Estimating Ethanol-Corn-Oil price relations in the US market with Smooth Transition Vector Error Correction Models.

A.J. de Jong

July 7, 2011

Abstract

This paper investigates price relations between ethanol, corn and oil. We find that a significant cointegrating relationship exists between corn and ethanol. Furthermore we use non-linear Smooth Transition Vector Error Correction Models to estimate long-run and short-run price behavior. With GR functions we visualize the effects of price shocks to the commodities on the prices of the other commodities. Our findings are in line with the idea that ethanol can have an upward price effect on corn prices.

Contents

1 Introduction

Problem definition The goal of this thesis is to show how ethanol is related to its main replacement - oil - and its primary resource - corn. This is clarified in terms of:

1. What kind of equilibria exist between prices of ethanol, oil and corn?
2. Which of the commodities is leading in a relation and which follows?
3. How does one commodity's price react to a price change of another commodity?
4. How long do these changes persist?

This is achieved by using cointegration and bivariate linear VEC and non-linear STVECM models. First of all we find long run price equilibria between commodities with cointegration tests. Then linear and non-linear VEC models are used to describe short and long run deviations from these equilibria over time and the speed with which adjustments are made. Finally, generalized response functions give an impression of price behavior over time as a result of a shock in one of the commodity's prices ceteris paribus. We only treat the bivariate case, because multivariate models can result in ambiguous outcomes that are hard to interpret.

Motivation One reason for researching this topic is its novelty. Bio-ethanol is still a relatively new commodity. Hardly as many papers have been written for ethanol as there have been for oil and corn. From those, only a few use advanced econometric methods and even fewer use data that contain the Global Financial Crisis of 2008-2009.

Another reason is found in the article of Serra et al. (2011). They state that the ethanol industry in the US was in the midst of an ongoing expansion in 2008, which was the end of their dataset. They recommended further research to see whether the price patterns they discovered maintain over time. Although we use a different kind of data, it will be insightful to see if we find similar patterns or not.

Relevance As oil resources become more scarce and prices keep rising, alternative fuels will become more and more important. Understanding the price relations between ethanol, oil and corn is therefore very useful for traders, industries and policy makers. It is especially helpful for hedging strategies.

An important political and sociological aspect is the so called 'ethanol-corn-sugar' nexus. Corn and sugar are raw materials for the production of ethanol. An increase in ethanol demand is therefore likely to cause increase in sugar/corn demand. Thus cause upward price effects and increase scarcity of food. This has resulted in a worldwide debate on increasing food prices due to biofuels.

Literature review In this section we discuss related papers that also make use of linear and non-linear models to find price relations between ethanol, corn and oil.

Serra et al. (2011) performed a study on price relations and transmission patterns in the US Corn-Ethanol-Oil-Gasoline market. Their goal was to contribute to the debate of increasing food prices. Their used monthly prices from the National Agricultural Statistics Service and that covered a period of 18 years from 1990 to 2008. They use STVEC models and GR functions to quantify the magnitude, timing and duration of price adjustments. They find a strong link between corn and energy prices: an increase in energy prices is followed by an increase in corn prices. Increased corn prices lead to increased ethanol prices. Vice versa only holds for the short run. In conclusion they find that ethanol prices reduce as gasoline increases, but if ethanol increases gasoline does so as well.

Very recently a paper was written on the biofuel nexus by Hammoudeh et al. (2011). They performed research on price relations between corn, oil, ethanol, soybeans, sugar and open interest rate. The research uses day-to-day data for the period June 2, 2006 to January 13, 2011, covering a period of 5 years. In fact it uses the same Ethanol futures data this paper uses. To find price relations a linear VEC model is used. Ethanol is found to be leading in the price discovery process on the long run and a significant short-run adjuster.

Another interesting paper was written by Escalante et al. (2009). In this case price relations are established through cointegration, VECM, IR and Multivariate GARCH models. They find that shocks in oil and ethanol can in the short-run increase agricultural commodity prices, but in freely operating markets this effect should die out. Surprisingly there are no long-run relations among ethanol, oil and gasoline over recent years. The dataset covers a period of 18 years from March 1989 to December 2007 with weekly prices.

Balcombe and Rapsomanikis (2008) investigate non-linearities in the sugar-ethanol-oil nexus and conclude that non-linear adjustment is present in ethanol-oil and sugar-oil, but this is not the case for ethanol-sugar. They compare

VAR, TVECM, STVECM and Markov Chain models. Their dataset is made up of weekly prices covering 6 years from 8 July 2000 to 20 May 2006.

Overview Section 2 holds a description the commodities ethanol, corn .and oil. Furthermore the source and format of the dataset is presented, including some choices on how we have treated the dataset. Section 3 is completely devoted to the methods used: cointegration, VAR, VEC and the nonlinear STVECM. In the results section (4) the outcome of the different approaches is presented, as well as an discussion. In the last section we summarize the findings in this paper.

2 Ethanol, corn and oil

2.1 Some basic facts

Ethanol Ethanol is one of the most popular bio-fuels. It can be made from sugars contained in plants like sugarcane, potato, cassava and corn. The net amount of energy released can vary. Sugarcane has by far the highest energy yield. In Brazil sugarcane is the main production resource, whereas in the US corn is used the most.

Both in the US and the EU there has been a substantial increase in the production of ethanol in recent years (Figure 1). Even more, in the US ethanol has shown almost exponential growth over the past twenty years. Although ethanol is not as commonly traded as oil, there are various markets where ethanol can be obtained at spot and future prices. For instance, CBOT, NY, Rotterdam and NYMEX.

Corn Corn (also known as Maize, but in the US commonly referred to as corn), is an agricultural product. The US is the largest producer of corn, then followed by China and Brazil. Corn is of course far more commonly traded than ethanol. Many future contracts are available.

