated finite sample distribution. However, the Monte Carlo experiments performed in analysis preceding this article do not draw a connection between their simulation experiments and the type of data series they intend to mimic. Are they simulated daily, weekly, monthly, quarterly, or yearly data? This would be characterized by the variance of innovations in the process, but the Monte Carlo studies considered above usually set the variance in a convenient range (such as $\sigma \in [0.25, 1]$); no explicit link between the data generating process (DGP) and a specific type of data series is made.

However, in many time series studied by applied researchers this distinction is important. While 400 observations of yearly data is a longer series than social scientists ever enjoy, 400 observations of daily data contain scarcely more than a year's worth of information. For price series like corn or other agricultural commodities, one year's worth of data only includes one crop yield realization. The day-to-day variation of these is determined mostly by the market's incentive to store, and a year is not likely to be nearly enough time to capture the variation required in the data to test for cointegrating relationships between variables in this context.

Further, the evidence is mixed about whether temporal aggregation of the data helps or hinders the power of cointegration tests to detect equilibrium relationships in small samples (see (Hooker, 1993, Lahiri and Mamingi, 1995, Otero and Smith, 2000, Shiller and Perron, 1985) for discussion). The level of temporal aggregation will certainly influence the distribution of the test statistics in a small sample, and whether it increases or decreases the tests' power will depend on the nature of the DGP.

Therefore, it seems appropriate to investigate how much data is required before the tests statistic appears to approach asymptotic behavior under a range of assumptions about the DGP. Further, it seems worthwhile to provide small sample critical values that are associated more directly with a particular type of temporally aggregated data set, and for varying lengths of data available.

In the next section we perform a small Monte Carlo study that provides critical values for Johansen's statistics on cointegrating rank. The experiment is tailored explicitly to 'daily data'. We provide both critical values of the distribution of the statistic under the null hypothesis as well as a small study of the power of the test statistic under the alternative hypothesis.

**A Monte Carlo Study**

The DGP we use closely resembles that used in prior Monte Carlo studies done by Banerjee et al. (1986) and Haug (1996). In this study, we restrict our attention to a bivariate system, since our original interest was in exploring the nature of a possible equilibrium relationship between corn and ethanol prices. The data generating process is

(7) $y_t - x_t = v_t, \quad v_t = \rho v_{t-1} + w_t,$

(8) $y_t + x_t = \psi_t, \quad \psi_t = \psi_{t-1} + r_t, \quad r_t = \varphi_t + \theta\varphi_{t-1}$

(9) $\begin{bmatrix} w_t \\ \varphi_t \end{bmatrix} \overset{iid}{=} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \eta \\ \eta & \sigma_2^2 \end{bmatrix}\right).$

With this DGP a moving average component in the error terms exists when $\theta$ is non-zero. For $\rho = 1$ the data is generated under the null hypothesis of no cointegration, while a value of $|\rho| < 1$ corresponds to the alternative hypothesis that the two series are cointegrated. We choose values of $\eta$ and $\sigma$ so that the covariance matrix of $[y_t \quad x_t]'$ matches levels typical for the two series of interest, in this case the corn and ethanol prices, and run the simulations for each parameter scenario $(\rho, \theta)$, where $\rho \in \{0.85, 0.90, 1\}$, and $\theta \in \{-0.80, 0, 0.80\}$. We follow Haug (1996) in the choice of these parameter values, which allows us to illustrate the effect of a moving average component in the error term on the size distortion and power of the Johansen tests.

Deriving the variances and covariance of $[y_t \quad x_t]'$ implied by the data generating process in (7) and (8), one finds that the covariance matrix of $[w_t \quad \varphi_t]'$ relates to the covariance of $[y_t \quad x_t]'$ by the equations

(10) $\sigma_1^2 = 2\left(\sigma_x^2 - \sigma_{xy}\right)$

(11) $\sigma_2^2 = 2\left(\sigma_y^2 + \sigma_{xy}\right)/\left(1+\theta^2\right)$

(12) $\eta = \sigma_y^2 - \sigma_x^2$

Where $\Sigma_{yx} = \begin{bmatrix} \sigma_y^2 & \sigma_{yx} \\ \sigma_{yx} & \sigma_x^2 \end{bmatrix}$ is the covariance matrix of the $\left[\Delta y_t \quad \Delta x_t\right]'$ series. These equations allow us to tailor the Monte Carlo experiment to mimic data of any temporal aggregation level.

