THE IMPACT OF BIOFUEL PRODUCTION ON ENERGY AND AGRICULTURAL PRICE RELATIONSHIPS

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Doug Hermanson, B.A. Economics, B.A. Psychology

****

The Ohio State University

2008

Thesis Committee:

Dr. Matthew C. Roberts, Adviser

Dr. Stanley R. Thompson

Approved by

Graduate Program, Agricultural, Environmental, and Development Economics
ABSTRACT

The production of biofuels has increased significantly within the United States from 2005 to 2008. Soybean oil and corn have been the two major feedstocks in the production of biodiesel and ethanol in the United States. The use of soybean oil in the production of biodiesel and corn in the production of ethanol potentially transforms two independent relationships between both corn and gasoline and soybean oil and diesel into a substitute or complementary relationship. Through cointegration testing and the calculation of a confidence interval of correlation coefficients the hypothesis was tested that price relationships between corn and gasoline and also soybean oil and diesel have changed due to increased production of biofuels. The results of the cointegration tests were inconclusive in testing the hypothesis due to stationarity in several of the price series. The calculation of a bootstrap confidence interval of correlation coefficients from 1980 to 2008 showed that correlation between soybean oil and diesel prices and corn and gasoline prices from 2005 to 2008 was greater than the mean correlation over the entire 28 year period. The results provide evidence that increased biofuel production has led to a new economic relationship between agricultural and energy commodities. As a result food, agriculture and livestock industries are subject to a new type of price risk that is influenced by energy markets, not just agriculture markets.
ACKNOWLEDGMENTS

I wish to thank Dr. Matthew C. Roberts (my advisor) for his intellectual support and encouragement in the process of researching and writing this thesis. I would also like to thank Dr. Stanley R. Thompson for his time and support through his participation in my thesis committee.
VITA

2007 … Research Assistant to Dr. Matthew C. Roberts, Department of
Agricultural, Environmental, and Development Economics, The Ohio
State University
2007-present  Research Assistant, Retail Market Analysis Program, Department of
Agricultural, Environmental, and Development Economics, The Ohio
State University

FIELDS OF STUDY

Major Field: Agricultural, Environmental, and Development Economics
Specialization: Financial Economics and Risk Management
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>II</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>III</td>
</tr>
<tr>
<td>Vita</td>
<td>IV</td>
</tr>
<tr>
<td>List of Tables</td>
<td>VI</td>
</tr>
<tr>
<td>List of Figures</td>
<td>VII</td>
</tr>
<tr>
<td>Chapters:</td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1. Production and Consumption of Biofuels</td>
<td>2</td>
</tr>
<tr>
<td>2. Biofuel, Agriculture and Energy Price Relationships</td>
<td>11</td>
</tr>
<tr>
<td>2.1 Substitute and Complementary Relationships</td>
<td>11</td>
</tr>
<tr>
<td>2.2 Derived Demand</td>
<td>17</td>
</tr>
<tr>
<td>3. Methodology and Results</td>
<td>31</td>
</tr>
<tr>
<td>3.1 Data</td>
<td>31</td>
</tr>
<tr>
<td>3.2 Cointegration Review and Methodology</td>
<td>31</td>
</tr>
<tr>
<td>3.3 Cointegration Estimation and Results</td>
<td>36</td>
</tr>
<tr>
<td>3.4 Correlation Review and Methodology</td>
<td>39</td>
</tr>
<tr>
<td>3.5 Correlation Estimation and Results</td>
<td>41</td>
</tr>
<tr>
<td>4. Conclusion</td>
<td>49</td>
</tr>
<tr>
<td>Bibliography</td>
<td>51</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Augmented Dickey Fuller Unit Root Tests</td>
<td>37</td>
</tr>
<tr>
<td>2. Cointegration Test</td>
<td>38</td>
</tr>
<tr>
<td>3. Bootstrap Confidence Intervals of Correlation Coefficients</td>
<td>43</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Gasoline and Ethanol Differential without the Tax Credit</td>
<td>3</td>
</tr>
<tr>
<td>2.</td>
<td>Diesel and Biodiesel Differential without the Tax Credit</td>
<td>4</td>
</tr>
<tr>
<td>3.</td>
<td>Perfect Complements</td>
<td>11</td>
</tr>
<tr>
<td>4.</td>
<td>Perfect Substitutes</td>
<td>11</td>
</tr>
<tr>
<td>5.</td>
<td>Perfect Substitutes with Budget Lines</td>
<td>13</td>
</tr>
<tr>
<td>6.</td>
<td>Perfect Complements with Budget Lines</td>
<td>13</td>
</tr>
<tr>
<td>7.</td>
<td>U.S Gasoline and Ethanol Markets</td>
<td>21</td>
</tr>
<tr>
<td>8.</td>
<td>Alternative Gasoline and Ethanol Markets</td>
<td>21</td>
</tr>
<tr>
<td>9.</td>
<td>Gasoline and Ethanol Flex Differential with the Tax Credit</td>
<td>25</td>
</tr>
<tr>
<td>10.</td>
<td>Gasoline and Ethanol Mix Differential with the Tax Credit</td>
<td>26</td>
</tr>
<tr>
<td>11.</td>
<td>U.S Gasoline and Ethanol Markets with Inelastic Ethanol Supply</td>
<td>28</td>
</tr>
<tr>
<td>13.</td>
<td>Heating Oil and Soybean Oil Correlation 1980-2008</td>
<td>44</td>
</tr>
<tr>
<td>14.</td>
<td>Heating Oil and Soybean Oil Correlation 2005-2008</td>
<td>46</td>
</tr>
<tr>
<td>15.</td>
<td>Gasoline and Corn Correlation 2005-2008</td>
<td>47</td>
</tr>
</tbody>
</table>
INTRODUCTION

The production of biofuels has increased significantly in the last seven years within the United States. Soybean oil and corn have been the two major feedstocks in the production of biodiesel and ethanol in the United States. The use of soybean oil in the production of biodiesel and corn in the production of ethanol potentially transforms two independent relationships between both corn and gasoline and soybean oil and diesel into a substitute or complementary relationship. The existence of a substitute or complementary relationship may result in the cointegration or correlation of the prices of soybean oil and corn with the prices of diesel and gasoline, respectively.

The importance of this occurrence would be that food, agriculture and livestock markets would be subject to not only price risks associated with agriculture markets, but also price risks associated with energy markets. Evidence of the cointegration of or evidence of significant correlation between corn and gasoline and soybean oil and diesel provides evidence of a new economic relationship between the commodities and a new type of price risk for food, agriculture and livestock industries. In the presence of this new type of price risk, hedging strategies in food, agriculture and livestock industries that account for price changes in energy markets may need to be further researched and utilized.
CHAPTER 1

PRODUCTION AND CONSUMPTION OF BIOFUELS

According to the Renewable Fuels Association (RFA) ethanol production increased from 1.6 billion gallons in 2000 to 6.5 billion gallons in 2007 (2008b). Biodiesel production has increased from under 20 million gallons in 2000 to roughly 400 million gallons in 2007 (Bange).