Oil Crude oil, or simply oil, is the number one fuel source in the world. It is sometimes referred to as the 'Mother of all Commodities' because it is important in the production of all sorts of materials (Wikipedia, 2011b). Crude oil is heterogeneous in the sense that there is no such thing as a single oil type. The industry classifies oil by its geographic origin (e.g. Brent, Oman, Intermediate), its density (Light or Heavy) and its sulfur content (sweet if little, sour if much). This produces the names we recognize from financial markets, such as Light Sweet Brent Oil.

Annual Fuel Ethanol Production by Country (2007–2009) [2][52] Top 10 countries/regional blocks (Millions of U.S. liquid gallons per year)				
World rank	Country/Region	2009	2008	2007
1	United States	10,750.00	9,000.00	6,498.60
2	Brazil	7,264.73	7,053.39	5,943.87
3	European Union	1,039.52	733.60	570.30
4	China	541.55	501.90	486.00
5	Thailand	435.20	89.80	79.20
6	Canada	290.59	237.70	211.30
7	India	91.67	66.00	52.80
8	Colombia	83.21	79.30	74.90
9	Australia	56.80	26.40	26.40
10	Other	247.27		
	World Total	20,221.83	17,916.48	14,026.37

Figure 1: Annual Fuel Ethanol Production by Country

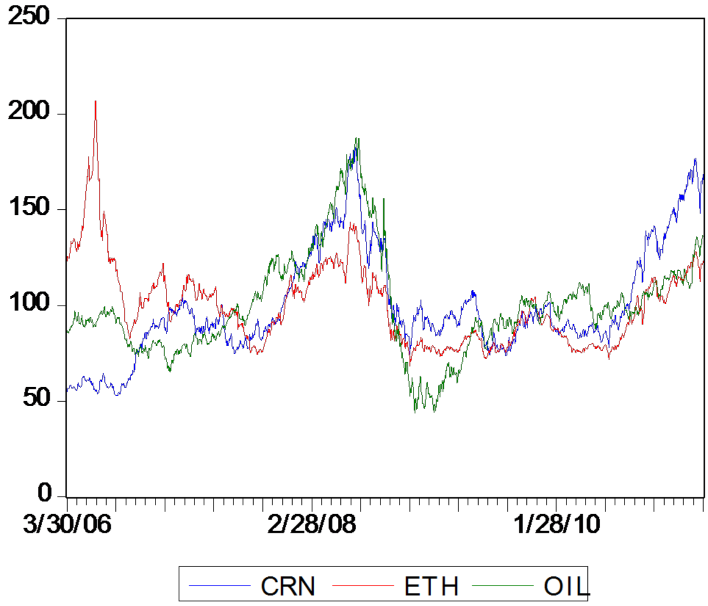

Figure 2: Daily Settlement Prices Ethanol, Corn and Oil

2.2 Description of the dataset

This research uses daily settlement prices of futures on Ethanol, Corn and Oil. These were obtained from the Datastream Databank:

- Ethanol Futures, Continuous series from Chicago Board of Trade;
- Light Crude Oil Futures Continuous series from New York Merchantile;
- Corn, Continuous series from Chicago Board of Trade.

The Datastream continuous series are equivalent to a 1-month future series. For consistency reasons and comparison with Serra et al. (2011) we have chosen explicitly for US time series. Figure 2 shows a plot of the time series from 30-3-2006 to 23-03-2011. The Global Financial Crisis is clearly recognizable in the graph as well as the recent increase in oil prices.

Trimmed data because of MTBE ban Ethanol demand has increased a lot from 2000 to 2011 due to regulatory changes. Especially the ban on Methyl-Tertiary Butyl Ether (MTBE) in the spring of 2006 has caused a very significant increase, which has great influence on parameter estimation. As the spring is the start of the dataset, we have trimmed it slightly to begin in 1-9-2006, where the effect of the MTBE ban has diminished.

The correlations in table 1 give a first impression on price relations.

	Corn	Ethanol	Oil
Corn	1.00	0.78	0.73
Ethanol	0.78	1.00	0.65
Oil	0.73	0.65	1.00

Table 1: Correlations among the commodities

The correlations are positive and relatively high. Ethanol and corn have the highest correlation, followed by ethanol-oil. The connection oil-corn has lowest correlation.

3 Methods

3.1 Cointegration

The derivation of statistical properties of time series is often based upon the assumption that series are stationary. When this is not the case, the statistical tests are no longer valid. A good method to transform a non-stationary stochastic series into a stationary one is by differencing. For instance, by differencing the level of stock prices one obtains a series of growth rates, which - usually - is stationary. Through differencing the interpretation of levels is lost however.

Another method is cointegration. The concept of cointegration was first introduced by Clive Granger Granger (1981). He states that when two time series x_t and y_t are both integrated of order 1, but have similar disturbances ϵ_t then $z_t = c + y_t - \alpha * x_t$ can be a stationary process (integrated of order 0). For instance, stock A and B could be individually non-stationary stochastic series, but the difference between the two could be stationary. This is also economically intuitive in the sense that cointegrating equations describe long run equilibria between variables.

This concept has turned out to be extremely important in the analysis of economic time series. Later on Johansen (1990) and Johansen (1988) introduced the necessary statistical tests for estimating and testing cointegration equations.

3.2 Linear VECM

In order to understand the concept of non-linear VEC models, it is helpful to describe the linear models first. The Linear Vector Error Correction Model (VECM) belongs to the family of autoregressive models. A well known autoregressive model is the so called AR(p) model which is given by: $x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + ... + \phi_p x_{t-p} + \varepsilon_t$. This concept has proven very useful and can easily be extended to a multivariate case, such that:

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \tag{1}$$

The above equation is an example of a VAR(1) model, where 1 stands for the number of lagged endogenous variables. More lags can be added as well as exogenous variables.

