

4-2008

# The Impact of Ethanol Production on U.S. and Regional Gasoline Prices and on the Profitability of the U.S. Oil Refinery Industry

Xiaodong Du

*Iowa State University*, [xdu@iastate.edu](mailto:xdu@iastate.edu)

Dermot J. Hayes

*Iowa State University*, [dhayes@iastate.edu](mailto:dhayes@iastate.edu)

Follow this and additional works at: [http://lib.dr.iastate.edu/card\\_workingpapers](http://lib.dr.iastate.edu/card_workingpapers)



Part of the [Econometrics Commons](#), [Industrial Organization Commons](#), and the [Oil, Gas, and Energy Commons](#)

---

## Recommended Citation

Du, Xiaodong and Hayes, Dermot J., "The Impact of Ethanol Production on U.S. and Regional Gasoline Prices and on the Profitability of the U.S. Oil Refinery Industry" (2008). *CARD Working Papers*. 485.

[http://lib.dr.iastate.edu/card\\_workingpapers/485](http://lib.dr.iastate.edu/card_workingpapers/485)

This Article is brought to you for free and open access by the CARD Reports and Working Papers at Iowa State University Digital Repository. It has been accepted for inclusion in CARD Working Papers by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

---

# The Impact of Ethanol Production on U.S. and Regional Gasoline Prices and on the Profitability of the U.S. Oil Refinery Industry

## **Abstract**

Using pooled regional time-series data and panel data estimation, we quantify the impact of monthly ethanol production on monthly retail regular gasoline prices. This analysis suggests that the growth in ethanol production has caused retail gasoline prices to be \$0.29 to \$0.40 per gallon lower than would otherwise have been the case. The analysis shows that the negative impact of ethanol on gasoline prices varies considerably across regions. The Midwest region has the biggest impact, at \$0.39/gallon, while the Rocky Mountain region had the smallest impact, at \$0.17/gallon. The results also indicate that ethanol production has significantly reduced the profit margin of the oil refinery industry. The results are robust with respect to alternative model specifications.

## **Keywords**

crack spread, crude oil prices, ethanol, gasoline prices

## **Disciplines**

Econometrics | Industrial Organization | Oil, Gas, and Energy

# **The Impact of Ethanol Production on U.S. and Regional Gasoline Prices and on the Profitability of the U.S. Oil Refinery Industry**

Xiaodong Du and Dermot J. Hayes

***Working Paper 08-WP 467***

April 2008

**Center for Agricultural and Rural Development  
Iowa State University  
Ames, Iowa 50011-1070  
[www.card.iastate.edu](http://www.card.iastate.edu)**

*Xiaodong Du is a research assistant in the Department of Economics and Center for Agricultural and Rural Development (CARD) at Iowa State University; Dermot Hayes is a professor in the Department of Economics and in CARD, Iowa State University..*

Selected Paper prepared for presentation at the annual meeting of the American Agricultural Economics Association, Orlando, Florida, July 27–29, 2008.

This paper is available online on the CARD Web site: [www.card.iastate.edu](http://www.card.iastate.edu). Copyright 2007 by Xiaodong Du, Dermot J. Hayes. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

Questions or comments about the contents of this paper should be directed to Xiaodong Du, 271 Heady Hall, Iowa State University, Ames, IA 50011-1070; Ph: (515) 294-6042; E-mail: [xdu@iastate.edu](mailto:xdu@iastate.edu); or Dermot Hayes, 568C Heady Hall, Iowa State University, Ames, IA 50010-1070; Ph: (515) 294-6185; E-mail: [dhayes@iastate.edu](mailto:dhayes@iastate.edu).

Iowa State University does not discriminate on the basis of race, color, age, religion, national origin, sexual orientation, gender identity, sex, marital status, disability, or status as a U.S. veteran. Inquiries can be directed to the Director of Equal Opportunity and Diversity, 3680 Beardshear Hall, (515) 294-7612.

## **Abstract**

Using pooled regional time-series data and panel data estimation, we quantify the impact of monthly ethanol production on monthly retail regular gasoline prices. This analysis suggests that the growth in ethanol production has caused retail gasoline prices to be \$0.29 to \$0.40 per gallon lower than would otherwise have been the case. The analysis shows that the negative impact of ethanol on gasoline prices varies considerably across regions. The Midwest region has the biggest impact, at \$0.39/gallon, while the Rocky Mountain region had the smallest impact, at \$0.17/gallon. The results also indicate that ethanol production has significantly reduced the profit margin of the oil refinery industry. The results are robust with respect to alternative model specifications.

**Keywords:** crack spread, crude oil prices, ethanol, gasoline prices.

# **The Impact of Ethanol Production on U.S. and Regional Gasoline Prices and on the Profitability of the U.S. Oil Refinery Industry**

## **Introduction**

Fuel ethanol production in the United States increased from 1.63 billion gallons in 2000 to 7.22 billion gallons in 2007 (RFA). In comparison, the U.S. consumed approximately 146 billion gallons of petroleum in 2007 (EIA). The purpose of this paper is to estimate the impact of this increase in ethanol supply on the U.S. gasoline market.

Ethanol is blended with gasoline to improve octane and performance in about 50% of the nation's gasoline supply. Typically, a gallon of ethanol blend will have 10% ethanol and 90% gasoline. This gallon of ethanol blend will contain approximately 97% of the energy of a gallon of gasoline (Tokgoz et al. 2007) and will use approximately one-tenth as much fuel energy to produce as it contains (Wang et al. 2007). Therefore, ethanol has essentially added to U.S. gasoline supplies by utilizing solar energy to grow the crop, coupled with energy from natural gas and coal to manufacture the farm equipment and fertilizer used in crop production.

In order to identify the separate impact of ethanol on gasoline prices, we need to separate the impact of ethanol from the other forces driving gasoline prices. We do so by examining the price of gasoline relative to the price of crude oil. We also estimate the impact of ethanol on the profits made by refiners. Both estimates are calculated for the U.S. as a whole and for each of five regions within the U.S. The motivation for conducting the regional analysis is that if ethanol is affecting gasoline prices, then we hypothesize that this impact will be largest in the Midwest where regional ethanol production and utilization is at its maximum.

The paper proceeds as follows. First, background information regarding previous work, relative gasoline prices, and the use of the crack spread as a measure of industry profitability are introduced. We then describe the five regional "Petroleum Administration for Defense Districts" (PADDs) that are the basis for the analysis. Next, we present a detailed description of and

motivation for the explanatory variables. We also provide a description of and motivation for the three estimation methods that are used. The last section summarizes the results.

### **Previous Work**

Quantitative analysis of the effect of ethanol on gasoline prices and on the profitability of the refinery industry has been largely neglected in the literature. Eidman (2005) points out that ethanol largely acts as a fuel extender. He also shows that there has been a strong positive correlation between ethanol and gasoline prices. Employing an international ethanol model consisting of behavioral equations for production, consumption, and trade, Tokgoz and Elobeid (2007) analyze the price linkage between ethanol and gasoline markets. They conclude that ethanol is mainly used as an additive to gasoline and that the complementary effect of ethanol dominates the substitution effect on gasoline prices. Szklo, Schaeffer, and Delgado (2007) conclude that by replacing methyl tertiary butyl ether (MTBE), which is a traditional additive used as an oxygenate to raise the octane number, ethanol blending will not reduce gasoline use until flexible fuel vehicles become widely available. Vedenov et al. (2006) apply a continuous-time option pricing method to calculate the decision threshold of switching to ethanol. Their empirical analysis suggests that blending ethanol into gasoline would generate lower gasoline price volatility and that switching from conventional gasoline to an ethanol blend is an economically sound decision.