The increased production of biofuels within the United States is partly attributed to an increase in the price of crude oil and the price of crude oil based products. According to the Energy Information Administration (EIA) crude oil prices have increased from $27 per barrel in January of 2000 to $91 per barrel in December of 2007. Over the same period the United States has seen a 137% and 147% rise in gasoline and diesel prices respectively (2008a; 2008b).
Notes: (a) Weekly cash prices. (b) Differential = gasoline – ethanol. (c) Data source: Commodity Research Bureau

**Figure 1: Gasoline and ethanol differential without the tax credit**

In addition to increased prices of gasoline and diesel, United States government policies have also contributed to the increased production of biofuels. A look at monthly price differentials between gasoline and ethanol in Figure (1) and monthly price differentials between diesel and biodiesel in Figure (2) indicate that in the past ethanol or biodiesel would not have been very competitive with gasoline or diesel without the existence of certain government policies. Kruse and de Gorter have both suggested that ethanol would not have been very competitive with gasoline in the recent past without certain government policies.
Figure 2: Diesel and biodiesel differential without the tax credit

There are several United States government policies that have provided additional incentives for the consumption and production of ethanol. Foreign competition with domestic ethanol markets has been reduced through a 2.5% ad volorem tariff and an added duty of 54 cents per gallon from certain source countries including Brazil, a major producer of sugar based ethanol. Caribbean countries, such as Costa Rica and Jamaica under certain conditions can import ethanol in to the United States duty free (Yacobucci).

The Jobs Creation Act in 2004 and the Energy Policy Act in 2005 established mandates for the consumption of ethanol and a blender’s tax credit for biodiesel. The act also kept in place a blender’s tax credit for ethanol that has been in place since the early 1980’s.

Prior to 2005 the ethanol blender’s tax credit ranged from 60 cents per gallon in the 80’s to 53 cents per gallon in early 2000 (de Gorter). Since 2005, a 51 cent per gallon
blender’s tax credit for ethanol and a 1 dollar per gallon blender’s tax credit for pure biodiesel have been established (Koplow). The blender’s tax credit for ethanol and biodiesel acts much like a consumption subsidy for the two biofuels. The Internal Revenue Service states that the credit is allowable to the person that produces the mixture of ethanol and gasoline or biodiesel and diesel for sale or use. If biodiesel is sold as pure biodiesel and not a mixture then the credit is allowable to the person selling the biodiesel at a retail establishment or to the person using the pure biodiesel.

Ten states additionally have some form of a purchase mandate for ethanol or biodiesel. Minnesota has a mandate that ten percent of gasoline sold within the state must be a ten percent blend of ethanol with gasoline (referred to as E10) and all diesel fuel must contain two percent biodiesel (referred to as B2) (Carriquiry; Koplow).

Recent research on the use of the oxygenate MTBE has also increased the demand for ethanol. MTBE and ethanol can both be used as oxygenates for reformulated gasoline and oxygenated gasoline. Prior to the year 2000 MTBE was the dominant oxygenate used by gasoline producers, but in 1999 MTBE was found by the EPA to be a potential contaminate of ground water and reservoirs. Several states prohibited the use of MTBE in response to the study, but it wasn’t until 2005 that major petroleum companies announced they would completely switch from MTBE to ethanol due to increased liability concerns (EIA, 2006).

As of 2007, ethanol accounts for almost all oxygenates used within the United States (EIA, 2007). Oxygenates are added to motor vehicle fuels to make them burn more cleanly, thereby reducing carbon monoxide emissions. It is for this reason that
oxygenates are used in the production of two different types of gasoline called reformulated gasoline and oxygenated gasoline. Some regions of the United States mandate the use of reformulated or oxygenated gasoline. The demand for mandated oxygenated gasoline is very small in comparison to total gasoline production, but mandated reformulated gasoline makes up roughly 1/3 of the total amount of gasoline produced. Reformulated gasoline requires a minimum 5.8 percent (by volume) ethanol in its production. Oxygenated gasoline requires a minimum 7.4 percent (by volume) ethanol in its production (EIA, 2007).

Gasoline producers also have the option of using additives, such as ethanol, to increase the octane of gasoline (EIA, 2007). A ten percent use of ethanol raises the octane level of gasoline 3 points (Lee). E10 and lower percentages of ethanol are certified to be used within conventional gasoline engines and gasoline pumps; therefore it is assumed that consumers of E10 or gasoline fuels with lower percentages of ethanol are most often unaware of the percentage of ethanol that exists within their gasoline.

E85, 85% ethanol and 15% gasoline mix, would provide a choice for end users of fuel between a majority gasoline based fuel or a majority ethanol based fuel, but E85 requires a flex fuel vehicle and E85 fueling pumps. Currently there is an insignificant number of flex fuel vehicles and E85 fueling pumps operating in the United States, therefore it is largely assumed that the majority of consumers of automobile fuel do not have a choice between a majority ethanol or a majority gasoline based fuel. While the
performance of ethanol is similar to gasoline except a slight increase in torque, consumers receive 30% fewer miles per gallon of ethanol compared to a gallon of gasoline (Al-Hasan; de Gorter; Roberts).

Biodiesel can be used in almost all diesel engines in its purest form of B100 or in its most common forms that use 2%, 5% or 20% biodiesel mixed with diesel fuel referred to as B2, B5 or B20 (Strong). Biodiesel is commonly used in proportions of 20% or lower due to decreasing cold weather performance as the percentage of biodiesel increases. Standardization and warranty issues are additional reasons why 20% biodiesel or less are most commonly used (Strong). Most studies have shown no appreciable difference in engine durability or fuel economy with the use of biodiesel compared to diesel fuel (Strong).

Ethanol is produced by two different methods; either a dry grind or wet mill process. The difference between the two processes is that the dry grind process is able to maximize the amount of capital return per gallon of ethanol, while the wet mill process allows for further separation of other valuable resources, such as corn gluten meal, within the process. The majority of plants use the dry grind process. The wet mill process produces 2.5 gallons of ethanol per bushel of corn, while the dry grind process produces 2.8 gallons of ethanol per bushel of corn (Bothast). Within both methods, fermentable sugars are produced from converting starch originating from the corn in to sugar. The sugars are in turn fermented into ethanol (Bothast).
Biodiesel is derived by a process called transesterification in which feedstocks, such as vegetable oil, are mixed with alcohol, most commonly methanol or ethanol, in the presence of a catalyst, such as sodium hydroxide, to produce methyl esters. Biodiesel can be produced using most edible oils including canola oil or palm oil, but within the United States it is most commonly produced with soybean oil (Strong). One bushel of soybeans can produce approximately 1.49 gallons of biodiesel or one gallon of soybean oil can produce one gallon of biodiesel (Gray).

The increased production of biofuels indicates that the production capacity within both biofuel markets is rapidly expanding. Ethanol production capacity has increased from 1.7 billion gallons in 2000 to an average production capacity in 2007 of 5.5 billion gallons per a year (RFA, 2008a). Capacity for biodiesel has at least doubled since the beginning of 2006 (Koplow).

Despite the rapid expansion of production capacity within biodiesel and ethanol markets, the capacity utilization rates of each industry are vastly different. In 2007 the capacity utilization rates within the biodiesel industry were around 50% (Koplow). Although the EIA reports that capacity utilization within the ethanol industry is above 95%, ethanol industry reports regularly state that plants are running 10-20 percent above their nameplate capacity, suggesting that it might be more accurate to state that utilization rates within the ethanol industry are at or above 100% (Koplow).

Corn and soybean oil have been the two major feedstocks in the production of biofuels within the United States. In the case of corn usage in the production of ethanol there has been an increase from an estimated 7.5% to 23% of total corn produced from
2001-2007 (United States Department of Agriculture (USDA), 2007). In the case of soybean oil usage in the production of biodiesel there has been an increase from 5% to 18% of United States consumption from 2005 to 2007 (USDA, 2008a).