From the VAR model it is only a small step to get to the VEC model. We have seen the idea of cointegration where stationary series are created from non-stationary series through cleverly weighted subtraction. We have also seen the multivariate VAR model in which we can model multiple time series at once. If we combine these two concepts we immediately get the VEC. For a bivariate case for instance we have:

$$\begin{bmatrix} \Delta x_{1,t} \\ \Delta x_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{0,1} \\ \phi_{0,2} \end{bmatrix} + \begin{bmatrix} \alpha_1(z_{t-1}) \\ \alpha_1(z_{t-1}) \end{bmatrix} + \sum_{j=1}^{p-1} \begin{bmatrix} \phi_{j,1,1} & \phi_{j,1,2} \\ \phi_{j,2,1} & \phi_{j,2,2} \end{bmatrix} \begin{bmatrix} \Delta x_{1,t-j} \\ \Delta x_{2,t-j} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \tag{2}$$

where:
- α denotes the speed of adjustment;
- $z_{t-1} = x_{1,t} - \beta_1 x_{2,t} - \beta_0$ the cointegrating equation.

The cointegrating equation denotes the deviation from the long run equilibrium with estimated β_1 and β_0. A higher value of α equals to a bigger adjustment towards equilibrium per time period. Also note that the equation now is given in differenced form, that is $\Delta x_{1,t}$ instead of $x_{1,t}$. ΦX_{t-1} still captures the short run effects of price changes, just as in the VAR model. We can estimate the parameters with plain OLS.

If two time series are indeed cointegrated, then there exists a long run equilibrium between these series.If at any moment in time the models is in disequilibrium, that is z_{t-1} is different from zero, then the model corrects itself towards equilibrium through the effects that α_1 and α_2 have on ΔX. Hence, the name Vector Error Correction model.

3.3 Smooth Transition Vector Error Correction (STVECM)

Although the linear VEC model is already a very powerful tool, it can be extended for various reasons. Here we discuss the addition of smooth regime changes. For a normal VEC model the parameters must remain fixed over the entire sample. It therefore assumes a constant speed of adjustment, constant level and constant slope parameters. But if we look at the ethanol-corn-oil case, it is well possible that some parameters actually changed over time. This is also known as regime switching. Economic explanations for this behavior are found in bans, market rigidities, transaction costs, adjustment costs, market power and risk in the ethanol market (Serra et al. (2011)). Some of these economic causes will induce smooth changes over time, rather than instant regime switching. This is all possible with STVECM.

3.3.1 Model specification

Consider the VEC model with 1 lag and suppose we use the left hand side of (2) as an abbreviation of the model. That is ΔX. Then it quite easy to understand the STVECM. First we describe a switching VEC model consisting of 2 VEC models and an indicator function:

$$\Delta X = \Delta X_{Reg1} I[s_t > 0] + \Delta X_{Reg2} I[s_t < 0] \tag{3}$$

where:
- ΔX contains the differenced time series
- ΔX_{Regi} contains VEC model with parameter estimates corresponding to regime i
- $I[s_t > o]$ is an indicator function which is 1 for $s_t > 0$ and 0 otherwise
- s_t is a transition variable. For $s_t > 0$ Regime 1 is active, for $s_t < 0$ Regime 2 is active.

This reveals the basic concept of STVECM. As you can see there are 2 linear VEC models combined with 2 indicator functions. For positive values of s_t we get regime 1, for negative values we get regime 2. An economic interpretation to this is that prices of economic goods can respond differently to positive and negative disequilibria.

3.3.2 Transition function

In the example above the switching behavior is quite simple and instantaneous. The model can experience a regime switch from one instant to another. To allow for smooth transition we need to replace the indicator function with a function that maps every input variable to the interval $[0, 1]$ and does so in a smooth and monotonically increasing fashion. (This concept is also incorporated in probit and logit models.)

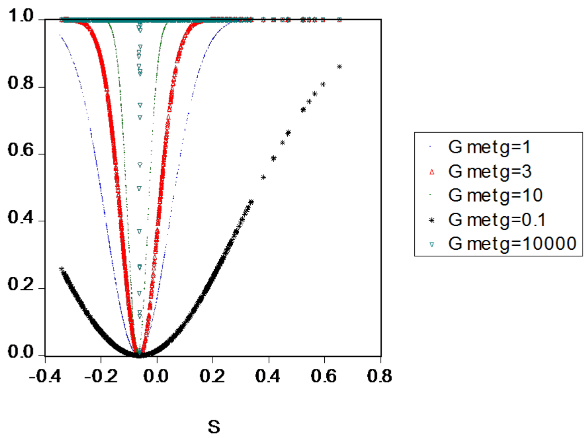

Figure 3: Response Functions VEC models

A popular choice is the first order logistic function:

$$G(s_t; \gamma, c) = (1 + \exp -\gamma(s_t - c))^{-1}, \gamma > 0 \tag{4}$$

where:
- γ is a shape parameter for the logistic function;
- s_t is the transition variable;
- c is a threshold level parameter.

Applied to (5) the result is:

$$\Delta X = \Delta X_{Reg1}(1 - G(s_t; \gamma, c)) + \Delta X_{Reg2}(G(s_t; \gamma, c)) \tag{5}$$

For this paper we actually use the exponential transition function:

$$G(s_t; \gamma, c) = 1 - \exp\{-\frac{\gamma(s_t - c)^2}{\sigma^2(st)}\}, \gamma > 0 \tag{6}$$

Shape parameter γ $G(s_t; \gamma, c) \to 1$ for both $s_t \to -\inf$ and $s_t \to \inf$ and $G(s_t; \gamma, c) = 0$ for $s_t = 0$. In other words: γ is a measure for the smoothness of the transition. For large γ we get abrupt regime switching behavior. For γ close to zero we get very smooth regime switching. Figure 3 shows transition function for different values of γ.