We match the covariance structure of the simulated data series to that of the first difference of logged corn and ethanol futures prices in the analysis above. So that $\Sigma_{y,x} = \begin{bmatrix} 0.000406 & 0.000172 \\ 0.000172 & 0.000435 \end{bmatrix}$.

We perform the simulation using 100,000 replications, and for sample lengths varying from one month to 100 years. Table 6 contains the finite sample critical values for different assumptions on the error term (values of $\theta$), lag specification of the VECM, and for varying data series lengths. When the moving average component of the DGP is zero or positive, ($\theta \geq 0$), the small sample size distortion disappears with little more than six months of data, and the statistics have surprisingly small size distortion with as little as three months of data.

On the other hand when data is generated with a negative moving average component in the errors, the size distortion is severe. A lag length of $k = 2$ is not enough to purge the system of autocorrelation, and we see size distortion persist even after the statistic seems to have converged. Increasing the lag length is helpful in reducing the size distortion, but this costly in terms of power of the test statistic under the alternative.

Table 6: Finite sample critical values for Johansen's cointegration tests – "Daily data"

| | Sample length $T^*$[29] | Empirical Size[28] | Trace | | Max | |
|---|---|---|---|---|---|---|
| | | | 95% c.v. | 99% c.v. | 95% c.v. | 99% c.v. |
| $\theta = 0$ | 1 mo | 0.19 | 24.17 | 30.97 | 20.22 | 26.44 |
| $k = 2$ | 3 mos | 0.07 | 19.36 | 24.45 | 16.17 | 20.68 |
| | 6 mos | 0.06 | 18.71 | 23.80 | 15.59 | 20.02 |
| | 2 yrs | 0.05 | 18.32 | 23.09 | 15.28 | 19.56 |
| | 4 yrs | 0.05 | 18.10 | 22.88 | 15.01 | 19.30 |
| | 10 yrs | 0.05 | 18.11 | 22.80 | 15.01 | 19.33 |
| | 20 yrs | 0.05 | 18.13 | 22.88 | 15.08 | 19.29 |
| | 100 yrs | 0.05 | 18.12 | 22.91 | 15.03 | 19.34 |
| | $\infty$ | | 18.17 | 23.46 | 16.87 | 21.47 |
| $\theta = 0.8$ | 1 mo | 0.20 | 25.33 | 32.56 | 21.18 | 27.56 |
| $k = 2$ | 3 mos | 0.07 | 19.41 | 24.94 | 16.18 | 21.15 |
| | 6 mos | 0.06 | 18.54 | 23.58 | 15.44 | 20.12 |
| | 2 yrs | 0.05 | 17.90 | 22.81 | 14.92 | 19.44 |
| | 4 yrs | 0.05 | 17.85 | 22.67 | 14.84 | 19.25 |
| | 10 yrs | 0.04 | 17.74 | 22.57 | 14.78 | 19.28 |
| | 20 yrs | 0.04 | 17.79 | 22.69 | 14.83 | 19.26 |
| | 100 yrs | 0.04 | 17.75 | 22.65 | 14.79 | 19.19 |
| | $\infty$ | | 18.17 | 23.46 | 16.87 | 21.47 |
| $\theta = -0.8$ | 1 mo | 0.22 | 25.23 | 31.68 | 21.12 | 27.11 |
| $k = 2$ | 3 mos | 0.40 | 28.41 | 34.36 | 24.86 | 30.52 |
| | 6 mos | 0.60 | 35.19 | 42.75 | 31.98 | 39.26 |
| | 2 yrs | 0.73 | 52.47 | 66.73 | 49.67 | 63.73 |
| | 4 yrs | 0.74 | 59.17 | 77.65 | 56.31 | 74.82 |
| | 10 yrs | 0.76 | 64.23 | 86.60 | 61.41 | 83.58 |
| | 20 yrs | 0.76 | 66.54 | 91.73 | 63.74 | 88.39 |
| | 100 yrs | 0.76 | 67.94 | 94.73 | 64.98 | 92.28 |
| | 200 yrs | 0.76 | 68.12 | 94.04 | 65.28 | 91.36 |
| | $\infty$ | | 18.17 | 23.46 | 16.87 | 21.47 |

$n = 2$, $k$ is the number of lags used to perform the tests, 100,000 replications
Asymptotic critical values from Osterwald-Lenum (1992)

[28] If the 95% asymptotic critical value is used.