The industrial use of corn and soybean oil in the production of biofuels has replaced some of their prior use, but they are still used in the production of food products and most significantly as a feed for livestock. The majority of corn and soybean oil usage in food products is often done indirectly and without the knowledge of consumers. Soybean oil is an ingredient in many everyday food products such as candy, sandwich spreads and margarine. The EPA estimates that Americans consume on average 60 pounds per capita per year of soybean oil (2001). Corn is consumed as a food product indirectly through products, such as high fructose corn syrup that is most commonly used in soft drinks. The USDA estimates that Americans consume 41.5 lbs per capita per year of high fructose corn syrup (2008b).

Corn and soybeans also make up a majority of the ingredients used to feed livestock in the United States. According to the USDA, in the years 2001 and 2007, 74% and 58% respectively of the domestic use of corn was categorized as feed and residual (2007).

Nearly all soybeans are crushed to separate the oil from the meal. Soybean meal obtained through the soybean crushing process is the world’s most important protein feed, accounting for nearly 65 percent of world supplies. In addition livestock feeds account for 98 percent of soybean meal consumption (Ash).
The establishment of how important a role soybean oil and corn play in the production of livestock, food products and biofuels is essential in understanding both the implications and the probability of a change in the price behavior of soybean oil and corn. A change in the price behavior of soybean oil and corn in the form of correlation or cointegration with diesel and gasoline prices would present a new type of price risk to livestock, agriculture and food industries. The new type of price risk would be related to price changes of corn and soybean oil caused not only by variables such as weather patterns, livestock demand and food demand, but also by variables such as seasonal car driving patterns, economic growth and crude oil supply shocks. The correlation or cointegration of soybean oil and diesel and also corn and gasoline would be the result of a new economic relationship caused by the production of biofuels. The new economic relationship would be that soybean oil and corn have become substitutes or complements for diesel and gasoline respectively.
CHAPTER 2

BIOFUEL, AGRICULTURE AND ENERGY PRICE RELATIONSHIPS

2.1 Substitute and Complementary Relationships

Defining and understanding complement and substitute relationships starts with the work of Edgeworth-Pareto and the use of indifference curves. Hirschleifer states indifference curves represent a constant level of utility across differing ratios of two consumption goods (p. 72). Figure (3) is a case of the indifference curves for two goods that are perfect complements. Figure (4) is a case of the indifference curves for two goods that are perfect substitutes.

![Figure 3: Perfect complements](image1)

![Figure 4: Perfect substitutes](image2)
The slope of the indifference curves can be defined as the negative of the marginal rate of substitution. The marginal rate of substitution represents an increase in the first good needed to offset a reduction of one unit of the second good while maintaining the same utility. If the marginal rate of substitution is two then in order for an individual to give up one unit of good Y and maintain the same utility they would have to obtain two units of good X.

The marginal rate of substitution is undefined in figure (3), because no amount of the first good is able to replace a reduction of one unit of the second good while maintaining the same utility level. A reduction in the quantity of either good lowers the utility level of the consumer regardless of an increase in the consumption of the other good.

The marginal rate of substitution in figure (4) is constant for any particular point on an indifference curve resulting in a linear indifference curve. Starting at any particular bundle of goods on a linear indifference curve, the amount of good Y that needs to be increased by a reduction of one unit of good X to maintain the same utility is constant. Imperfect complements and imperfect substitutes have indifference curves that are non-linear, representing varying marginal rates of substitution for a particular indifference curve that depends on the initial bundle of goods.

The optimal bundle of goods occurs where the indifference curve is tangent to the budget line. The budget line is all combinations of the two goods that can be purchased, given a level of income. The budget line is assumed to be linear and the slope of the
budget line is the negative of the price ratio of the two goods. It is also important to recognize that the tangency of the two curves is the point at which the marginal rate of substitution is equal to the price ratio of the two goods.

Perfect substitutes in figure (5) will have a corner solution in which only the relatively cheaper good will be purchased and consumed by the individual. In figure (5) the consumer consumes all of good X if good X is the cheaper good represented by the steep dotted budget line with the set of prices \( P_{X2} \) and \( P_{Y2} \). Only Y is consumed if good Y is the cheaper good represented by the flatter dotted budget line with the set of prices \( P_{X1} \) and \( P_{Y1} \). Perfect complements in figure (6) are only consumed at a constant ratio therefore the reduction in the price of good Y will lead to an increase in the consumption of both goods at the same ratio given the goods are infinitely divisible.
For imperfect complements and substitutes an increase in the price of good Y while holding the other price constant can be broken down into two effects. The substitution effect occurs as individuals substitute the consumption of good Y to the now relatively less expensive good X and the income effect as individuals purchase more of good Y as a result of an increase in real income due to a reduction in the relative price of good X. Both effects occur simultaneously for imperfect complements and substitutes, but the effects can be shown in isolation. Gross substitutes have a greater substitution effect resulting in an increase in the demand for good X due to an increase in the price of Y. Gross complements have a greater income effect resulting in a decrease in the demand for X due to an increase in the price of Y.

The representation of complements and substitutes through this method provides a somewhat straightforward presentation of static price theory. Samuelson states that representing static theory with this method was popular with Marshall and many other economists in the early part of the twentieth century, but Hicks pointed out that sometimes theoretically contrary results can occur if changes in real income are not compensated for. The interpretation of the Hicksian equation is that for substitutes an increase in the relative price of one good will decrease the demand for the second good while holding real income constant.

The previous analysis was based on consumer choice between two goods, but two factors of production can be analyzed in a similar way with the replacement of indifference curves with isoquants. Isoquants are curves representing differing bundles of two inputs that produce the same level of output. The marginal rate of technical
substitution represents the substitution ratio of one resource for another while holding output constant along an isoquant. Firms will additionally choose a bundle of input resources based on the tangency of the isoquant and the isocost line. The isocost line represents the different bundles of resources that can be purchased given a constant level of income and its slope is the negative of the price ratio of the two resources.

Static price theory can be used to understand the price relationships within biodiesel and ethanol markets by replacing a standard good X and good Y with biodiesel, diesel, ethanol and gasoline.

Biodiesel is most accurately analyzed as a consumer good. While biodiesel is mixed with diesel by refiners customers are still able to choose between biodiesel and diesel. The relationship between diesel and biodiesel therefore can be represented through indifference curves.

Ethanol is most accurately analyzed as a factor of fuel production and therefore through isoquants. The decision to substitute ethanol for gasoline in the production of fuel and the requirement that ethanol be used as a complementary additive in the production of reformulated and oxygenated gasoline is the responsibility of fuel producers. Currently it would not be as accurate to analyze ethanol as a consumer good, because only a small number of flex fuel vehicles and E85 fuel pumps exist in the United States. The majority of customers within a fuel market cannot choose between gasoline and ethanol. It is assumed that customers of fuel markets do not specifically demand ethanol and do not take into account the proportion of ethanol within the gasoline they purchase.
Biodiesel and diesel are perfect substitutes. According to static price theory, if they are perfect substitutes an increase in the price of diesel above the price of biodiesel produces a corner solution similar to figure (5) in which markets that have access to biodiesel will demand only biodiesel. The cross price elasticity, which is defined as the change in the quantity demanded for one good given a change in the price of another good, is positive as a decrease in the price of diesel reduces the quantity demanded for biodiesel.