As the switching becomes abrupt for γ large, we can interpret γ as a goodness of fit parameter for the non-linear model. For very large gamma, one should try different specifications of the STVECM (transition function, transition variable, number of lags, etc.) or perhaps drop the assumption of non-linearity completely.

Interpreting γ economically is dependant on the specification of the transition variable. For instance, if the transition variable is defined as the amount of disequilibrium, than a large γ will narrow the bandwidth on which the transition variable can influence the value of transition function. So small deviations from equilibrium could lead to a different regime very quickly. A small γ yields a very large bandwidth, so only very distant disequilibria states lead to pure regime 2. Between that the model mixes between the two regimes. In short, γ determines *how fast* the market responds *differently* to deviations from equilibrium.

Transition variable s_t Another important ingredient of the transition function is the transition variable s_t. This serves as indicator for determining the active regime at a moment in time. Different people have come up with different types of variables. The STAR as discussed in Terasvirta (1994) uses a lagged endogenous variable: s_t where $d > 0$ and integer. Alternatively van Dijk et al. (2000) one can choose exogenous variable ($s_t = z_t$), where z_t has explanatory power for the regime. Other possibilities are (non-linear) functions of lagged endogenous variables ($s_t = h(y_t, \alpha)$ or a linear trend model ($s_t = t$). For simplicity reasons we set s_t equal to the disequilibrium in the model at $t-1$, thus equal to CE_{t-1}. This way we can see how ethanol reacts to its own disequilibrium and whether the corn and crude oil market are affected by these equilibrium conditions.

Threshold parameter c Furthermore, one can choose to estimate c with NLS or give it a preset value. For instance when $c = 0$ then the result is a model that distinguishes between positive and negative deviations from equilibrium. One can also choose $c = \bar{s}$. In that case the model distinguishes between small and large states of disequilibria.

The model Having described all the necessary concepts of the STVECM in a somewhat stylized fashion, we now conclude with the model as it used throughout the research phase.

$$\begin{bmatrix} \Delta x_{1,t} \\ \Delta x_{2,t} \end{bmatrix} \left(\begin{bmatrix} \phi_{0,1} \\ \phi_{0,2} \end{bmatrix} + \begin{bmatrix} \alpha_{11}(z_{t-1}) \\ \alpha_{12}(z_{t-1}) \end{bmatrix} + \right) (1-G(s_t;\gamma,c)) + \left(\begin{bmatrix} \psi_{0,0} \\ \psi_{0,1} \end{bmatrix} + \begin{bmatrix} \alpha_{21}(z_{t-1}) \\ \alpha_{22}(z_{t-1}) \end{bmatrix} \right) (G(s_t;\gamma,c)) \tag{7}$$

3.3.3 Estimation and diagnostics

STVECM models can be estimated using Non-linear Least Squares. Whereas VAR is often available in any software package, STVECM has to be programmed separately. NLS is sensitive to the choice of start values. A grid search gives the optimal value for γ and c if necessary. Note that the model is only non-linear in these two values. Having set γ and c the model could also be solved with OLS.

Interpreting the STVECM model is as follows: Regime 1 is when disequilibrium is relatively small. Regime 2 is when distance from equilibrium is large.

3.3.4 Generalized Impulse Response Functions

For the linear case the IR functions could be derived analytically. For non-linear models this is no longer the case. The solution is using simulation to measure the effect of a shock in one of the variables. To this end we simulate 100 instances of the variables twice. Once with a shock δ to $y1$ and once without. We then compute the paths that each of these instances walk. Afterwards we take averages per time period and subtract the averaged instances with and without the shock δ. The result is the impulse response functions for variable y1 and y2 (where y1 could be ethanol and y2 corn for instance).

4 Results

This section shows and discusses the results of the linear and non-linear modeling of ethanol, oil and corn.

4.1 Preliminary analysis

For convenience we perform a log transformation on all three series. Because cointegration is based on the assumption that the difference of two I(1) series becomes stationary, we first perform a number of Augmented Dickey Fuller tests. These show that all levels contain unit roots, but that the first differences are stationary. This suggests that all series are I(1).

4.2 Cointegration

Standard Johansen cointegration tests reveal that ethanol and corn contain 1 cointegration relationship: $P_{t,eth} = 0.9765(0.00806)P_{t,crn}$, with the standard error in parenthesis. The coefficient is significantly different from 1, but there is no clear-cut economic interpretation to this. For the remainder of this research the cointegrating equation nr. 1 (CE1) is assumed to be 1-to-1, that is:

$$\text{CE1} : P_{t,eth} = P_{t,crn}$$

For ethanol and oil the presence of an cointegration relationship is not significant. If there is one, then it is given by $P_{t,eth} = 0.9881(0.0222)P_{t,oil}$. This time the coefficient is not significant from 1. Therefore it is assumed that:

$$\text{CE2} : P_{t,eth} = P_{t,oil}$$

ADF tests on CE1 and CE2 show that these no longer contain unit roots at 1% resp. 5% significance. Note that the existence of cointegrating equations among variables does not mean that there must be a causal relation (Serra et al. (2011)).

4.3 VEC Model

Table 2 shows the parameter estimates of linear VEC models for the ethanol-corn case and the ethanol-oil case. None of the estimates are significant at 10%. α_{CE1} is negative for Ethanol as expected. When the price of ethanol is higher than corn, the price of ethanol is corrected towards equilibrium. Consequently α_{CE1} for corn should be positive. This is not the case. This relation does hold true for ethanol and oil: both are corrected towards equilibrium. Oil adjusts fastest.

	Δ Eth	Δ Crn		Δ Eth	Δ Oil
c	0.0133	0.0006	c	0.0000	0.0000
α_{CE1}	-0.0028	-0.0030	α_{CE2}	-0.0019	0.0056
R^2	0.0005	0.0020		0.0006	0.0005

Table 2: Parameter estimates linear VEC model. No significant values at 10%.