[29] Based on the assumption that there are approximately 250 trading days in one year. I.e., if T = 4 years this corresponds to *nobs* = 1000 observations.

Table 6 (cont.): Finite sample critical values for Johansen's cointegration tests – "Daily data"

| | | | | | | |
|---|---|---|---|---|---|---|
| $\theta = -0.8$ | 1 mo | 0.48 | 37.32 | 49.05 | 32.07 | 42.84 |
| $k$ = 4 | 3 mos | 0.18 | 23.44 | 29.07 | 19.85 | 25.06 |
| | 6 mos | 0.24 | 25.44 | 31.39 | 21.98 | 27.79 |
| | 2 yrs | 0.34 | 30.58 | 39.28 | 27.43 | 36.06 |
| | 4 yrs | 0.37 | 32.33 | 42.20 | 29.34 | 39.02 |
| | 10 yrs | 0.38 | 33.67 | 44.54 | 30.69 | 41.42 |
| | 20 yrs | 0.39 | 34.26 | 45.42 | 31.25 | 42.50 |
| | 100 yrs | 0.39 | 34.56 | 46.25 | 31.67 | 43.19 |
| | 200 yrs | 0.39 | 34.47 | 46.24 | 31.52 | 43.56 |
| | $\infty$ | | 18.17 | 23.46 | 16.87 | 21.47 |
| $\theta = -0.8$ | 1 mo | 0.81 | 71.82 | 98.12 | 63.68 | 89.80 |
| $k$ = 5 | 3 mos | 0.17 | 23.07 | 28.93 | 19.35 | 24.73 |
| | 6 mos | 0.18 | 23.45 | 29.28 | 20.00 | 25.70 |
| | 2 yrs | 0.24 | 26.41 | 33.60 | 23.18 | 30.30 |
| | 4 yrs | 0.25 | 27.27 | 35.62 | 24.16 | 32.27 |
| | 10 yrs | 0.26 | 27.93 | 36.59 | 24.84 | 33.32 |
| | 20 yrs | 0.27 | 28.31 | 37.22 | 25.22 | 34.02 |
| | 100 yrs | 0.27 | 28.49 | 37.60 | 25.51 | 34.54 |
| | 200 yrs | 0.27 | 28.41 | 37.70 | 25.46 | 34.60 |
| | $\infty$ | | | | | |

$n$ = 2, $k$ is the number of lags used to perform the tests, 100,000 replications

Table 7 contains the results of a small power study of the test statistic under the alternative hypothesis that $|\rho| < 1$. The reported values are *size adjusted* powers, so it is the probability of rejecting the false null hypothesis using the appropriate small sample critical values. Performing the simulations for $\rho = 0.85$, $0.90$, and $0.95$ demonstrates how the test loses ability to discern a cointegrating relationship from a persistent alternative.

Table 7: Size-adjusted finite sample power of Johansen's trace cointegration test under the alternative $|\rho|<1$ – "Daily data"

| | Sample length $T$ | $\theta=0$ | $\theta=0.8$ | $\theta=-0.8$ |
|---|---|---|---|---|
| $\rho=0.85$ | 1 mo | 0.05 | 0.05 | 0.05 |
| | 3 mos | 0.10 | 0.10 | 0.08 |
| | 6 mos | 0.28 | 0.27 | 0.12 |
| | 2 yrs | 1 | 1 | 0.49 |
| | 4 yrs | 1 | 1 | 0.99 |
| | 10 yrs | 1 | 1 | 1 |
| | 20 yrs | 1 | 1 | 1 |
| | 100 yrs | 1 | 1 | 1 |
| $\rho=0.90$ | 1 mo | 0.05 | 0.05 | 0.05 |
| | 3 mos | 0.08 | 0.07 | 0.06 |
| | 6 mos | 0.15 | 0.15 | 0.09 |
| | 2 yrs | 0.99 | 0.99 | 0.22 |
| | 4 yrs | 1 | 1 | 0.70 |
| | 10 yrs | 1 | 1 | 1 |
| | 20 yrs | 1 | 1 | 1 |
| | 100 yrs | 1 | 1 | 1 |
| $\rho=0.95$ | 1 mo | 0.05 | 0.05 | 0.05 |
| | 3 mos | 0.06 | 0.06 | 0.06 |
| | 6 mos | 0.08 | 0.08 | 0.0 |
| | 2 yrs | 0.55 | 0.56 | 0.09 |
| | 4 yrs | 0.99 | 0.99 | 0.17 |
| | 10 yrs | 1 | 1 | 0.90 |
| | 20 yrs | 1 | 1 | 1 |
| | 100 yrs | 1 | 1 | 1 |