Ethanol and gasoline markets are both perfect complements and perfect substitutes. Ethanol has a necessary role as a complement for producers in its application as a mandated additive in the production of reformulated and oxygenated gasoline. For perfect complements the proportions of the two goods remain constant no matter the output level or price ratio. A reduction in the price of gasoline, holding the price of ethanol constant, increases the quantity demanded for both ethanol and gasoline at the same proportion as existed before the price change; similar to figure (6). Therefore when ethanol is used as a complement, the cross price elasticity is negative, because a decrease in the price of gasoline increases the quantity demanded for ethanol.

Ethanol can also be used as a perfect factor substitute for gasoline in the production of fuel. An increase in the price of gasoline above the price of ethanol causes fuel producers that have access to ethanol to demand ethanol up to the allowable 10% proportion of ethanol to gasoline. When ethanol is used as a substitute the cross price elasticity is positive as a decrease in the price of gasoline reduces the quantity demanded for ethanol.
2.2 Derived Demand

While there is both an income effect and a substitution effect when analyzing the net effect of a change in the price of gasoline on the demand for ethanol, we cannot use the same analysis as with imperfect complements or imperfect substitutes. It is similar in that we are dealing with one individual producer of fuel, but there are two separate sets of isoquants for that one producer. The producer faces a linear isoquant, the same shape as the utility curves in figure (4), when the producer is substituting between gasoline and ethanol. The producer faces an isoquant that has the same shape as the utility curves in figure (3), when the producer is producing reformulated or oxygenated gasoline. The net effect cannot be shown through isoquants, but must be analyzed using demand curves for gasoline and ethanol.

In the general case, the net effect a price change within gasoline markets has on ethanol markets will be shown to be determined by the demand elasticities of gasoline and ethanol.

Based on the laws of derived demand, the demand elasticity of gasoline and the demand elasticity of ethanol as a substitute are assumed to be very different from each other with the demand elasticity of ethanol as a substitute being much more elastic. Hirshleifer states that derived demand is the employment of factors and goods by firms and consumers insofar as their employment contributes to the production of other goods demanded by producers and consumers (p. 431). Gasoline, ethanol, diesel and biodiesel can be broadly labeled as being derived demand for automobile fuel. As in the previous
discussion of consumer and producer substitutes and complements the derived demand within these goods can be further specified as derived demand from consumers or producers, but for general purposes it will be assumed that they are all derived demand for automobile fuel.

A change in the demand for automobile fuel affects the demand for gasoline, ethanol, biodiesel and diesel. As stated by Samuelson and alluded to by Hicks the relation of two goods as complements or substitutes is defined by derived demand or stated differently they are defined by their relation to other goods (Ogaki; Hicks; Sato). For instance the derived demand for gasoline and ethanol from fuel markets defines their relationship as substitutes or complements depending if ethanol is being demanded by producers of automobile fuel as an additive or substitute for gasoline.

Marshall stated the four laws of derived demand which can be applied to help gain a better understanding of the demand elasticities of the different markets. As stated by Pigou and published by Sato:

1. The demand for anything is likely to be more elastic, the more readily substitutes for that thing can be obtained.

2. The demand for anything is likely to be more elastic, the more elastic is the demand for any further thing it contributes to produce.

3. The demand for anything is likely to be less elastic, the less important is the part played by the cost of that thing in the total cost of some of other thing, in the production of which it is employed.
4. The demand for anything is likely to be more elastic, the more elastic is the supply of its co-operant agents of production.

All with the exception of the third law were proven mathematically by Hicks in the case where the substitution between two factors is not necessarily equal to zero. Hicks states that the third rule is only true if the elasticity of demand for the final product is greater than the elasticity of substitution between factors of production (Bronfenbrenner).

The first law can be affirmed graphically by a similar analysis previously used for perfect substitutes. If they are perfect substitutes and there is enough supply in both markets to individually meet demand then the demand elasticity for both is perfectly elastic as a price increase in good X over the price of good Y results in zero good X being demanded. Ethanol is a near perfect substitute for gasoline therefore the law leads us to assume that the demand for ethanol is highly elastic when it is being used as a substitute. The cross price elasticity of ethanol in its use as a substitute for gasoline is positive and very elastic.

Similarly biodiesel as a near perfect substitute for diesel leads us to assume that the demand for biodiesel is also very elastic. The cross price elasticity of biodiesel is positive and very elastic.

Gasoline and diesel markets are assumed not to be elastic even though they are also perfect substitutes for ethanol and biodiesel. Gasoline and diesel have a much greater production capacity and as a result have much larger share of the automobile fuel
market. Switching from a small share of gasoline and diesel to ethanol and biodiesel will not have a significant effect on the percentage change in demand of gasoline and diesel.

The second law is important in analyzing gasoline and diesel markets, because of their dominant share in the fuel market. Given the large dependence of Americans on automobiles and deficiency in public transportation in many areas in the United States there is assumed to be short term inelasticity in the demand for miles driven with fuel based transportation (Parry and Small). According to the second law the inelasticity in the market for miles driven with fuel-based transportation results in inelasticity in the demand for fuel including gasoline and diesel. The cross price elasticity of ethanol in its use as a complement with gasoline can be defined as the demand elasticity of gasoline times the percentage of ethanol used as an additive in a unit of gasoline. The cross price elasticity of ethanol in its use as a complement with gasoline is negative, but inelastic.

The third law does not apply to ethanol markets given the caveats stated by Hicks. It has already been established that the substitution elasticity of ethanol and gasoline is much greater than the demand elasticity for miles driven by fuel based automobiles.

The fourth law can be applied toward the demand elasticity of ethanol when being used as an additive. The supply elasticity of gasoline affects the demand elasticity of ethanol as shifts in gasoline demand result in a large increase or decrease in the quantity of gasoline and ethanol supplied. Due to high capacity utilization among oil refineries the supply of gasoline is also assumed to be relatively inelastic.

In the general case if it is assumed that the demand and supply elasticity of gasoline is inelastic, while the demand elasticity of ethanol when it is being used as a
substitute is very elastic the net effect of a price change in gasoline on the demand for ethanol can be shown to be positive even though a greater proportion of ethanol is used as a complement rather than a substitute.

A gasoline price increase in an inelastic gasoline market will only slightly reduce the quantity demanded for ethanol to be used as a complement. A gasoline price increase will greatly increase the quantity demanded for ethanol to be used as a substitute, because of the very elastic demand curve within the substitute ethanol market. The net effect will be an overall increase in the quantity demanded for ethanol as a result of an increase in the price of gasoline.

Figure 7: U.S. gasoline and ethanol markets

Figure 8: Alternative gasoline and ethanol markets
In figure (7) we see a case in which the demand curve for gasoline ($D_g$) in the upper graph is very inelastic and the demand curve for ethanol ($D_e$) in the lower graph is very elastic when the price of ethanol is below the price of gasoline. The supply curve of gasoline is presented as being somewhat elastic for visual purposes even though it is assumed to be inelastic. The elasticity of the supply curve will have little effect on the net outcome given the shift in supply rather than demand. A shift inward of the gasoline supply curve causes the price of gasoline to rise from $P_{g1}$ to $P_{g2}$ and the quantity demanded of gasoline to be slightly reduced from $Q_{g1}$ to $Q_{g2}$. Within the ethanol market below the decrease in the quantity demanded of gasoline causes the ethanol demand curve to shift in as the quantity of ethanol that is demanded as a complement decreases, but the increase in the price of gasoline also causes the demand curve for ethanol to shift up as firms have more opportunity to substitute ethanol for gasoline. Initially the price of gasoline was equal to the price of ethanol and all ethanol up to the quantity $Q_{e1}$ was demanded to be used as an additive for gasoline. A small quantity of ethanol demanded as an additive was reduced represented by the difference between $Q_{e1}$ and $Q_{e3}$, but the quantity of ethanol that was used as a substitute was increased from zero to the difference between $Q_{e3}$ and $Q_{e2}$. In the case of an elastic ethanol demand curve and an inelastic gasoline demand curve the result of an increase in the price of gasoline greater than the price of ethanol causes the overall demand for ethanol to increase from $Q_{e1}$ to $Q_{e2}$.