4.4 STVECM Model

Outcomes for the ethanol-corn case Table 3 shows the parameter estimates of the STVECM model. Under regime 1 the ethanol and corn prices are very close to equilibrium, because of the way we have defined the transition function. Consequently we find that α for ethanol is not significantly different from zero. The positive sign is therefore not very worrying. α under regime 2 is significant at a 1% level and has negative sign. When the system

Regime 1	Δ Eth	Δ Crn		Δ Eth	Δ Oil
ψ	0.0030	0.0027	c	-0.0002	0.0070
α_{CE1}	0.0495	0.0370	α_{CE2}	-0.0127	*-0.0873
Regime 2	Δ Eth	Δ Crn		Δ Eth	Δ Oil
ψ	0.0002	**0.0019	c	0.0003	0.0007
α_{CE1}	***-0.0120	-0.0051	α_{CE2}	-0.0011	***0.0083
γ	*2.7987		**4.4476		
R^2	0.0103	0.0043		0.0011	0.0064

Table 3: Parameter estimates STVECM model - * = 10%, ** = 5%, *** = 1% significance

is in disequilibrium, this causes the ethanol prices to be corrected towards those of corn.

The α of corn in regime 1 has positive sign. This is consistent with adjustment towards equilibrium. It is however not significant. Under regime 2 the adjustment speed α is not significant, but ψ is at 5% level. This causes corn prices to increase in case of disequilibrium. Because $P_{t-1,eth} - P_{t-1,crn} = s_t$ and the transition function moves towards 0 as $s - c$ becomes large this in fact has the same effect as a positive α.

With γ set at 2.799 the transition function is moderately smooth. Figure 4 shows that the transition takes place between values of -0.3 and +0.2 for s. γ is only significantly different from zero at the 10% level. R^2 is relatively high for explaining ethanol prices and still reasonable for corn prices. If we ignore the not significant values of α we can conclude that ethanol is leading in terms of price adjustment.

Outcomes for the ethanol-oil case For the ethanol-oil relation we find α's for ethanol price changes under regime 1 and 2 both insignificantly different from zero, as well as the constant terms. The sign of the α's is in line with the idea of price adjustment towards equilibrium. Now the α's for the price change are both significant.

Strangely again the sign of α under regime 1 causes disequilibrium. The sign of regime 2 does lead to equilibrium.

The explanatory power for price changes in ethanol is much less. For oil it is actually larger than that of corn in the ethanol-corn case and reasonably close to 1%, which is considered high for day-to-day data. γ is larger than it was for the ethanol-corn case. Figure 4 shows the smoothness of the transition function.

Figure 6 shows the evolution of the transition function of CE1, which is

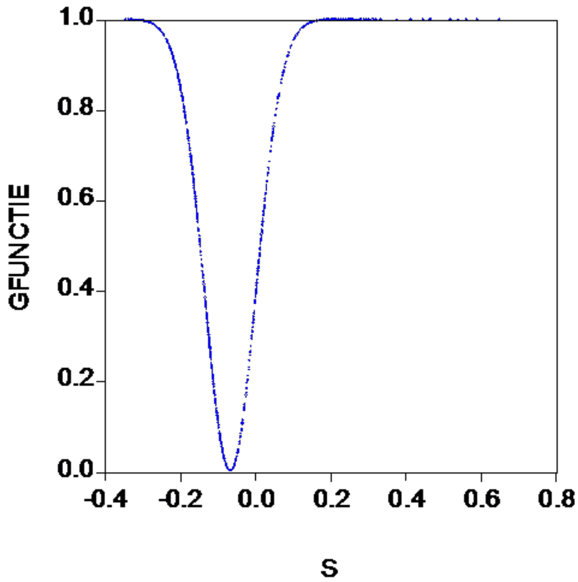

Figure 4: Scatter plot of s_t and G for ethanol-corn

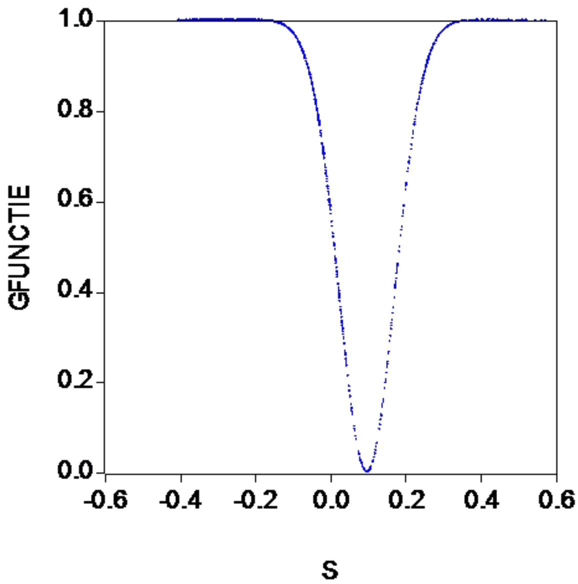

Figure 5: Scatter plot of s_t and G for ethanol-oil

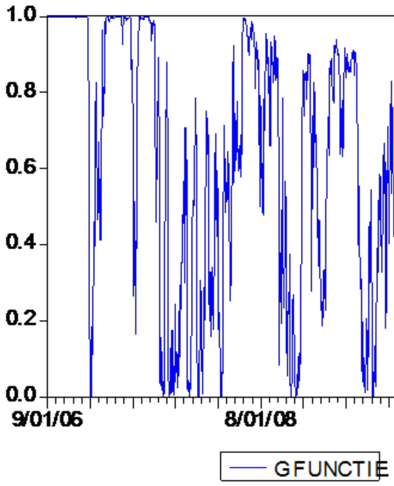

Figure 6: Transition function voor CE1

equated as $s_t = P_{t-1,eth} - P_{t-1,crn}$. Comparing this with figure 2, we see that G=1 roughly corresponds with periods where ethanol and corn are far apart. These moments correspond with the period just after the MTBE ban, the financial crisis and the rising oil prices. One could also say the regimes make distinction between periods of small and large volatility. Figure 7 shows the evolution for CE2. Again, we recognize periods that correspond to periods of smaller and larger international market volatility. Overall, when market volatility increases, the number of regime switches increases.