The “3:2:1 crack spread” is used as one of the significant indicators of refinery profitability. It is a term used in the oil industry and futures trading as a proxy for the profitability of refineries. Although there is some qualitative description of its determinants, formal quantitative analysis is limited in the literature. Asche, Gjolberg, and Völker (2003) examine the price relationships among crude oil and refined products. They find that the crude oil price is weakly exogenous and that the spread is constant among some of the prices. Girma and Paulson (1998) examine the crack spread of daily futures prices of crude oil and heating oil. Girma and Paulson (1999) investigate the long-run relationship among crude oil, gasoline, and heating oil futures prices and find the prices are co-integrated. They also find a stationary relation between crude oil and its end products.

In the literature on mergers in the refinery industry, several studies rely on analysis of the price margin, which is defined as wholesale prices of gasoline less crude oil prices. The Government Accounting Office (GAO 2004) models the price margin as a function of the crude oil price, inventory ratio, utilization rate, and dummy variables representing a merger and acquisition event. Geweke (2003) provides a comprehensive survey on this subject.

The degree of market concentration has been long recognized and analyzed in the literature seeking to explain price changes and adjustment in the wholesale gasoline market. Focusing on Gulf Coast, Los Angeles, and New York whole spot gasoline markets, Oladunjoye (2007) investigates the effects of market structure on the pattern of price adjustment and finds that market concentration has a significant asymmetric effect on gasoline price changes responding to crude price shocks. The GAO (2004) concludes that mergers and increased market concentration generally led to higher wholesale gasoline prices in the United States from the mid-1990s through 2000. Examining wholesale price responses in 188 gasoline markets in the U.S., Borenstein and Shepard (2002) find that refinery firms with market power generally choose a different adjustment rate and adjust prices more slowly than do competitive firms.

## **Background**

The 3:2:1 crack spread is defined as

$$\pi = \frac{2}{3}P_G + \frac{1}{3}P_H - P_O$$

where  $P_G$ ,  $P_H$ , and  $P_O$  are the prices of regular gasoline, no. 2 heating oil, and crude oil, respectively.

The 3:2:1 crack spread has been institutionalized over the years as a way to measure the refinery margin. The use of the 3:2:1 crack spread is justified by the fact that among all finished products converted from crude oil in the refinery process, gasoline and distillate fuel oil are the two primary product classes. The relative proportion of these two products is approximately two barrels of gasoline to one barrel of distillate fuel. Together, gasoline and distillate fuel comprise about 80% of the refinery yield. The average refinery yield of finished motor gasoline is about 46% and has been stable over the 1993-2007 sample period (DOE).

The West Texas Intermediate (WTI) crude oil price, which is priced at Cushing, Oklahoma, is chosen to represent the crude oil price in this study. The reason is that WTI-Cushing is one of the most widely traded and price-transparent crude oils in the U.S. crude oil market.

We use the Petroleum Administration for Defense Districts (PADDs) to define refinery product markets. This market definition was formed during World War II for the purpose of administering oil allocation. The PADDs are still used by the Department of Transportation (DOT) and Energy Information Administration (EIA) for statistical and reporting purposes. The five regions are East Coast (PADD I), Midwest (PADD II), Gulf Coast (PADD III), Rocky Mountain (PADD IV), and West Coast (PADD V). These five geographically distinct regions are also very different in terms of their economic conditions, oil and petroleum characteristics, oil-related pipeline infrastructure, and local product supply and demand conditions.

Because of its high population density, the East Coast PADD I has the highest demand for refined products in the country, but it has very limited refinery capacity. Its regional demand is largely satisfied by the Gulf Coast and by foreign imports. The Midwest PADD II is distinct in its coexistence of a highly industrialized section and a rural agricultural section. It also leads the nation in ethanol production, mainly because of its leading role in corn production, the primary feedstock for ethanol production. For example, Iowa had 30 ethanol plants in operation by the end of 2007 and produces nearly 2.1 billion gallons of ethanol annually. Much of the crude oil used in the Midwest is piped in from the Gulf Coast and Canada. One place worth mentioning in this region is Cushing, Oklahoma, which is the major crude oil transportation hub for the Midwest.

The Gulf Coast region, including Texas, Louisiana, New Mexico, Arkansas, Alabama, and Mississippi, produces over 50% of the nation's crude oil and 47% of its final refined products. This region also serves as a national hub for crude oil and is the center of the pipeline system. The Rocky Mountain region, or PADD IV, has the smallest and fastest-growing oil market in the U.S., with only 3% of national petroleum product consumption. The West Coast region, PADD V, is the largest oil-producing and consuming region. This region's oil supply is independent of



other regions since it is geographically separated by the Rocky Mountains. In addition, the refinery market of this region is highly concentrated.

## **Data**

The gasoline price relative to that of crude oil is used as a dependent variable to measure ethanol's possible substitution effect on the gasoline price, while the 3:2:1 crack spread is employed as a dependant variable to quantify the effect of ethanol on the refinery profit margin. Figure 1 presents the relative gasoline to crude oil price over the 1995-2007 period. Figure 2 is for the 3:2:1 crack spread deflated by Producer Price Index (PPI) for crude energy material for five PADD regions over the same sample period. The PPI data are obtained from the U.S. Bureau of Labor Statistics.

The relative gasoline price is similar to crack spread in the sense that both are measurements of profitability of the refinery industry. The difference is that relative gasoline prices only account for the contribution of gasoline to the profit margin. It is employed in this study to quantify the substitution effect of ethanol production on gasoline prices. Relative gasoline prices and the refinery profit margin are mainly determined by similar explanatory variables. The explanatory variables included in this study are market demand and supply conditions, refinery capacity and utilization rate, market concentration and structure, unexpected supply disruptions, gasoline imports, seasonality, and ethanol production. Each of these chosen variables and its relationship with the relative gasoline price and refinery profitability is discussed in greater detail in this section.

## ***Crude and Product Market Conditions***

The gasoline price and refinery profitability are affected by the supply and demand balances of the crude market and product market. When the crude oil market has ample stocks, refinery profit should increase because of lower crude oil prices. Alternatively, when there are large stocks of gasoline and other refinery products, refinery profits should fall because of lower product prices. A tight product market will generate upward pressure on product prices even when there is an ample supply of crude oil. That is, product prices are bid up by more than any underlying cost increases. This upward movement relative to crude oil prices will be seen as an

increase in the relative price and crack spread. We use monthly crude oil inventory and gasoline inventory data collected by the EIA to represent the conditions in these two markets. The gasoline stock and crude oil stock data for the East Coast region from 1995 to 2007 are shown in Figures 3 and 4, respectively.

### ***Refinery Capacity and Capacity Utilization Rate***

Refinery capacity is a critical factor influencing the profitability of the refinery industry. Figure 5 presents the operable crude oil distillation capacity in the five PADD regions from 1995 to 2007. In this figure, refinery capacity is represented by monthly data of atmospheric crude oil distillation units (barrels per calendar day). Total refinery capacity increased by 13% over the past 12 years, with PADD III, the Gulf Coast, having the highest growth of 19%. The lowest increase in capacity occurred in the Midwest, with a 4% growth over the same period.

The monthly percent refinery capacity utilization rates for 1995 to 2007 for PADDs II, III, and V are shown in Figure 6. Here, refinery capacity utilization is based on gross input to atmospheric crude oil distillation units divided by the refinery operable distillation capacity. The average rate over five regions is 92%, which means that capacity utilization has increased significantly and refineries are running at high rates of utilization. Refinery capacity and its utilization rate are variables that will affect gasoline price and refinery profits via higher prices for products and possible increases in marginal costs.