In figure (8) there is an alternative case in which the gasoline demand curve is somewhat elastic, but the demand curve for ethanol is inelastic. The result of the same
shift in the supply curve in the gasoline market results in a large reduction in the quantity
demanded of gasoline from \( Q_{g1} \) to \( Q_{g2} \). The price of gasoline increases less than it did in
figure (7) from \( P_{g1} \) to \( P_{g2} \). There is still substitution within the ethanol market represented
by the difference between \( Q_{e3} \) and \( Q_{e2} \), but the amount of ethanol that has been demanded
for its use as an additive has been reduced from \( Q_{e1} \) to \( Q_{e3} \). In the case of an inelastic
demand curve in the ethanol market and an elastic demand curve within the gasoline
market the result of an increase in the gasoline price on the overall quantity of ethanol
demanded is reduced from \( Q_{e1} \) to \( Q_{e2} \).

Figure (7) with an elastic demand curve within the ethanol market and an inelastic
demand curve within the gasoline market is more representative of the current market
within the United States, but there are several caveats in the theory that an increase in the
price of gasoline will result in a net increase in the demand for ethanol.

The theory assumes that gasoline producers are not already substituting ethanol at
the maximum allowable mix of 10% prior to the increase in gasoline prices. Given the
requirement that ethanol cannot exceed the 10% proportion at this limit producers will
not be able to substitute more ethanol if they are already applying it at the maximum
level. In 2007 141 billion gallons of gasoline was consumed, but only 6.5 billion gallons
of ethanol was produced. (EIA, 2008c) A rough calculation shows that in 2007 there was
only a 4.7% proportion of ethanol to gasoline being consumed within the United States; a
number that is well under the maximum level of 10%. The high level of capacity

23
utilization among ethanol producers would also indicate that the market is under the limit. If the market was saturated with ethanol at the 10% level we would expect to see slack in the capacity utilization of ethanol plants.

The theory also assumes that the price of gasoline rises above the price of ethanol for the entire market. In reality due to transactions costs, different markets face different prices for ethanol. The price of gasoline may rise above the price of ethanol in Illinois, but gasoline may still be below the price of ethanol in Texas. It is therefore possible that markets in Illinois will increase their demand for ethanol in response to an increase in gasoline prices, while markets in Texas will decrease their demand for ethanol in response to an increase in gasoline prices. Within this case the proportion of ethanol used as a complement compared its use as a substitute within the fuel market may influence the net effect that a change in the price of gasoline has on the demand for ethanol. If there is only a small portion of the market that is facing a higher relative gasoline price even with an increase in gasoline prices than there may be little substitution of ethanol, but there may be a small but still bigger reduction of ethanol in the additive market. The net effect may be a decrease in the demand for ethanol in response to an increase in gasoline prices if the average ethanol price is higher than the average gasoline price.
Notes: (a) Weekly cash prices. (b) Flex differential with credit = (gasoline * .7) – ethanol + tax credit - (.3*fuel tax). (c) Data source: same as figure 1.

Figure 9: Gasoline and ethanol flex differential with the tax credit

Using calculations suggested by de Gorter Figure (9) and Figure (10) show the difference between gasoline and ethanol prices. In figure (9) it shows the price difference as it would exist within a market in which the majority of fuel consumers had flex fuel vehicles and access to E85 fuel pumps. The tax credit is also subtracted from the price of ethanol. Within this market consumers are aware of the amount of ethanol that exists within their fuel and as a result take into account the miles per gallon difference between ethanol and gasoline. This type of market is referred to as a “flex” market.
Notes: (a) Weekly cash prices. (b) Mix differential with credit = gasoline − ethanol + tax credit. (c) Data source: same as figure 1.

**Figure 10: Gasoline and ethanol mix differential with the tax credit**

In figure (10) it shows the difference between gasoline and ethanol as it exists within the current United States market in which consumers are assumed to be unaware of the amount of ethanol within the fuel they use and therefore do not take into account the miles per gallon difference between ethanol and gasoline. This type of market is referred to as a “mix” market. In the “mix” market differentials above zero represent time periods in which ethanol is cheaper than gasoline when the tax credit is taken into account. It shows that on average there are market incentives for substitution in 2005 and 2007, but in 2006 ethanol was on average more expensive than gasoline making substitution less likely. The data used was averaged weekly data for national gasoline and ethanol cash prices. de Gorter reported the same conclusions when averaging by year.
Based on the results it is fair to assume that from 2005 to 2007 on average ethanol producers have had incentive to substitute ethanol for gasoline even without the current ethanol mandate, but with the establishment of the mandate it makes it possible that producers will still react to changes in gasoline price even if ethanol is more expensive than gasoline. If producers are mandated to consume ethanol at a certain level then they will still react to more expensive gasoline by increasing demand for ethanol in order to fulfill the mandate when the price differential has become more favorable.

The last assumption is that fuel producers have access to more ethanol if demand increases. There are two foreseeable cases in which this assumption could be wrong. One case is there could be transportation constraints in which certain fuel producers do not have access to ethanol. Given the use of ethanol in the production of reformulated and oxygenated gasoline, 1/3 of the total amount of gasoline produced, it is reasonable to assume that a large amount of gasoline producers have physical access to ethanol.

The other case is ethanol producers may not be able to increase production of ethanol due to a capacity constraint within the ethanol industry. Statistics show that ethanol plants are at very high capacity utilization levels implying that the demand curve may be crossing the supply curve at a very inelastic point in the supply curve. A very inelastic linear supply curve \( S_{c2} \) was added to figure (7) in figure (11). With the somewhat elastic supply curve \( S_{c1} \) the quantity of ethanol demanded increased from \( Q_{c1} \) to \( Q_{c2} \), but with the inelastic supply curve the quantity of ethanol demanded increases very little or does not increase.
Another important implication of an inelastic supply curve within the ethanol market relates to how potential supply constraints affect the gasoline and corn price relationship. The affect can be understood by recognizing that derived demand relationships also link the feedstock markets of corn and soybean oil to both their respective biofuel markets and the related energy markets.

An increase in the demand for ethanol due to an increase in the price of gasoline should increase the demand for corn and as a result the price of corn. In the case of an inelastic supply curve for ethanol the demand for corn and the price of corn increases.

Figure 11: U.S. gasoline and ethanol markets with inelastic ethanol supply
very little in response to a shift in the demand for ethanol. If the price of gasoline increases, but the production of ethanol is constrained, the price relationship between corn and gasoline then becomes disconnected as changes in gasoline prices affect the price of corn very little.

An increase in the demand for biodiesel due to an increase in the price of diesel should increase the demand for soybean oil and as a result the price of soybean oil. Within biodiesel markets it is assumed that the demand curve is elastic and the supply curve is somewhat elastic given the spare production capacity that exists within the biodiesel industry. An increase in the price of diesel causes an outward shift in the demand for biodiesel and an increase in the quantity of biodiesel produced. The increased production of biodiesel also increases the demand for soybean oil and increases soybean oil prices. The expected positive price relationship between soybean oil and diesel prices is not constrained as it potentially is in the case of gasoline and corn.