From figure 6 we see that the model transits 'evenly' over the entire sample period. A transition from 1 to 0 can take from one or two weeks to a couple of months and occurs quite often. This is not the case for ethanol-oil (figure 7). The function is mostly 1 and has 7 clearly distinctive transitions to 0 (figure 7). The Global Financial Crisis is recognizable from 2008 to 2009. For oil the transition function shows periods of ethanol-oil market instability, where G is close 1. Economic interpretation of figure 6 is not straightforward.

An important reason for the 'lack' of transitions is the specification of the transition function. c is equal to 0.11, whereas the mean of s for CE2 is -0.2. However, the model does not converge for $c = -0.2$. This could be due to the fact that preliminary testing concluded there is no significant cointegrating equation between ethanol and oil. (For the Ethanol-Corn the model did converge for c equal to average of s.). This means that for values of s close to 0.11 the model behaves like regime 1. For values of values of s larger than 0.3 or smaller then -0.20 the model behaves more like regime 2 (figure 6).

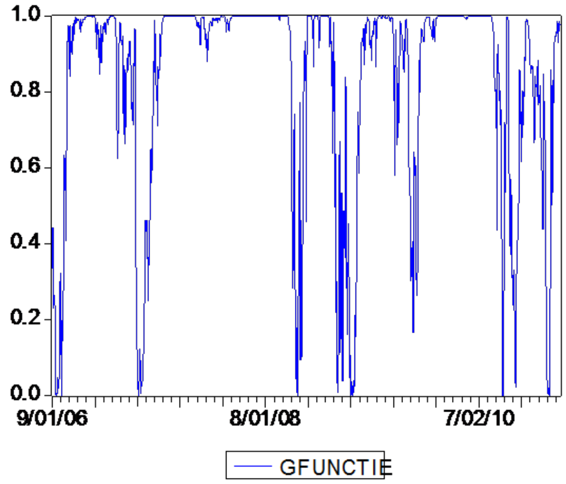

Figure 7: Transition function voor CE1

4.5 Generalized Response functions

Conditional expectations are constructed from the estimated models. Iteration is used to predict 100 ΔP_{t+h} forecasts: One set with and one without the initial shock to Y1. The result is averaged and then subtracted from each other. The result is the isolated effect of the initial shock. This is done for multiple states of the regimes. The corresponding Response Function graphs are given in Figures 8 and 9 in the Appendix. (Note that the scales differ among the different graphs!). We will discuss the results for the ethanol-corn and ethanol oil-case and highlight interesting features.

Ethanol-Corn case We first look at the situation where ethanol and corn are close to equilibrium. This is not exactly equilibrium, as we have set c equal to the average of s, rather than 0. So at G=0 we actually observe the case where prices are close to their average long run disequilibrium (which is still reasonably close to equilibrium). Interestingly we find that a shock δ to ethanol multiplies over time and has a lasting effect on both ethanol and corn of roughly the same size. Conversely, a δ sized shock reduces at first and then, through a feedback loop in ethanol, returns with the size of $\frac{1}{2}\delta$. This feedback loop could have an economic interpretation. As corn is a resource for ethanol, we would expect ethanol prices to increase over time through the production process. This may take some time.
When G=1 and ethanol price is above corn price, ethanol behaves very separate from corn. A δ increase dies out and has negligible effect on corn prices.

A shock in corn actually reduces both prices, but with small amounts ($< \frac{1}{2}\delta$). For G=0.5 and ethanol prices higher than corn prices, we find high correlation among the commodities. Shocks generate continuous price increases and are 1-to-1 transferred to the other commodity. For G=0.5 and corn prices higher than ethanol prices, we find that a shock δ in ethanol generates a lasting price increase in corn of size δ. Vice versa a shock to corn dies out and has a very small upward effect on ethanol.
For G=1 and corn prices higher than ethanol prices we see that both commodities behave quite separately again and shocks tend to die out. Apparently this is characteristic for disequilibrium circumstances.

Ethanol-Oil case Just as for the ethanol-corn case, we consider the outcomes at different states of price (dis)equilibria. As described earlier the threshold c for the ethanol-oil case is set to 0.11. This is important for interpreting the GR functions.

For G=0 and s=0.10 (so s-c is roughly 0.00), we find that a shock δ to ethanol reduces at first and then increases quite fast (in less than 100 days) to level at 1.5δ. Oil follows this pattern almost 1-to-1. So ethanol has an increasing upward effect on oil.
A δ shock to oil simply levels at δ, whereas ethanol adjusts with an amount almost equal to δ. An ethanol price increase boosts both ethanol and oil prices to amounts larger than the original shock, but an oil boost simply causes ethanol to adjust to the original shock.
For G=1 and ethanol prices much higher than oil prices we have more or less the same result except that price increases due to ethanol are only δ and not 1.5δ. So again ethanol boosts both commodities prices and whereas oil only causes ethanol to adjust.
For G=0.5 and ethanol prices almost equal to oil prices the result is different. A δ shock to either ethanol or oil stimulates both commodities to take equilibrium between 0 and δ. So at equilibrium prices both commodities become price adjusters.
For G=1.0 and oil higher than ethanol prices ($s = -0.13$), a boost in ethanol is sustained and is followed by oil. Surprisingly a boost to oil is almost doubled over time and ethanol follows this result.