### ***Market Concentration***

Mergers and acquisitions among refinery firms may potentially further reduce the competition in the refinery market, thus possibly leading to a higher refinery margin. To measure the level of market concentration, the Herfindahl-Hirschman Index (HHI) is commonly applied in the literature. The HHI of a market is calculated by summing the squares of the percentage market shares held by the respective firms as

$$HHI_t = \sum_{i=1}^{N_t} S_{it}^2$$

where  $S_{it}$  is the market share of a specific firm in the corresponding production market with total

firms of  $N_t$  at year  $t$ . A market with an HHI less than 1,000 is considered to be a competitive market; 1,000-1,800 to be a moderately concentrated market, and greater than 1,800 to be a highly concentrated market.

We constructed an HHI for the five PADD regions over the period 1995 to 2007, and we present this information in Figure 7. The HHI for the refinery market in PADD 1 increased from 1,558 to 2,335 from 1995 to 2007 and changed from a moderately concentrated to a highly concentrated market using Department of Justice definitions. Since much of this region's refinery product supply is from other regions, the impact of this increased concentration may be small. The refinery market in PADD II, the Midwest, suggests that this is a competitive market, although its HHI increased to 960 in 2007. Similar to the Midwest region, PADD III, the Gulf Coast, also has a competitive refinery market as of the end of 2007. The HHI for PADD IV, the Rocky Mountain region, decreased from 1,025 to 930, which suggests that its refinery market became less concentrated than before. The HHI for the PADD V, the West Coast region, increased from 914 to 1,155, and this refinery market changed its definition to a moderately concentrated market by 2007.

### ***Unexpected Supply Disruptions***

On August 29, 2005, Hurricane Katrina hit the U.S. Gulf Coast at New Orleans. On September 24, 2005, Hurricane Rita hit at the border between Texas and Louisiana. Both were category four storms when they did significant damage to the refineries' facilities and pipeline in the Gulf Coast region. Refinery operations were reduced by 1.8 million barrel/day in September and October 2005. Retail gasoline prices jumped by \$0.50 to over \$3.00 per gallon on a national average basis after Hurricane Rita. Prices were distinctly higher than before. In order to control for the effect of this event on the gasoline and refinery profit margin, we include dummy variables for September and October in 2005, when the disruptions were most severe.

### ***Gasoline Imports***

A significant share of total gasoline demand in the U.S. is met by imports. The net import share of total gasoline consumption in 2007 is 14%. Figure 8 presents U.S. finished motor gasoline imports from all countries over the period 1995 to 2007. Imports reached their highest level in

October 2005, the month after Hurricanes Katrina and Rita. Major sources of gasoline imports include Canada, Europe, and the Virgin Islands. A structural surplus in gasoline production in Europe means that gasoline production costs are lower when derived from foreign sources than they would be if the U.S. built and operated additional refinery capacity domestically. Growth in imports is expected to be tempered because of the increased use of domestically produced ethanol. Also, with increases in imported gasoline, refinery profitability is expected to be negatively affected.

### ***Ethanol Production***

Figure 8 presents the monthly ethanol production over the 1995-2007 period. There are 68 ethanol plants under construction or expanding. Iowa leads the nation with about 2 billion gallons of ethanol production capacity. Our hypothesis is that this additional production has had a negative impact on gasoline prices and on the margins of crude oil refiners.

### ***Seasonality***

The gasoline market is highly seasonal due to stronger demand in spring and summer. Gasoline price tends to gradually rise before and after summer. Demand for distillate fuel including heating oil and diesel fuel typically peaks in winter and thus has a counter-cyclical price pattern from gasoline. We include a set of monthly dummies to account for the seasonal pattern.

### **Estimation Method**

The regression model is specified as follows:

$$(1) \quad \pi_{it} = X_{it}'\beta + \varepsilon_{it} \quad i = 1, \dots, N; t = 1, \dots, T.$$

where  $\pi_{it}$  is the price of gasoline divided by the price of crude oil or the 3:2:1 crack spread of region  $i$  at month  $t$ , and  $X_{it}$  is the  $K$ -dimensional vector of explanatory variables described earlier.

There are several options for estimating equation (1), including pooled OLS regression and panel data models. The pooled OLS regression simply pools together data series for all PADD regions

and applies the ordinary least squares method. The OLS estimates of the standard errors may be highly inaccurate if the data exhibits heteroskedasticity and/or cross-sectional and serial correlation. The panel data models increase precision of estimates and allow us to control for an unobservable individual region's heterogeneity and temporal effects without aggregation bias.

The Hausman test for misspecification (Greene 2003, p. 301) is employed to help us select from two principal types of panel data models: the fixed effect model and the random effect model. Under the null hypothesis, the random effects estimator is consistent and efficient, while under the alternative, it is inconsistent. The random effect model is chosen if we fail to reject the null hypothesis. In the case of relative gasoline price (3:2:1 crack spread), the  $\chi^2$  test statistic was calculated at 26.92 (48.99) and significant at the 5% (1%) significance level. This suggests that the fixed effect estimator is consistent and asymptotically efficient in both cases.

Different specification tests are applied on the data set to better specify the panel data model. Applying the Wooldridge test for autocorrelation in panel data for the relative gasoline price (or crack spread) (Wooldridge 2002, p. 282), we get an F-test statistic of 917 (1,708), which is highly significant, and the null hypothesis of no first-order autocorrelation is rejected. Tests developed by Pesaran (2004) and Frees (1995) of cross-sectional independence are applied and both null hypotheses are rejected; this confirms the existence of cross-sectional correlation across regions.

Based on these diagnostic results, we used a fixed effect panel data model with correction for first-order serial correlation. We also estimated a feasible generalized least squares (FGLS) model with generalized error structure to allow for the presence of AR(1) autocorrelation within panels, as well as for heteroskedasticity and cross-sectional correlation across panels. By using three alternative model specifications we hope to provide information on the robustness of the results.

The fixed effect model is specified as

$$(2) \quad \pi_{it} = \alpha_i + X_{it}'\beta + \varepsilon_{it} \quad i = 1, \dots, N; t = 1, \dots, T .$$

where  $\alpha_i$  represents the individual regional effect. The fixed effect model is typically estimated by the least squares dummy variable (LSDV) method (Greene 2003, p. 287).

The FGLS estimation method takes into account heteroskedasticity, and cross-sectional and serial correlation. The error terms can be written as

$$E[\varepsilon\varepsilon'] = \Omega = \begin{bmatrix} \sigma_{11}\Omega_{11} & \sigma_{12}\Omega_{12} & \dots & \sigma_{1N}\Omega_{1N} \\ \sigma_{21}\Omega_{21} & \sigma_{22}\Omega_{22} & \dots & \sigma_{2N}\Omega_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1}\Omega_{N1} & \sigma_{N2}\Omega_{N2} & \dots & \sigma_{NN}\Omega_{NN} \end{bmatrix}$$

where

$$\Omega_{ij} = \begin{bmatrix} 1 & \rho_j & \rho_j^2 & \dots & \rho_j^{T-1} \\ \rho_i & 1 & \rho_j & \dots & \rho_j^{T-2} \\ \rho_i^2 & \rho_i & 1 & \dots & \rho_j^{T-3} \\ \dots & \dots & \dots & \ddots & \dots \\ \rho_i^{T-1} & \rho_i^{T-2} & \rho_i^{T-3} & \dots & 1 \end{bmatrix}$$

An FGLS panel data model is also called the Parks-Kmenta method (Kmenta 1986). This method consists of the following steps. Estimate equation (1) by regular OLS. Then use the estimation residuals to estimate assumed error AR(1) serial correlation coefficient  $\rho$ . Use this coefficient to transform the model to eliminate error serial correlation. Substitute  $\hat{\Omega}$  for  $\Omega$  using estimated  $\rho$  and  $\sigma^2$ , then obtain the FGLS estimator of  $\beta$  as

$$\hat{\beta}_{GLS} = (X' \hat{\Omega} X)^{-1} X' \hat{\Omega}^{-1} y.$$

### Analysis of Estimation Results

Using the relative gasoline price as the dependent variable, we get estimation results for the pooled OLS regression, a fixed effect panel data model, and a panel FGLS method; these are shown in Table 1. The corresponding estimation results for 3:2:1 crack spread are shown in Table 2.