The testing of the corn and gasoline price relationship and the soybean oil and diesel price relationship allows for somewhat of a comparative analysis that may suggest that the corn and gasoline price relationship is affected by capacity constraints within the ethanol industry. Corn and soybean oil are used in similar proportions compared to their total production. Despite two different types of ethanol application, as a substitute and as a complement, it has been shown that a relationship between gasoline and corn prices that is similar to diesel and soybean oil prices, should theoretically exist. Therefore a strong
change in the price relationship between soybean oil and diesel, but not corn and gasoline may be an indication that capacity constraints are causing an inconsistent connection between corn and gasoline prices.
CHAPTER 3

METHODOLOGY AND RESULTS

3.1 Data

The dataset used to test the relationship among the commodity prices through cointegration tests and the estimation of a confidence interval of correlation coefficients is made up of daily cash prices for soybean oil, gasoline, heating oil and corn from January 1, 1980 to January 11, 2008. Heating oil prices will be used as a proxy for diesel prices. The NYMEX states that the heating oil futures contract is also used to hedge diesel fuel therefore it is assumed that heating oil futures prices are very strongly reflective of diesel prices. The cash prices were provided by the Commodity Research Bureau.

3.2 Cointegration Review and Methodology

Rao states that standard methods of estimation are based on an assumption that the means and variances of the variables are well defined constants and independent of time
Variables whose means and variances change over time are known as being non-stationary and may also contain a unit root process. Some commodity price series have a unit root, because they are characterized by a ‘random walk’. A ‘random walk’ is defined as the value in period t is equal to the value in period t-1 plus a random error,

\[ p_t = p_{t-1} + e_t \]  

Kennedy states that Student’s-t, Durbin-Watson and R² statistics do not retain their traditional characteristics in the presence of non-stationary data and the application of an OLS regression with non stationary data produces spurious results (p. 325).

Engle and Granger state that while an individual time series may be non-stationary a combination of two non-stationary time series, such as

\[ y_t - \alpha - \beta x_t = e_t \]  

may be stationary. If so, these two series are said to be cointegrated. The presence of the cointegrated relationship represents a long run equilibrium relationship between the two variables. Rao also states that the general interpretation of this relationship is that due to an underlying economic relationship the two individual time series cannot move “too far” away from each other and thus suggesting a long run link between the two variables (p. 4). Cointegration testing has been applied to test the long run relationship between two or more non-stationary series in several areas.

Cointegration testing has been used to test the spatial integration of two or more markets of a single homogenous commodity by Baffes (1991), Ardeni (1989), Goodwin and Schroeder (1991) and Mohanty, Peterson and Smith (1996). International
commodity arbitrage in efficient markets implies that for a single homogenous commodity under a common currency unit with no transportation costs the prices should be equal if the law of one price holds. Even in the presence of transportation costs the law of one price can be tested through a cointegration test on two or more homogenous commodity price series that are individually non-stationary.

The long run relationship between spot and futures prices was tested through cointegration by Thraen and Petrov (1997), Kellard, Newbold, Rayner and Ennew (1999), Zhou (2005) and Brockman and Tse (1995). If the futures prices and spot prices of the same commodity are found to be non-stationary as individual series cointegration testing can be performed to establish a long run relationship between the futures price and spot price of the same commodity. As stated by Thraen and Petrov the existence of a long-run equilibrium is a necessary condition for effective hedging opportunities.

Cointegration can also be used to test the long run relationship between two or more different commodities. The use of cointegration testing on different commodities can provide evidence of a substitute or complementary relationship between the different commodities.

Malliaris and Urrutia find evidence of a long-run relationship between U.S. grown corn, wheat, oat, soybean, soybean meal and soybean oil futures prices. Booth and Ciner assert the long-run relationship between these commodity products may not be evidence of an underlying economic relationship between the commodities, but may be the result of a co-movement of prices fostered by similar government farm policy and weather within the United States. After testing corn, redbean, soybean and sugar commodities
produced in different countries, Booth and Ciner were unable to find cointegration between the different commodities, except for corn and soybean. Booth and Ciner conclude that the findings were evidence that traders did not display a “herding” mentality across all commodities without the establishment of an economic relationship between them. Williams and Bessler use cointegration to test the long-run price relationship between the substitute’s refined sugar and high fructose corn syrup and found evidence of cointegration from 1984 to 1991.

Substitute and complementary relationships between agriculture, biofuel and energy commodities have also been tested through cointegration. Yu, Bessler and Fuller tested the cointegration of the oil markets of soybean, sunflower, rapeseed, palm and crude. Using data from 1999 to March of 2006 Yu found no evidence of cointegration between any of the edible oil markets with crude oil markets.

Higgins, Bryant, Outlaw and Richardson, using data from 1989 to 2005, tested the cointegration of several price relationships within the ethanol industry. Higgins et al. found evidence of cointegration between ethanol and gasoline, between ethanol and corn and between MTBE, ethanol and natural gas.

Zhang and Reed tested the cointegration of Chinese corn, soy meal and pork prices with world crude oil prices. Using data from 2000 to October 2007 Zhang found no evidence of cointegration among any of the price relationships.

Campiche et al. tested the cointegration of corn, soybean, soybean oil, palm oil, sugar and crude oil in two different time periods. No cointegrated relationships could be established in the time period 2003-2005. Corn and soybeans, but not soybean oil were
found to be cointegrated with crude oil from 2006 through the first half of 2007. Calculation of correlation coefficients using the Fisher Transformation provided evidence that corn and crude oil had a negative correlation from 2003 to 2006 that was significant and had a negative correlation in the first half of 2007 that was insignificant. Soybean oil and crude oil were insignificantly correlated from 2003 to 2006, but had a very high positive correlation in 2007 that was significant.

The results obtained by Campiche et al. are not theoretically consistent. If soybeans are cointegrated with crude oil in the time period after 2006 then it is expected that soybean oil would also be cointegrated given it is the intermediate product in the derived demand relationship between soybeans, soybean oil and biodiesel. The negative correlation between corn and crude oil markets that includes data points within the year 2006 are also inconsistent with the results of the cointegration tests. Campiche et al. acknowledges the inconsistency of the results and suggests a low number of data points in the second time period as a possible explanation for the inconsistent results. Kennedy also notes that estimates of the cointegrating regression have considerable small sample bias (p. 328).

The current analysis hopes to receive more conclusive results by utilizing a larger dataset and by testing more directly the substitute or complementary relationships of corn and gasoline and soybean oil and diesel.
3.3 Cointegration Estimation and Results

For the unit root and cointegration tests the dataset was divided up into two different time periods to test if the increased production of biofuels has caused corn and soybean oil to be cointegrated with gasoline and diesel respectively. The first time period of 1980 to the end of 2004 will be used to test if there is cointegration in the presence of very low levels of biofuel production. The second time period of 2005 to January 2008 will be used to test for cointegration in the presence of much higher levels of biofuel production in comparison to the average level of biofuel production in the previous 25 years.