5 Conclusion

Problem definition and related literature In the introduction we presented the problem definition. Through this research we gain insight in the price relations between ethanol, corn and oil. We discussed related papers that use both linear and non-linear models. Comparison is not easy, as often the time intervals do not overlap. Hammoudeh et al. (2011) use a linear model and find that ethanol is leading in the price discovery process on the long run and a significant short run adjuster. Serra et al. (2011) use nonlinear models as well and find strong linkage between corn and ethanol.

Data In section 2 we provided basic information about the commodities. Furthermore we explained the origin of our data. The ethanol market endured a couple of significant changes over the last years. A very significant one is the MTBE ban in 2006. The data we use therefore captures only the last part of 2006 and runs up to March 2011. We examine two main relationships: ethanol-corn and ethanol-oil.

Equilibria We estimated the corresponding equilibrium. For both the results are stationary time series. For ethanol-corn the cointegrating equation is actually significant. Presence of an equilibrium between oil and ethanol prices cannot be statistically confirmed. Because the estimated coefficients are very close to 1, we continue assuming the relations are 1-to-1.

Long run price adjustments We estimated linear and non-linear VEC models. The linear VEC models yield no significant parameters for either the ethanol-corn or ethanol-oil case. According to the parameter estimates oil is the fastest price adjuster. According to the non-linear STVECM ethanol adjusts faster than corn in equilibrium and disequilibrium regimes. Compared to the linear model R^2 for ethanol is greatly improved (20x). For corn it is only doubled.For the ethanol-oil case we find that oil is the faster price adjuster. R^2 for oil improved greatly (13x) compared to the linear model. For ethanol it only doubled. We also found that for both cases regime G=1 in general coincides with periods of increased market volatility.

Reactions to price shocks For the ethanol-corn case both commodities react quite strongly and fast to an increase in ethanol prices in equilibrium state. When prices are far from equilibrium we find both behave quite separately. Between these states there is great potential for prices to increase due to a shock in either corn or ethanol. Both adjust strongly towards one another. Overall our findings are in line with the idea that ethanol prices can have upward effect on corn prices. Corn tends to follow.

For the ethanol-oil case we find different results. When ethanol prices are above oil prices in relation to the cointegration equilibrium, ethanol has an upward price effect on itself and on oil that was larger than the original shock. A shock to oil usually causes the oil price to increase with shock size and remain at that new level. Only when oil prices are above ethanol (with respect to equilibrium), we see that a shock to oil causes prices to keep increasing. Overall oil seems to behave more separate from ethanol then corn does. As a result it is not clear whether oil or ethanol leads on the short run.

References

Balcombe, K. and Rapsomanikis, G. (2008). Bayesian estimation and selection of nonlinear vector error correction models: The case of the sugar-ethanol-oil nexus in brazil. *American Journal of Agricultural Economics*, 90(3):658–668.

Escalante, C., Wetzstein, M., Zhang, Z., and Lohr, L. (2009). Ethanol, corn, and soybean price relations in a volatile vehicle-fuels market. *Energies*, 2(2):320–339.

Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16:121–130.

Hammoudeh, S., Sarafrazi, S., Chang, C., and McAleer, M. (2011). The dynamics of energy-grain prices with open interest. Working Papers in Economics 11/24, University of Canterbury, Department of Economics and Finance.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3):231–254.

Johansen, S. (1990). Maximum likelihood estimation and inference on cointegration - with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52(2):169–210.

Serra, T., Zilberman, D., Gil, J. M., and Goodwin, B. K. (2011). Nonlinearities in the u.s. corn-ethanol-oil-gasoline price system. *Agricultural Economics*, 42(1):35–45.

Terasvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89(425):208–218.

van Dijk, D., Franses, P. H., and Terasvirta, T. (2000). *Smooth Transition Autoregressive Models*, volume EI2000-23/A. Erasmus School of Economics, Rotterdam.