In the case of the relative gasoline price, three estimation methods generate similar results. The only difference is that standard errors of coefficient estimates get bigger after taking into account cross-sectional and temporal autocorrelation, which in turn lead to a comparatively lower significance level for corresponding variables. Crude oil and gasoline inventories, refinery capacity, short-run supply disruption, and dummy variables for some summer months all significantly influence the relative gasoline price. Ethanol production has a considerably negative impact on the gasoline price, which is highly significant at the 1% level in all three estimation results. This indicates that over the sample period, ethanol has a significant substitution effect on gasoline. Evaluating at the sample mean, we find that the gasoline price is lowered by 39.5¢, 28.7¢, and 34.1¢ per gallon because of the substitution effect of ethanol.

For the 3:2:1 crack spread, the estimation results of the fixed effect and panel FGLS models are quite different from that of the pooled OLS regression. In addition, the pooled OLS regression model generates highly significant estimates for all explanatory variables except the dummy variables for January, February, and November. As previously mentioned, ignoring cross-sectional and serial correlation as well as individual heterogeneity typically leads to highly inaccurate standard error estimation; i.e., the significance estimation results are not reliable. Hence, we focus on the fixed effect and panel FGLS estimation results.

From these two sets of estimates, all the explanatory variables have intuitively correct signs. First, the profitability represented by the 3:2:1 crack spread presents a strong seasonal pattern. This is reflected by the fact that the dummy variables for months in the second and third quarters are all significant at the 1% significance level in the panel FGLS model and at the 5% level in the fixed effect model. Second, crude oil and refinery product market conditions, refinery capacity, ethanol production, and unexpected supply disruption significantly affect profit margins. For all five PADD regions, unexpected supply disruption, measured by dummies for Hurricanes Katrina and Rita, considerably increased profits in the months right after the occurrence. Gasoline imports and the HHI are found not to have statistically significant effects on crack spread nationally. Finally, we find that ethanol production generates negative pressure on crack spread

over the sample period. For the fixed effect and panel FGLS models, the marginal effect of ethanol production on the crack spread is estimated to be -0.000073 and -0.000077, respectively.

### ***Regional Analysis***

Pooling cross-sectional and time-series information provides more accurate estimation results. However, it is instructive to analyze the time-series data of each region individually. Each PADD region has unique supply and demand conditions of crude oil and refinery products, different market structures, and different ethanol production and usage. The effects of explanatory variables may differ considerably because of region-specific factors.

We apply regular OLS regression on individual region's monthly data series over the period 1995 to 2007. The estimation results for the relative gasoline price and 3:2:1 crack spread are summarized in Tables 3 and 4, respectively.

From the estimation results for the relative gasoline price, ethanol production has a significant negative effect on gasoline prices in all regions. And the magnitude of the effect varies with PADD regions, ranging from -0.000041 to -0.000095. As expected, in PADD II, the Midwest region, ethanol production has the largest impact on the gasoline price with a coefficient of -0.000095. The substitution effect is highly significant and reduces the gasoline price by 39.5¢ on average over the sample period. The West Coast and East Coast experience similar negative ethanol impacts with estimates of -0.000056, which means that the corresponding gasoline price is lowered by 23.3¢. The Gulf Coast region, PADD III, has a slightly higher coefficient estimate of -0.000059, or, equivalently, a 24.6¢ reduction in gasoline prices. The Rocky Mountain region, or PADD IV, experienced the smallest downward gasoline price change, at 17.1¢, probably because of its comparatively low total gasoline consumption. These results tell us what would have happened had we removed the entire ethanol industry at the mean of the data set, and they are not marginal effects of removing one unit of ethanol capacity in each region.

From the estimation results of the profit margin for individual regions, effects of some explanatory variables differ considerably across regions. In PADD regions III and V, the HHI has a significant positive effect on refinery profit. This result suggests that higher market



concentration in these two regional markets results in refinery profits. We did not find this pattern in our panel data model. Similarly, gasoline imports have a significant negative effect on the profit margin in both East Coast and Midwest regions, possibly because these two regions are more heavily dependent on imported refinery products to meet their regional demand. Ethanol production has a significant negative effect on the refiner's profit margin in all five PADD regions.

## **Conclusions**

We employ pooled OLS regression, a fixed effect panel data model, and a panel FGLS estimation method to quantify the possible impact of ethanol on regular gasoline in the U.S. as a whole and in five regions of the U.S. The models control for gasoline imports, refinery capacity, capacity utilization rate, hurricanes, market concentration in the refinery industry, stocks, and seasonality.

Estimation results show that over the period 1995 to 2007, ethanol production had a significant negative effect of \$0.29 to \$0.40 per gallon on retail gasoline prices. The results suggest that this reduction in gasoline prices came at the expense of refiners' profits. These results are statistically significant across a range of model specifications and across all regions.

Results for individual U.S. regions indicate that the largest impact of ethanol on gasoline is found in the Midwest region where gasoline prices were reduced by 39.5¢ per gallon. The Gulf Coast region is found to have experienced a 24.6¢ reduction in the retail gasoline price, while for the West Coast and East Coast, the average price drop is about 23.3¢. The smallest impact, a 17.1¢ reduction, is found in the Rocky Mountain region, mainly because of its comparatively low gasoline consumption.

These reductions in retail gasoline prices are surprisingly large, especially when one considers that they are calculated at their mean values over the sample period. The availability of ethanol essentially increased the "capacity" of the U.S. refinery industry and in so doing prevented some of the dramatic price increases often associated with an industry operating at close to capacity. Because these results are based on capacity, it would be wrong to extrapolate the results to

today's markets. Had we not had ethanol, it seems likely that the crude oil refining industry would be slightly larger today than it actually is, and in the absence of this additional crude oil refining capacity the impact of eliminating ethanol would be extreme. In addition, the impact of the first billion gallons of ethanol on this capacity constraint would intuitively be greater than the billions of gallons that came later. We did try a quadratic term to pick up this effect, and it was not significant.