The necessary condition for cointegration testing is that the price series individually are non-stationary and contain a unit root. The most widespread test for stationarity is the augmented Dickey-Fuller test. The augmented Dickey-Fuller test allows for lagged values and in the case of two lags is represented by

\[ \Delta y_t = \alpha + \theta y_{t-1} + \gamma \Delta y_{t-1} + \epsilon_t \]

where \( \theta = (\rho - 1) \) and \( \Delta y_t = y_t - y_{t-1} \). Holden and Perron in Rao state the number of lags needed is determined by using an information criterion (p. 68). The null hypothesis is that the series is non-stationary or that \( \rho = 1 \). A specialized test statistic must be used to test the null. The test statistic based on the number of observations and the form of the equation was calculated by Fuller and has been published by Hamilton (p. 763).

Two lags were chosen for each price series based on the minimization of the Akaike Information Criterion. Equation (3) that includes a constant and no time trend was tested for a unit root within the four price series for each time period. The results are
shown in Table 1. In the time period 1980-2004 for gasoline and heating oil \( \theta \) was significantly different from zero at the 1% level, for soybean oil \( \theta \) was significantly different from zero at the 5% level and for corn \( \theta \) was significantly different from zero at the 10% level. In the time period of 2005-2008 for gasoline \( \theta \) was significantly different from zero at the 5% level and for corn, soybean oil and heating oil \( \theta \) was not significantly different from zero. Only the pair of soybean oil and heating oil in the 2005-2008 time period were each individually non-stationary and therefore could be tested for cointegration.

<table>
<thead>
<tr>
<th>Variables</th>
<th>ADF</th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corn</td>
<td>-2.85*</td>
<td>0.40</td>
</tr>
<tr>
<td>Gasoline</td>
<td>-3.45***</td>
<td>-2.89**</td>
</tr>
<tr>
<td>Soybean Oil</td>
<td>-2.99**</td>
<td>2.71</td>
</tr>
<tr>
<td>Heating Oil</td>
<td>-3.81***</td>
<td>-1.52</td>
</tr>
</tbody>
</table>

Notes: (a) Critical values for ADF tests at the 10%, 5% and 1% test level are -2.57, -2.87 and -3.44 for \( n_{obs} = 500 \), respectively. (b) *, **, and *** indicate significance at the 10%, 5% and 1% test level, respectively.

**Table 1: Augmented Dickey Fuller unit root tests**

The Engle and Granger approach toward testing cointegration was used. It involves testing the stationarity of a linear combination of variables such as

\[
y_t = \alpha + \beta x_t + e_t
\]
Stationarity is determined by performing a unit root test on the residuals of the estimated co-integrating regression in equation (4). Engle and Granger recommend using the augmented Dickey-Fuller unit root test within this application. By replacing \( y_t \) in equation (3) with \( e_t \) and testing the null hypothesis of a unit root we are testing if there is a linear combination of \( y_t \) and \( x_t \) that is stationary. Note that if \( \beta \) is known to be one we are simply testing if \( y_t - x_t \) or the difference between the two variables is stationary.

Oftentimes \( \beta \) is not known and must be estimated through equation (4). \( \beta \) is referred to as the cointegrating vector. The presence of cointegration between \( x_t \) and \( y_t \) is a rejection of the null hypothesis that the residuals estimated from the equation contain a unit root. Holden and Perron in Rao state that the critical values for the Dickey Fuller statistics for a unit root in the residuals from the cointegrating regression differ from those for a unit root in the variables involved in that regression (p. 80). The test statistics are also calculated by Fuller and published by Hamilton (p. 766).

<table>
<thead>
<tr>
<th></th>
<th>ADF</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005-2008 Heating Oil and Soybean Oil</td>
<td>-1.51</td>
</tr>
</tbody>
</table>

Notes: (a) Critical values for ADF test at the 10%, 5% and 1% test level are -3.45, -3.77 and -4.31 for \( \text{nobs} = 500 \), respectively. (b) *, **, and *** indicate significance at the 10%, 5% and 1% test level, respectively.

**Table 2: Cointegration test**

The result in table (2) shows a non rejection of the null hypothesis of a unit root within the residuals. The interpretation of the result is that soybean oil and heating oil
prices are not cointegrated in the time period of 2005-2008. The results of the cointegration test do not provide evidence of a long run relationship between soybean oil and heating oil from 2005 to 2008. Additionally the results could not be compared to the previous time period, because cointegration prior to 2005 in the soybean oil and heating oil markets could not be tested with a stationary series of heating oil prices. Based on the low levels of biodiesel production from 1980 to 2004 it could be assumed that there was also no long run relationship between soybean oil and heating oil during this time period, but without the ability to test for cointegration during this time period there is no evidence of this or that increased biodiesel production has changed the long run price relationship between soybean oil and diesel. The method of cointegration testing for corn and gasoline could not provide any evidence that increased ethanol production has changed the price relationship between corn and gasoline, because the two commodities could not be tested for cointegration in either time period.

### 3.4 Correlation Review and Methodology

The hypothesis that increased production of biofuels has significantly changed the price relationship between soybean oil and diesel and also corn and gasoline could not be sufficiently tested through cointegration, but correlation testing can be used to test the hypothesis.

Bootstrapping a confidence interval of Pearson’s correlation coefficients may be preferable to testing using Fisher’s transformation, because Fisher’s transformation
requires normality of the distribution of observations (Wilcox, p. 200). Bootstrapping obtains distributions through simulation and therefore hypothesis tests can be performed without making any necessary assumptions about the normality of the distribution of observations.

There are several different methods for calculating a bootstrap confidence interval. For a review see Efron (1981), Diciccio and Romano (1988) and Hall (1988). Two methods for calculating a bootstrap confidence interval are the percentile bootstrap and the bootstrap-t. Both methods are presented by Wilcox (p. 207). Let $X_1, \ldots, X_n$ represent a random sample of observations, and let $X^*_1, \ldots, X^*_n$ represent a bootstrap sample of $n$ that is obtained by random sampling, with replacement, $n$ values from $X_1, \ldots, X_n$. The process of generating a bootstrap sample of size $n$ is performed $B$ times. For every bootstrap sample a Pearson’s correlation $r$ is calculated. In the case of a bootstrap-t confidence interval a bootstrap distribution of $r_1 < \ldots < r_b$ is formed. From the standard deviation of the distribution of $r$ a confidence interval is calculated for the mean $r$ across the observations $X_1, \ldots, X_n$.


Wilcoxon states that in cases of small samples the probability of a type 1 error may be higher than the desired level for a bootstrap-t and therefore the percentile bootstrap
may be recommended. Kilian finds that all bootstrap intervals including the percentile bootstrap tend to have excellent coverage accuracy within large sample sizes.

The percentile bootstrap method does not rely on the distribution of \( r \), but ranks \( r_1 < \ldots < r_n \) and forms a .95 confidence interval from \( (r^*_{(a)}, r^*_{(b)}) \) with (a) and (b) being equal to the .025 and .975 observations of \( r_1 < \ldots < r_n \).

Wilcox used a percentile bootstrap method to form a confidence interval and test \( H_0: \rho = 0 \) with \( \rho \) being the correlation of \( X_1, \ldots, X_n \). The inclusion of zero within the confidence interval is a failed rejection of the null that there is zero correlation between the two variables.