A Appendix

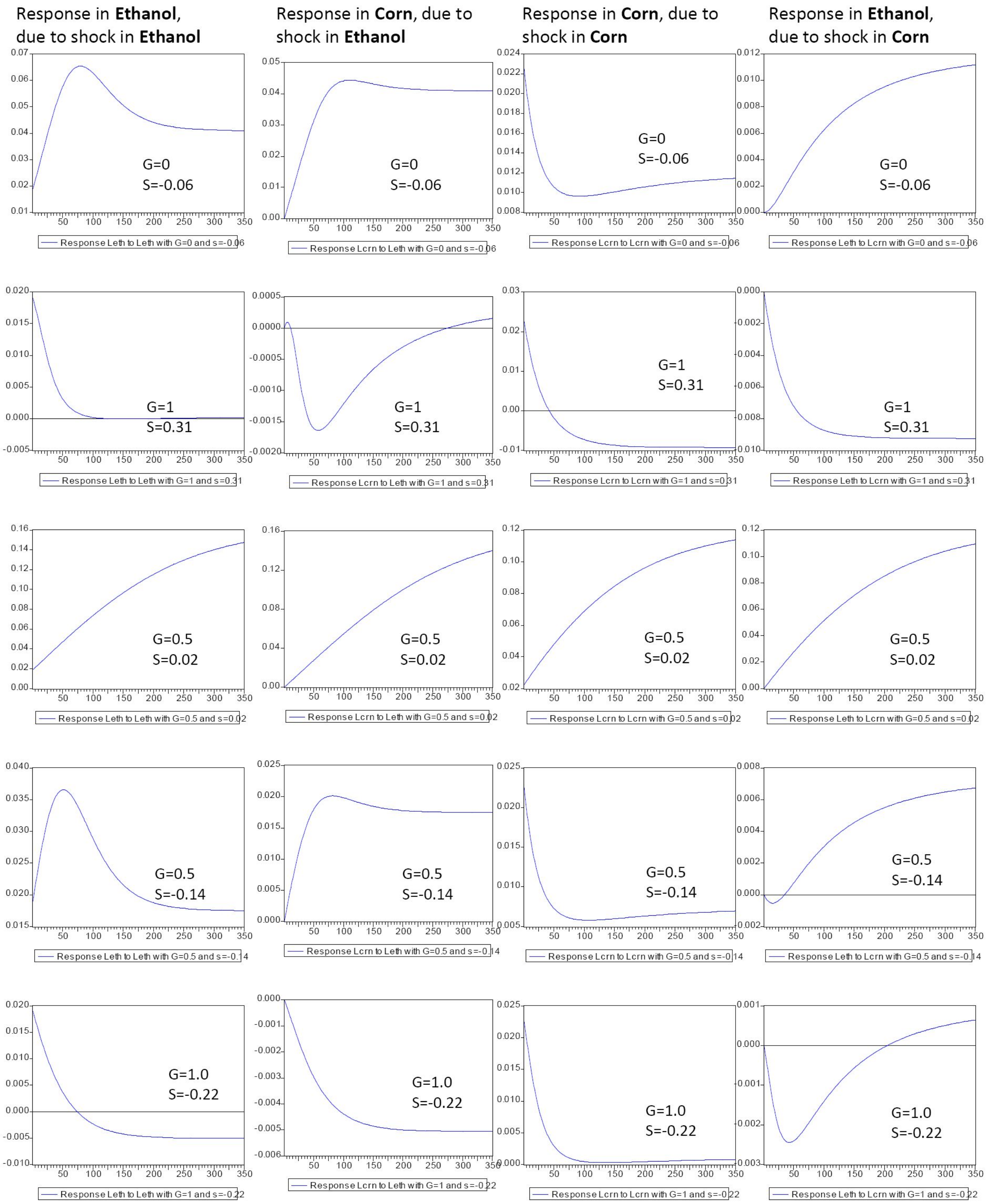

Figure 8: Generalized Response Functions Ethanol-Corn case with STVECM models

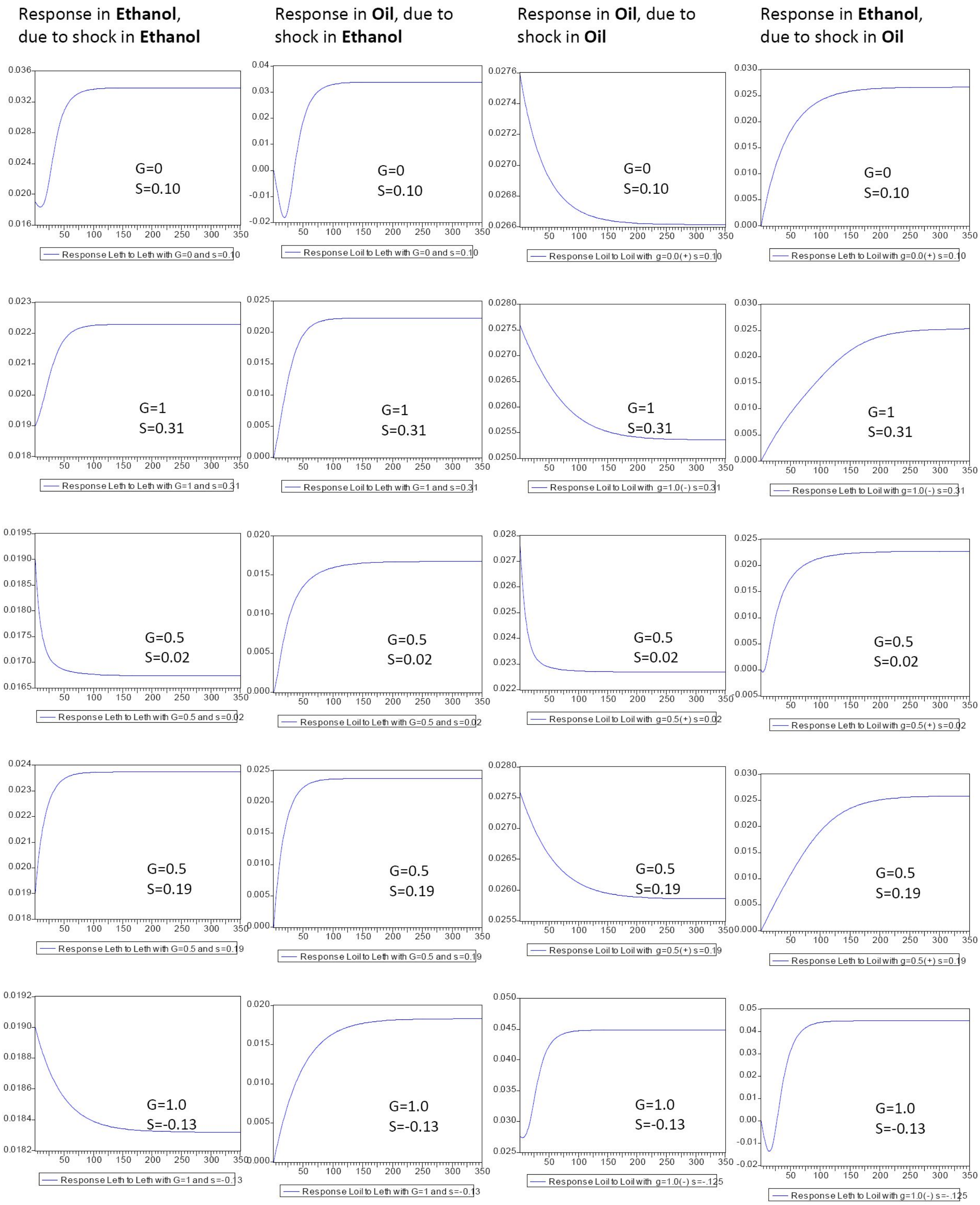

Figure 9: Generalized Response Functions Ethanol-Oil case with STVECM models