## References

- Asche, F., O. Gjolberg, and T. Völker. "Price Relationships in the Petroleum Market: An Analysis of Crude Oil and Refined Product Prices." *Energy Economics*, 2003: 25, 289-301.
- Borenstein, S., and A. Shepard. "Sticky Prices, Inventories, and Market Power in Wholesale Gasoline Markets." *Rand Journal of Economics*, 2002: 33, 116-139.
- EIA (Energy Information Administration). [http://tonto.eia.doe.gov/steo\\_query/app/papage.htm](http://tonto.eia.doe.gov/steo_query/app/papage.htm)
- Eidman, V.R. "Agriculture as a Producer of Energy." In *Agriculture as a Producer and Consumer of Energy*, edited by K.J. Collins, J.A. Duffield, and J. Outlaw. Cambridge, MA: CABI Publishing, 2005.
- Frees, E.W. "Assessing Cross-sectional Correlation in Panel Data." *Journal of Econometrics*, 1995: 69, 393-414.
- Geweke, J. "Empirical Evidence on the Competitive Effects of Mergers in the Gasoline Industry." Working Paper, University of Iowa, 2003.
- Girma, P.B., and A.S. Paulson. "Risk Arbitrage Opportunities in Petroleum Futures Spreads." *Journal of Futures Markets*, 1999: 931-955.
- Girma, P.B., and A.S. Paulson. "Seasonality in Petroleum Futures Spreads." *Journal of Futures Markets*, 1998: 581-598.
- Greene, W.H. 2003. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.
- Kmenta, J. *Elements of Econometrics (2nd Ed.)*. New York: Macmillan; London: Collier Macmillan, 1986.
- GAO (U.S. Government Accountability Office). *Energy Market: Effects of Mergers and Market Concentration in the U.S. Petroleum Industry*. United States Government Accountability Office Report, 2004.

Oladunjoye, O. "Market Structure and Price Adjustment in the U.S. Wholesale Gasoline Markets." *Energy Economics*, 2007: 1-25.

Pesaran, H. *General Diagnostic Tests for Cross Section Dependence in Panels*. Working Paper, University of Cambri.

RFA (Renewable Fuels Association). <http://www.ethanolrfa.org/industry/statisticsdgc>, 2004.

Szklo, A., R. Schaeffer, and F. Delgado. "Can One Say Ethanol Is a Real Threat to Gasoline?" *Energy Policy*, 2007: 35, 5411-5421.

Tokgoz, S., and A. Elobeid. "Understanding the Underlying Fundamentals of Ethanol Markets Linkage between Energy and Agriculture." Paper presented at the American Agricultural Economics Association annual meeting, Portland, OR, 29 July-1 August, 2007.

Tokgoz, S., A. Elobeid, J.F. Fabiosa, D.J. Hayes, B.A. Babcock, T-H. Yu, F. Dong, C.E. Hart, and J.C. Beghin. "Emerging Biofuels: Outlook of Effects on U.S. Grain, Oilseeds, and Livestock Markets." Staff Report 07-SR 101, Center for Agricultural and Rural Development, Iowa State University, 2007.

Vedenov, D.V., J.A. Duffield, and M.E. Wetzstein. "Entry of Alternative Fuels in a Volatile U.S. Gasoline Market." *Journal of Agricultural and Resource Economics*, 2006: 1-13.

Wang, M., M. Wu, and H. Huo. 2007. "Life-Cycle Energy and Greenhouse Gas Emission Impacts of Different Corn Ethanol Plant Types." *Environmental Research Letters* 2(April-June) Article 024001, available at <http://www.iop.org/EJ/abstract/1748-9326/2/2/024001/> (accessed January 2008).

Wooldridge, J. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press, 2002.