$n$ = 2, 100,000 replications

The trace statistic performs relatively well in terms of power also when $\theta \geq 0$. When the DGP has $\rho=0.85$ the power of the test is relatively good for sample length as small as six months, and has very high power for samples of 2 years or longer. The power becomes weaker when the DGP is closer to a the null hypothesis of $\rho=1$ with $\rho=0.90$ and $0.95$. In these cases, two years of data are required before the test statistic has reasonable power; however, the power is very good for data series of four years or more.

The results of this Monte Carlo study are consistent with the findings of earlier research on the topic. The Johansen trace statistic suffers from size distortion in small samples, the severity of which depends on the nature of the error term in the data generating process. A negative moving average component to the error terms results in severe size distortions, while a positive or no moving average component in the error term results in fairly good properties of the statistic under the null hypothesis. The power of the statistic also is lower when the data generating process has a moving average component, but is more severely affected by the persistence of the data series measured by the distance of $\rho$ from unity.

With respect to our original question of interest, the Monte Carlo study teaches us that the analysis conducted above on the cointegration of corn and ethanol prices, while instructive, should be interpreted with a healthy measure of caution. The data set contained daily data spanning less than three years. If the data generating process is either 1) highly persistent or 2) contains a negative moving average component or both, we can expect the statistic to perform poorly regardless of whether the null or the alternative hypothesis happens to be true. The test rejects the null hypothesis of no cointegration between ethanol and corn under the assumption that the DGP does not contain a negative moving average component in the errors, since the trace test statistic is 23.41 compared to the 99% critical values of approximately 23 and 22.7 respectively. But the test fails to reject the null hypothesis under the assumption that the DGP does contain a negative moving average component in the errors, since the 99% and 95% critical values are decidedly above the test statistic.

**Conclusion**

In this paper we explored the possibility of an equilibrium relationship between corn and ethanol prices. Using the equilibrium condition that in the long-run zero economic profit should be earned by the ethanol industry, we posit that futures contracts far from maturity should have prices related according to this breakeven relationship.

We demonstrated how this long run relationship could be transmitted to spot prices through intertemporal arbitrage. In the cost-of-carry model intertemporal arbitrage governs the relationship between the futures and spot price of a storable commodity. This could impose a long run relationship that is maintained by the price of two or more futures contracts on spot prices as well.

In order to test this theory we use the statistical methods of Johansen (1991). We find evidence that there exists an equilibrium relationship between corn and ethanol prices, but at the same time caution against a strong interpretation of the results.

We discussed the small sample properties of the Johansen trace and maximal eigenvalue statistics. The probability distributions of the statistics are only valid asymptotically; the statistics have been shown to have poor finite sample properties in previous Monte Carlo studies.

We tailor a Monte Carlo study with a data generating process that mimics the series we study, in order that we may determine how much data of this particular kind of series is required for good performance of the statistic. We found that our data set would not be sufficient to expect good performance of the statistic if a negative moving average is present in the error term, and obtained mixed results under different assumptions about the DGP.

## References

Ahn, S. K., and G. C. Reinsel. "Estimation for partially nonstationary multivariate autoregressive models." *Journal of the American statistical association* (1990): 813-823.

Baker, M. L., D. J. Hayes, and B. A. Babcock. "Crop-Based Biofuel Production under Acreage Constraints and Uncertainty." *CARD Working Paper Series* 08-WP 460(2008).

Banerjee, A., et al. "Exploring equilibrium relationships in econometrics through static models: some Monte Carlo evidence." *Oxford Bulletin of Economics and statistics* 48, no. 3(1986): 253-277.