### 3.5 Correlation Estimation and Results

In order to test if the correlation between soybean oil and diesel and the correlation between corn and gasoline have significantly increased, a percentile bootstrap confidence interval of Pearson’s correlation coefficients was formed for the two different pairs of commodities from log price changes from 1980 to 2008. The size of the bootstrap samples is 759 observations, equal to the number of observations from 2005 to 2008. 1000 non-block bootstrap samples were generated randomly with replacement. The samples were not drawn in blocks; because it is assumed any co-movement of the two variables is random. Given the assumption that all pairings of the two variables at a particular time \( t \) are randomly created we would expect that on average any particular block of samples would not be statistically different from a non-block random sample. If
the block sample is found to be statistically different from the randomly generated non-block samples then it is interpreted that the time period of the block sample is different from other time periods within the larger sample. A Pearson’s correlation coefficient was calculated for each bootstrap sample of soybean oil and heating oil log price changes and also corn and gasoline log price changes. Additionally a Pearson’s correlation coefficient $r_1$ was calculated for corn and gasoline and also soybean oil and heating oil for the time period 2005 to 2008.

The null hypothesis being tested is that the mean correlation coefficient for the 1980-2008 time period is greater than or equal to the calculated correlation coefficient $r_1$ for the 2005-2008 time period. The inclusion of $r_1$ within the range of the confidence interval is a failed rejection of the null. Conversely the rejection of the null can be interpreted as the correlation of the two variables from 2005 to 2008 is significantly greater than the mean correlation of the two variables over the entire time period of 1980-2008. Table 3 shows the results of the calculated confidence intervals and the correlation of the two pairings of variables from 1980 to 2008 and from 2005 to 2008. The null for both pairs of variables are rejected at the .95 level and the .99 level of confidence. Both the correlation of soybean oil and heating oil and the correlation of corn and gasoline from 2005 to 2008 are significantly greater than the mean correlation over the last 28 years.
The results also give an indication to what degree the price relationships have changed. By looking at figure (12) and figure (13) and even by looking at the two confidence intervals it can be seen that the historical price relationship between soybean oil and heating oil is somewhat similar to the historical price relationship between corn and gasoline for much of the last 28 years, but from 2005 to 2008 the correlation between soybean oil and heating oil was calculated as being roughly 18 percentage points higher than the maximum value of the .99 confidence interval, while the correlation between corn and gasoline was calculated as being roughly 2.4 percentage points higher than the maximum value of the .99 confidence interval.
Notes: (a) Data source same as figure 1. (b) 90 day rolling correlations of daily percentage price change.

**Figure 12. Gasoline and corn correlation 1980-2008**

Notes: (a) Data source same as figure 1 (b) 90 day rolling correlations of daily percentage price changes

**Figure 13: Heating oil and soybean oil correlation 1980-2008**
The results provide evidence that correlation between soybean oil and heating oil and corn and gasoline have significantly increased, but the results do not provide any evidence of why the correlation of soybean oil and diesel have increased more significantly compared to that of corn and gasoline. Suggested reasons include speculative behavior within the biodiesel market or capacity constraints within the ethanol industry.

Campiche et al. suggest that the difference in correlation in 2007 within their analysis is a result of the relative maturity of the ethanol industry compared to the biodiesel industry. Both industries have rapidly expanded, but biodiesel has seen a much greater percentage increase since the beginning of 2005; increasing from 28 million gallons in 2004 to an estimated 501 million gallons in 2007\(^1\) compared to ethanol increasing from 3400 million gallons in 2004 to 6500 million gallons 2007 (RFA, 2008b). In turn Campiche et al. suggests speculators have a played a role in creating high correlation within the biodiesel market. Uncertainty about how much the biodiesel industry will expand and how much soybean oil will be needed to feed the expansion may have created a speculative environment, while the more mature and slower expanding ethanol industry may have led to less speculation, because of a more accurate perception of how much corn will be needed to keep pace with the expanding ethanol industry.

---

\(^1\) Biodiesel Production data provided by Jacobson Reports and 2007 biodiesel production estimated by J.N Ferris.
Notes: (a) Data source same as figure 1. (b) 60 day rolling correlations of daily percentage price changes.

**Figure 14: Heating oil and soybean oil correlation 2005-2008**

While this may be a fair suggestion over a shorter period of time such as the 6 month period in 2007 that Campiche *et al.* has examined, a look at figure (14) shows that soybean oil and diesel prices have had a consistently higher correlation from 2005 to 2008 compared to corn and gasoline with similar if not less volatility. If high correlation was fueled by speculation it might be expected that over a three year period there would be more drops in the correlation of the prices, but it has remained consistently well above the mean 28 year average suggested by the calculated confidence interval. The corn and gasoline price relationship from 2005 to 2008 in figure (15) shows very little evidence that there ever was a “peak” correlation that may have been due to speculation. The figure also shows several periods of correlation that were within the calculated mean
confidence interval providing further visual evidence that corn and gasoline have significantly, but not drastically changed their price relationship from 2005 to 2008 compared to the previous 25 years.

Notes: (a) Data source same as figure 1. (b) 60 day rolling correlations of daily percentage price changes.

Figure 15. Gasoline and corn correlation 2005-2008

Speculation may be playing some role in the high level of correlation within soybean oil and diesel markets, but the consistent level of higher correlation from 2005 to 2008 compared to its 28 year mean indicates that the price relationship has significantly changed due to changes in the economic relationship, not just speculation.

If a significant reason for the high level of correlation in the soybean oil and diesel markets is due to a change in the economic relationship then it might indicate that corn and gasoline should have developed a similar economic relationship and as a result a similar level of correlation. Perhaps another reason why the corn and gasoline price
relationship has not been very similar to soybean oil and diesel is that high capacity utilization rates within ethanol markets may be causing an inconsistent connection between gasoline and corn prices. If capacity constraints are the reason for relatively lower and inconsistent correlation between corn and gasoline it may be possible to see a very large jump in the level of correlation between gasoline and corn as both production capacity expands and utilization rates decrease.
Through cointegration testing and the calculation of a confidence interval of correlation coefficients the hypothesis was tested that price relationships between corn and gasoline and also soybean oil and diesel have changed due to increased production of biofuels. The results of the stationarity and cointegration tests were inconclusive in testing the hypothesis. The gasoline price series from 1980 to 2005 and 2005 to 2008 and the heating oil price series from 1980 to 2005 were found to be stationary. As a result only soybean oil and heating oil from 2005 to 2008 could be tested for cointegration and were found not to be cointegrated. The results of the calculated confidence interval of Pearson’s correlation coefficients for soybean oil and diesel and also corn and gasoline provided evidence that correlation between soybean oil and heating oil and also corn and gasoline from 2005 to 2008 were statistically different from their 28 year mean correlation. The evidence of a significantly different level of correlation between the two commodity pairings compared to the past supports the hypothesis and leads to an inference that increased production of biofuels has changed the economic relationship between the two pairs of commodities from an independent relationship to a substitute
relationship. Such a relationship presents a new type of price risk to agriculture, food and livestock industries that may require hedging strategies that account for price changes in energy markets. Results of the calculated confidence interval also show that soybean oil and diesel prices have changed their level of correlation more significantly than corn and gasoline. One suggested reason for the difference is that there are capacity constraints within the ethanol industry.

This thesis research is the first step in establishing evidence of a new economic relationship between agriculture and energy commodities. More research is needed on what hedging strategies can best be utilized in the presence of an economic relationship between energy and agriculture markets and also how capacity constraints within the ethanol industry have potentially influenced the price relationship between corn and gasoline.
BIBLIOGRAPHY


Zhang, Qiang and Michael Reed. (2008). Examining the impact of the world crude oil price on china’s agricultural commodity prices: The case of corn, soybean, and pork. *Selected paper for presentation at the Southern Agricultural Economics Association Annual Meetings, Dallas, TX, February 2nd – 5th, 2008.*