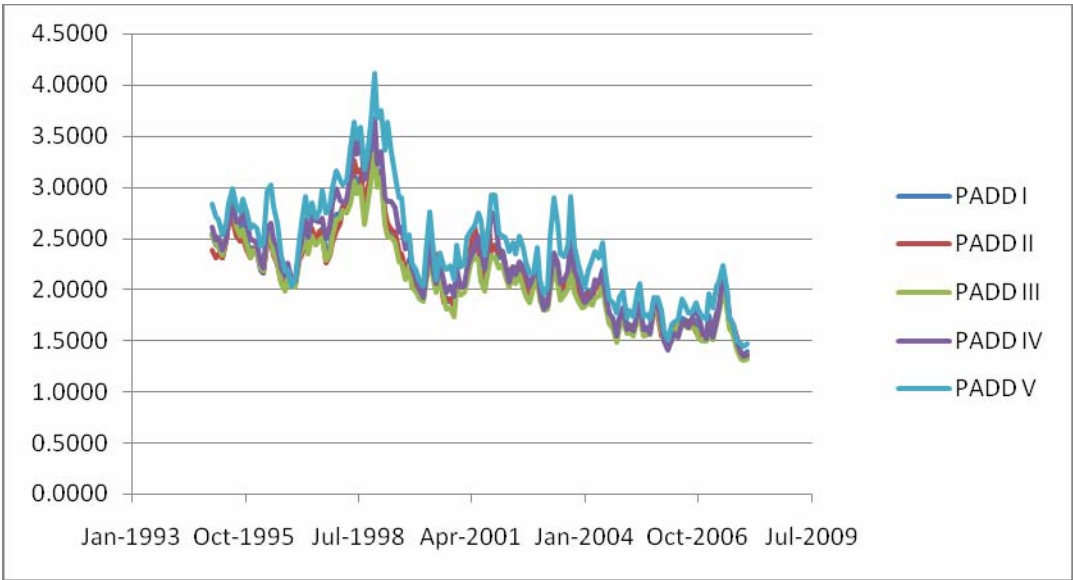


Figure 1. Relative Gasoline Price, 1995-2007

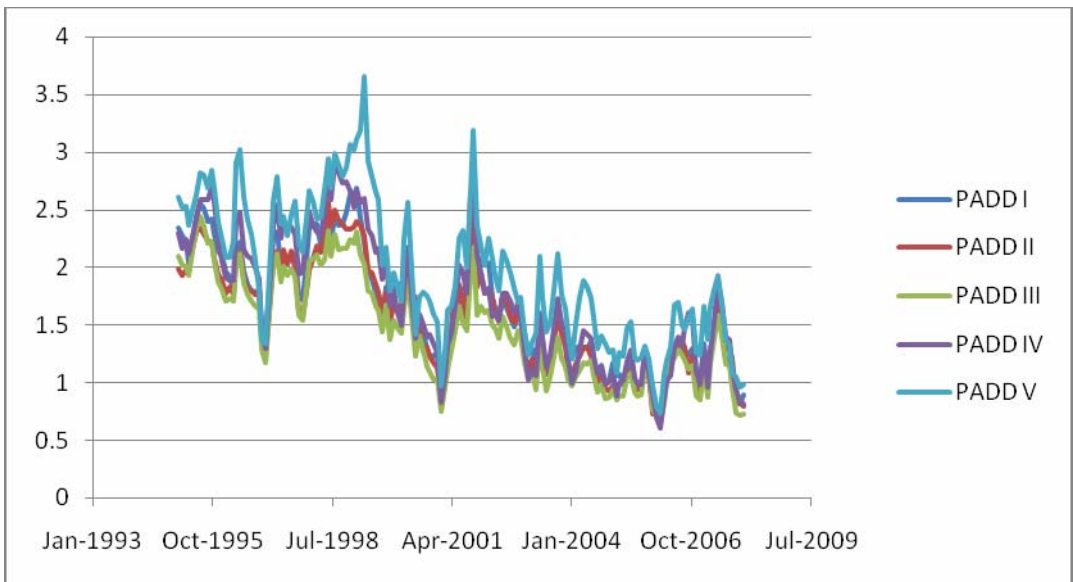


Figure 2. Deflated 3:2:1 Crack Spread, 1995-2007

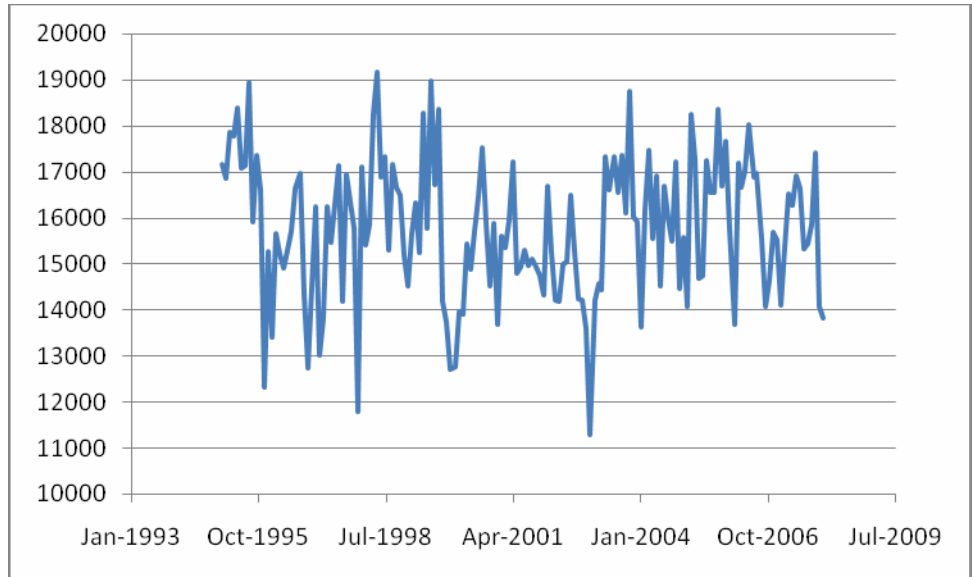


Figure 3. Month-End Oil Stock for PADD 1 (1995-2007)

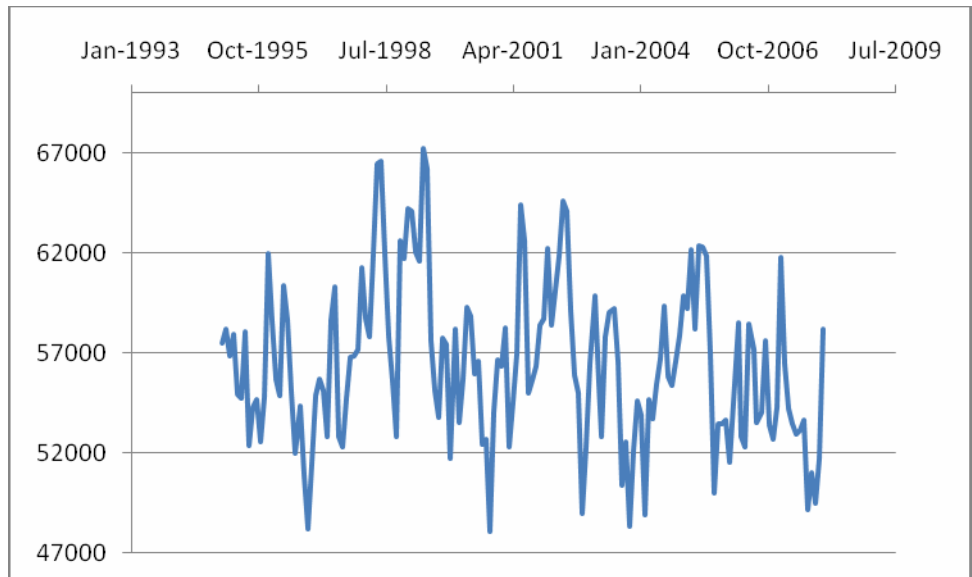


Figure 4. Month-End Gasoline Stock for PADD I (1995-2007)

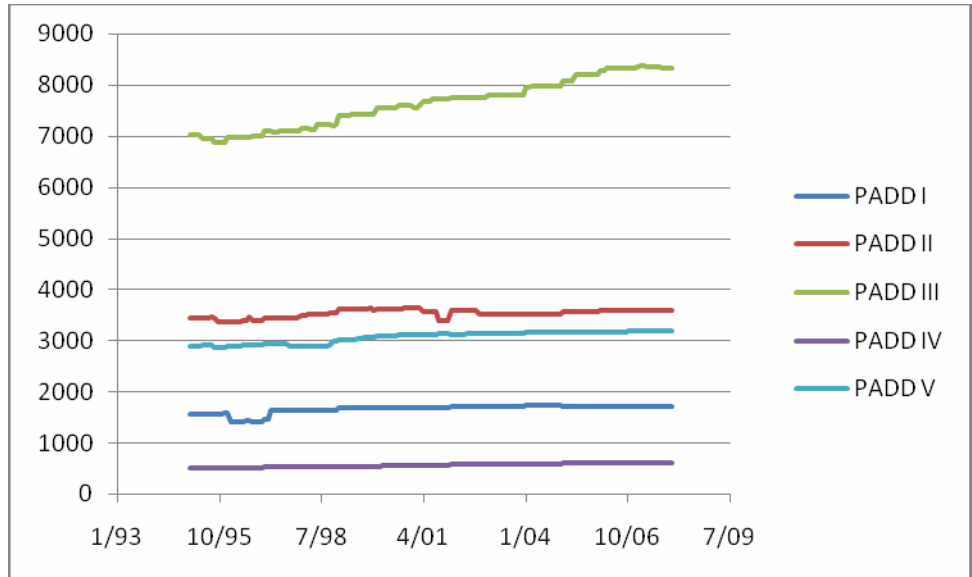


Figure 5. Operable Crude Oil Distillation Capacity (1,000 Barrels/Day)

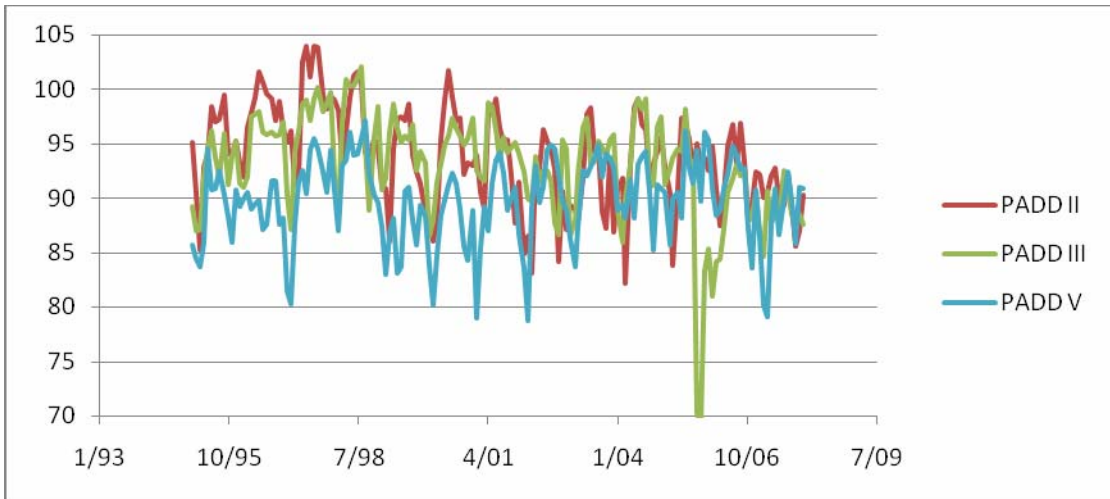


Figure 6. Percent Utilization of Refinery Operable Capacity (PADDs II, III, and V)