Brennan, M. J. "The supply of storage." *The American Economic Review* (1958): 50-72.

Cheung, Y. W., and K. S. Lai. "Finite-sample sizes of Johansen's likelihood ratio tests for cointegration." *Oxford Bulletin of Economics and statistics* 55, no. 3(1993): 313-328.

Elliott, G., T. J. Rothenberg, and J. H. Stock. "Efficient Tests for an Autoregressive Unit Root." *Econometrica* 64, no. 4(1996): 813-836.

Engle, R. F., and C. W. J. Granger. "Co-integration and error correction: representation, estimation, and testing." *Econometrica: Journal of the Econometric Society* (1987): 251-276.

Fama, E. F. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance* 25, no. 2(1970): 383-417.

Granger, C. W. J. "Some Properties of Time Series Data and Their use in Econometric Model Specification." *Journal of Econometrics* 16(1981): 121-130.

Granger, C. W. J., and A. A. Weiss. "Time series analysis of error correction models." *Essays in Econometrics: Collected Papers of Clive WJ Granger* (2001): 129.

Hamilton, J. D. *Time series analysis*: Princeton university press, 1994.

Haug, A. A. "Tests for cointegration: a Monte Carlo comparison." *Journal of Econometrics* 71, no. 1(1996): 89-116.

Hooker, M. A. "Testing for cointegration : Power versus frequency of observation." *Economics Letters* 41, no. 4(1993): 359-362.

Hull, J. *Fundamentals of futures and options markets*: Prentice Hall, 2002.

Johansen, S. "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models." *Econometrica: Journal of the Econometric Society* (1991): 1551-1580.

Johansen, S. "A small sample correction for the test of cointegrating rank in the vector autoregressive model." *Econometrica* (2002): 1929-1961.

Johansen, S. "Statistical analysis of cointegration vectors." *Journal of Economic Dynamics and Control* 12, no. 2-3(1988): 231-254.

Johansen, S., H. Hansen, and S. Fachin. "A simulation study of some functionals of random walk." *University of Copenhagen. Preprint* (2005).

Johansen, S., and K. Juselius. "Maximum likelihood estimation and inference on cointegration with applications to the demand for money." *Oxford Bulletin of Economics and statistics* 52, no. 2(1990): 169-210.

Kruse, J., et al. "Economic Impacts of Not Extending Biofuel Subsidies." *AgBioForum* 10, no. 2(2007): 94-103.

Lahiri, K., and N. Mamingi. "Testing for cointegration: Power versus frequency of observation—another view." *Economics Letters* 49, no. 2(1995): 121-124.

MacKinnon, J. G., A. A. Haug, and L. Michelis. "Numerical distribution functions of likelihood ratio tests for cointegration." *Journal of Applied Econometrics* 14, no. 5(1999): 563-577.

Malkiel, B. G. "The Efficient Market Hypothesis and Its Critics." *Journal of Economic Perspectives* 17, no. 1(2003): 59-82.

Osterwald-Lenum, M. "A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics." *Oxford Bulletin of Economics and statistics* 54, no. 3(1992): 461-472.

Otero, J., and J. Smith. "Testing for cointegration: power versus frequency of observation—further Monte Carlo results." *Economics Letters* 67, no. 1(2000): 5-9.

Perron, P., and S. Ng. "Useful modifications to some unit root tests with dependent errors and their local asymptotic properties." *The Review of Economic Studies* (1996): 435-463.

Phillips, P. C. B. "Time series regression with a unit root." *Econometrica: Journal of the Econometric Society* (1987): 277-301.

Phillips, P. C. B., and S. Ouliaris. "Asymptotic properties of residual based tests for cointegration." *Econometrica: Journal of the Econometric Society* (1990): 165-193.

Phillips, P. C. B., and P. Perron. "Testing for a unit root in time series regression." *Biometrika* 75, no. 2(1988): 335-346.

Reimers, H. E. "Comparisons of tests for multivariate cointegration." *Statistical papers* 33, no. 1(1992): 335-359.

Shapouri, H., and P. Gallagher. "USDA's 2002 Ethanol Cost-of-Production Survey." United States Department of Agriculture.

Shiller, R. J., and P. Perron. "Power versus frequency of observation." *Economics Letters* 18(1985): 381-386.