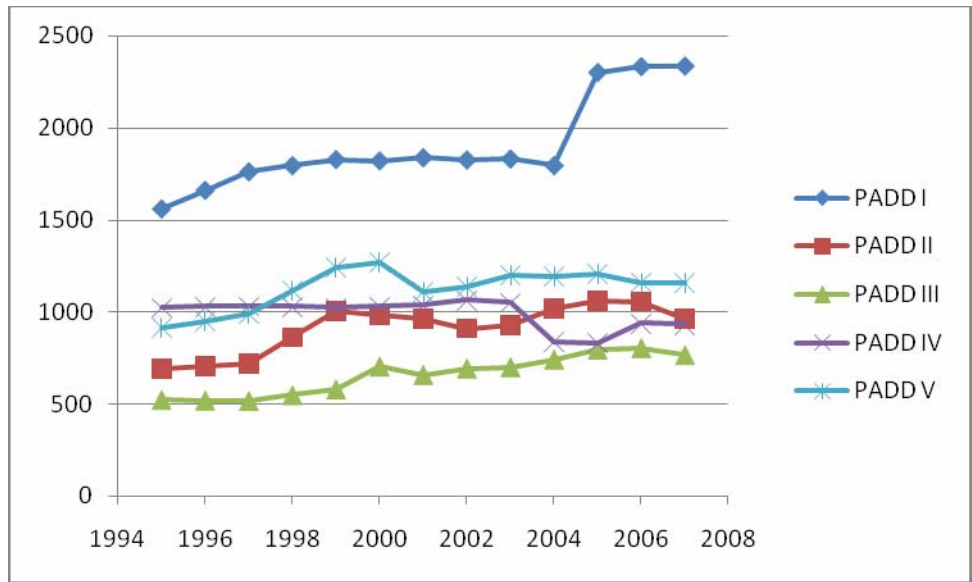


Figure 7. HHI, 1995-2007

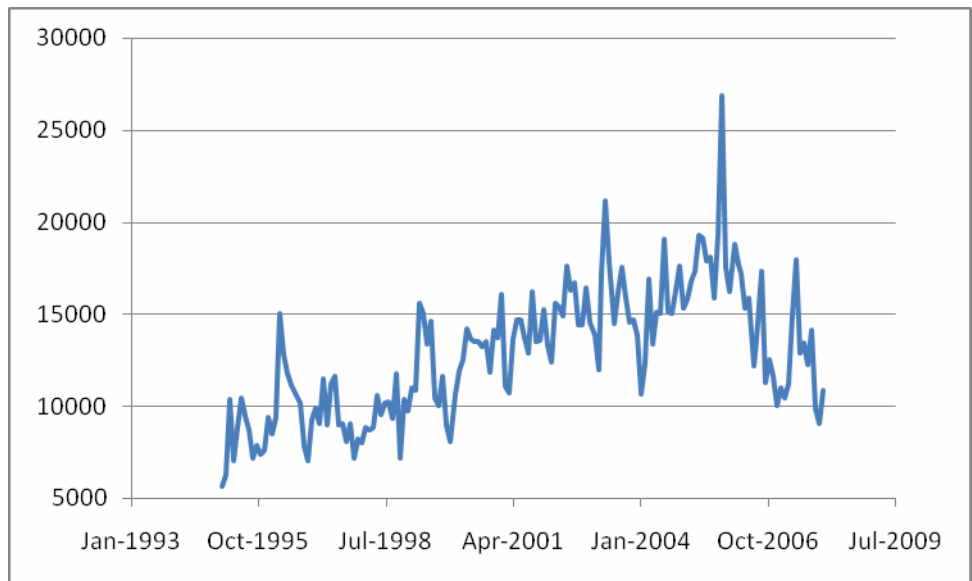


Figure 8. Monthly Motor Gasoline Imports, 1995-2007



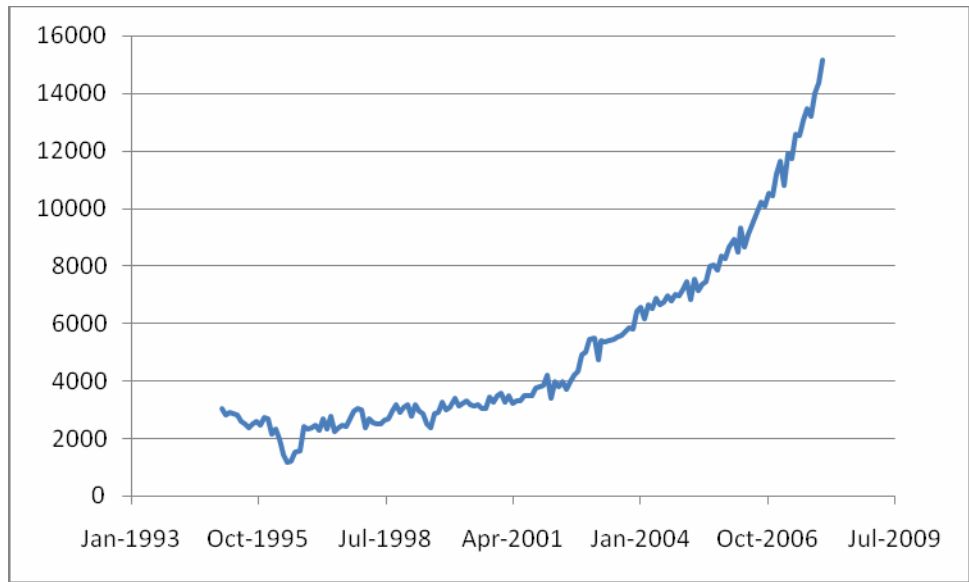


Figure 9. Monthly Ethanol Production, 1995-2007

Table 1. Regression Results for Pooled OLS, the Fixed Effect Model, and the Panel FGLS Method on Relative Gasoline Prices

Variable	Pooled OLS Regression		Fixed Effect Model with AR(1)		Panel FGLS Method	
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Oil stock	3.88e-6***	8.42e-7	1.97e-6**	9.28e-7	5.71e-7***	2.19e-7
Gasoline stock	-5.03e-6***	1.11e-6	0.000010***	2.70e-6	1.03e-6**	5.24e-7
Refinery capacity	-0.000099***	0.000029	-0.00038**	0.00019	-0.00040*	8.94e-6
Utilization rate	-0.0019	0.0028	0.00095	0.0015	0.00048	0.00041
Ethanol production	-0.000095***	3.96e-6	-0.000069***	0.000012	-0.000082***	0.000012
Supply disruption	0.32***	0.11	0.20***	0.055	0.20**	0.099
Gasoline import	-0.000037***	3.89e-6	7.37e-6***	2.75e-6	6.22e-6	4.78e-6
HHI	0.00028***	0.000062	-0.00019	0.00019	-0.000037	0.000025
January	-0.030	0.054	-0.047**	0.022	0.015	0.035
February	-0.083	0.054	-0.058**	0.028	0.00061	0.046
March	0.013	0.055	-0.0079	0.031	0.031	0.053
April	0.12**	0.055	0.055	0.035	0.069	0.058
May	0.19***	0.056	0.089**	0.036	0.099*	0.060
June	0.17***	0.055	0.10***	0.037	0.10*	0.060
July	0.11**	0.055	0.046	0.036	0.020	0.059
August	0.046	0.055	0.029	0.034	-0.0084	0.056
September	-0.0077	0.054	-0.012	0.031	-0.060	0.052
October	-0.014	0.054	-0.014	0.026	-0.069	0.046
November	-0.032	0.053	-0.0063	0.019	-0.05	0.034
Constant	3.12***	0.29	3.20***	0.076	2.46***	0.12
$R^2$	0.6014		$\rho = 0.87$			
Adjusted $R^2$	0.5914		F test	9.42		

\*\*\* 1% significance; \*\* 5% significance; \* 10% significance level.