Shurson, G. C., et al. (2003) Value and use of "new generation" distiller's dried grains with solubles in swine diets. Lexington, KY.

Sims, C. A. "Macroeconomics and reality." *Econometrica: Journal of the Econometric Society* (1980): 1-48.

Thompson, W., S. Meyer, and P. Westhoff. "How Does Petroleum Price and Corn Yield Volatility Affect Ethanol Markets With and Without an Ethanol Use Mandate?" *Energy Policy* 37, no. 2009(2009): 745-749.

Toda, H. Y. "Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions." *Econometric Theory* (1995): 1015-1032.

Tokgoz, S., et al. "Bottlenecks, Drought, and Oil Price Spikes: Impact on U.S. Ethanol and Agricultural Sectors." *Review of Agricultural Economics* 30, no. 4(2008): 604-622.

Tokgoz, S., et al. "Emerging Biofuels: Outlook of Effects on U.S. Grain, Oilseed, and Livestock Markets." *Center for Agricultural and Rural Development* 07-SR 101(2007).

# CHAPTER 5: GENERAL CONCLUSTIONS

This dissertation is a collection of three papers. Each paper dealt with a particular aspect of the relationship between energy and agriculture. The objective of the first paper was to create a model that would enhance informed policy decisions. A forward-looking stochastic model captured the effect of uncertainty in crude oil prices and commodity yields on biofuel industry development. Acreage limitations on feedstocks such as corn, soybeans, and switchgrass were shown to create competition for acreage among the crops, and lead to high commodity prices. Investors in the model were rational in the sense that they engaged in biofuel production only if returns exceeded what they could expect to earn from alternative investments.

The Energy Independence and Security Act of 2007 mandated the use of 36 billion gallons of biofuels by 2022 with significant requirements for cellulosic biofuel and biodiesel production. In the model, the price wedge created by mandated biofuel production at these levels was \$2.50 per gallon for biodiesel and \$1.07 per gallon for cellulosic biofuel. Long-run commodity prices were high in our simulation, with corn at \$7.38 per bushel and soybeans at \$19.57 per bushel.

The second paper developed a model of the corn, soybean, and wheat markets to calculate welfare effects of increased biofuel production in the United States. Demand was disaggregated into livestock feed, food, energy. Allowing for uncertainty in crop yields permitted the valuation of farm deficiency payments as options. Incorporating soybean and wheat markets captured indirect welfare effects through an equilibrium price increase. Net welfare loss ranged from \$200 million to \$750 million depending on the size of biofuel increase. Consumers made a sizable transfer to farmers. The sign of the net costs to taxpayers depended on the size of the biofuel industry.

In the third paper, the nature of the relationship between corn and ethanol prices was explored. Economic fundamentals should require that the price of corn and ethanol maintain a long run equilibrium relationship. The relationship is driven by a long run condition that says entry and exit in the industry will occur maintaining no sustained profits or losses for the industry. Both ethanol producers and traditional users of corn have a stake in the behavior of these markets, and their profitability will rely on their ability to determine accurately this relationship. I tested for cointegration of these price series and find evidence that corn and ethanol prices are indeed maintaining an equilibrium relationship.

Statistical cointegration tests are known to have problems in small samples. This is a potential issue in interpreting the results mentioned above because ethanol production only recently constituted a significant portion of the corn crop. With only a few years of the most recent data for which we suspect that an equilibrium relationship existed, the small sample properties of cointegration tests are important. I conducted a Monte Carlo study that was tailored to mimic the actual data set of corn and ethanol prices. I find that the corn and ethanol price series are not long enough to rely on the asymptotic properties of the cointegration statistics, and therefore one should use small sample critical values in this kind of analysis.

This dissertation demonstrates some of the important ways energy markets and agricultural markets are intertwined because of biofuel production. The use of agricultural commodities as an input into the production of energy fundamentally changes the way these markets interact. There are still many questions about how these two seemingly distinct sectors are related, but this dissertation fills some of these gaps. The three papers collected here consider, in turn, the tax credit or subsidy required to maintain a biofuel industry at EISA 2007 mandated levels, the welfare implications of biofuel production increase, and they way corn and ethanol price behavior may has been altered due to biofuel production.