Table 2. Regression Results for the Pooled OLS, the Fixed Effect Model, and the Panel FGLS Method on the 3:2:1 Crack Spread

Variable	Pooled OLS Regression		Fixed Effect Model with AR(1)		Panel FGLS Method	
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Oil stock	4.61e-6***	9.53e-7	7.27e-7***	3.19e-7	1.7e-6	1.18e-6
Gasoline stock	-4.56e-6***	1.26e-6	1.13e-6	7.55e-7	0.000011***	3.48e-6
Refinery capacity	-0.000015***	0.000032	-0.000063***	0.000012	-0.00039*	0.00022
Utilization rate	-0.015***	0.00032	-0.000066	0.00073	-0.00087	0.0019
Ethanol production	-0.000091***	4.49e-6	-0.000073***	0.000011	-0.000077***	0.000014
Supply disruption	0.32***	0.12	0.23*	0.13	0.13*	0.071
Gasoline import	-0.000062***	4.41e-6	-3.3e-6	6.0e-6	5.19e-6	3.53e-6
HHI	0.00026***	0.000069	-0.000027	0.000036	0.000079	0.00024
January	-0.058	0.06	0.00099	0.046	-0.075***	0.028
February	-0.075	0.06	0.022	0.059	-0.036	0.036
March	0.13**	0.062	0.14**	0.067	0.095**	0.039
April	0.30***	0.063	0.22***	0.072	0.18***	0.044
May	0.39***	0.063	0.23***	0.07	0.19***	0.046
June	0.36***	0.063	0.24***	0.074	0.20***	0.046
July	0.31***	0.062	0.17**	0.073	0.15***	0.045
August	0.28***	0.062	0.16**	0.07	0.18***	0.043
September	0.23***	0.061	0.12*	0.066	0.18***	0.039
October	0.19***	0.061	0.099*	0.059	0.18***	0.034
November	0.0022	0.060	-0.03	0.044	0.02	0.025
Constant	4.04***	0.33	2.06***	0.13	2.53***	0.10
$R^2$	0.6196		$\rho = 0.87$			
Adjusted $R^2$	0.6101		F test	3.98		

\*\*\* 1% significance; \*\* 5% significance; \* 10% significance level.

Table 3. Results for OLS Regression on Relative Gasoline Price with Individual PADD Regional Data

Variable	PADD I		PADD II		PADD III		PADD IV		PADD V	
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Oil stock	.000025*	.000015	.000012***	2.11e-6	3.89e-6***	9.38e-7	-7.88e-6	.000018	2.84e-6	5.62e-6
Gasoline stock	.000031***	6.48e-6	.000011	8.16e-6	.000024***	7.36e-6	.00015***	.000055	.000029	.000019
Refinery cap.	0.0048	0.00032	.00054	.00047	-.00012	.00021	-.0062***	.0023	-.0032***	.00074
Utilization rate	0.0051	0.0046	-.012	.0073	.010*	.0056	-.010	.0086	.0037	.011
Ethanol prod.	-.000056***	.000014	-.000095***	8.45e-6	-.000059***	.000014	-.000041***	.000020	-.000056***	.000015
Supply disrup.	.47**	.20	.19	.19	.54***	.20	.23	.26	-.069	.29
Gasoline import	-.000044***	7.58e-6	-.000012	9.59e-6	-5.07e-6	8.18e-6	-.000013	.000014	-9.37e-6	.000014
HHI	-.00029*	0.00017	-.00029	.00037	.00037	-.0030	.00051	.00053	.0019***	.00046
January	-.10	.097	-.047	.10	-.0057	.0084	-.14	.13	-.067	.15
February	-.11	.098	-.11	.10	-.024	.088	-.17	.13	-.0060	.15
March	.0079	.099	-.12	.11	-.006	.079	-.063	.12	.12	.14
April	.060	.10	-.013	.10	-.0044	.080	.15	.13	.22	.14
May	.023	.11	.11	.11	.019	.080	.26*	.13	.22	.14
June	.000052	.10	.17	.11	.0097	.078	.35**	.14	.22	.14
July	.018	.10	.065	.10	-.0013	.078	.33**	.15	.16	.14
August	.072	.10	.093	.10	.017	.079	.32**	.15	.11	.15
September	-.037	.10	.053	.098	-.047	.076	.23	.14	.074	.14
October	.035	.10	-.012	.099	-.056	.077	.18	.13	.096	.14
November	-.029	.097	-.025	.096	-.052	.075	.09	.12	.046	.14
Constant	0.17***	.67	.31	1.79	.048	1.20	5.56***	1.49	8.87***	2.79
$R^2$	.7109		.7119		.8229		.6610		.6415	
Adjusted $R^2$	.6705		.6717		.7981		.6136		.5915	

\*\*\* 1% significance; \*\* 5% significance; \* 10% significance level.

Table 4. Results for OLS Regression on 3:2:1 Crack Spread with Individual PADD Regional Data

Variable	PADD I		PADD II		PADD III		PADD IV		PADD V	
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Oil stock	.000012	.000018	.000014	2.3e-6	4.9e-6***	1e-6	.000017	.000019	.000012*	6.2e-6
Gasoline stock	.000038***	7.8e-6	.000015*	8.7e-6	.000029***	7.9e-6	.000023***	.000058	.000025	.000021
Refinery cap.	-0.000079	0.000038	-.00017	.00050	-.00028	.00023	-.012***	.0025	-.0029***	.000082
Utilization rate	-.01**	.0055	-.031***	.0078	-.012*	.0063	-.013	.0091	-.0065	.012
Ethanol prod.	-.000051***	.000017	-.00009***	9e-6	-.000039**	.000016	-4.2e-6	.000021	-.000047***	.000016
Supply disrup.	.59**	.24	.23	.21	.10	.21	.079	.27	-.24	.32
Gasoline import	-.000065***	9.1e-6	-.000029***	.00001	-.000019**	8.8e-6	-.000023	.000015	-.000019	.000016
HHI	-.00033	.24	-.00042	.00039	-.0023***	.000078	.00033	.00056	.0016***	.00050
January	-.098	.12	-.053	.11	-.13	.090	-.19	.13	-.11	.16
February	-.023	.12	-.098	.11	-.12	.10	-.17	.14	-.025	.16
March	.16	.12	-.069	.11	.05	.085	.044	.14	.22	.16
April	.22*	.12	.16	.11	.14	.086	.34**	.14	.41**	.16
May	.17	.13	.33***	.12	.16*	.086	.46***	.14	.36**	.16
June	.11	.12	.41***	.11	.15*	.084	.59***	.15	.36**	.16
July	.16	.13	.28**	.11	.15*	.084	.62***	.16	.36**	.16
August	.25**	.12	.37***	.11	.21**	.085	.68***	.16	.38**	.16
September	.14	.12	.32***	.10	.14*	.081	.63***	.15	.40**	.16
October	.20	.12	.18*	.10	.065	.083	.53***	.14	.34**	.16
November	.01	.12	.017	.10	-.053	.081	.21	.13	.068	.15
Constant	2.10***	.81	3.76***	1.91	3.52***	1.29	7.22***	1.58	8.04***	3.07
$R^2$	.7017		.7419		.8299		.7216		.6592	
Adjusted $R^2$	.6600		.7058		.8061		.6827		.6116	

\*\*\* 1% significance; \*\* 5% significance; \* 10% significance level.