How integrated are the main markets of ethanol?

Felippe Cauê Serigati[1]
Paulo Furquim de Azevedo[2]
Mario Antonio Margarido[3]

Resumo
Há um mercado internacional de etanol capaz de coordenar globalmente as decisões de consumo e produção? Este artigo responde esta questão por meio de um sistema com séries de preços de etanol, no Brasil e nos Estados Unidos, e do preço internacional de petróleo, que permite avaliar o nível de integração entre os principais mercados domésticos de etanol. Os resultados indicam que não que há evidências de fortes e diretas conexões entre os dois principais mercados regionais deste biocombustível, notadamente quando incluída no modelo a série de preços de petróleo. Consequentemente, ainda não é possível considerar a existência de um mercado internacional de etanol já suficientemente desenvolvido. Diferentemente do que explícita ou implicitamente presume parte relevante da literatura, os resultados encontrando neste artigo sugerem que não é possível modelar o mercado internacional de etanol como um mercado bem desenvolvido tal qual é observado para outras *commodities* como soja, milho ou trigo.

Palavras-chave: etanol, mercado internacional, análise de cointegração, transmissão de preços, biocombustíveis

Abstract
Is there an international ethanol market able to coordinate production and consumption behavior? This article tackles this question based on the Brazilian and American ethanol prices, and the international oil price, that allows for assessing how integrated are the main markets of ethanol. Findings indicate that there are no strong and direct connections between to two most important domestic markets for ethanol when the oil prices are also included in the model. Consequently, it is possible to affirm that there is not yet a well-developed international market for this biofuel. Differently from what has been assumed by the literature on the economic impacts of biofuel production, these findings suggest that it is not appropriate to model the international market for ethanol as an efficient coordination mechanism, as it is the case for other commodities such as soybeans, corn, or wheat.

Key Words: ethanol, international market, cointegration analysis, price transmission, biofuels

Área de interesse: Área 10 - Economia Agrícola e do Meio Ambiente

JEL Code: Q11 - Agricultural and Natural Resource Economics; Environmental and ecological Economics - Aggregate Supply and Demand Analysis; Prices

[1] Doutorando, Escola de Economia de São Paulo da Fundação Getúlio Vargas (FGV-EESP)
[2] Professor, Escola de Economia de São Paulo da Fundação Getúlio Vargas (FGV-EESP)
[3] Pesquisador Científico, Instituto de Economia Agrícola (IEA)

1. INTRODUCTION

Is there an international ethanol market able to coordinate production and consumption behavior? Although this is a legitimate question with no trivial answer, the literature on the global ethanol trade often assumes explicitly or implicitly the existence of a highly integrated international market for this biofuel. For instance, Elobeid and Tokgoz (2008), analyzing the removal of U.S. trade barriers over the American and world ethanol prices, take as granted that this biofuel market has similar behavior to that observed in well developed international commodities markets. The authors also mention a 'world ethanol price': "(...) *when the tariff is not prohibitive, import demand is positive making the domestic U.S. price dictated by the world ethanol price through a price transmission equation.*" (Elobeid e Tokgoz, 2008, p. 19)[1]

Fabiosa et al (2009), evaluating, through multimarket models, the impact of the ethanol demand shocks over the allocation of land between corn, soybeans and sugar cane, also implicitly assume that i) there is a market, albeit incipient, and ii) an international price of ethanol: "*The U.S. ethanol market is nearly insulated from the world ethanol market because of a high U.S. tariff imposed on non-preferential imports of ethanol.*" (Fabiosa et al, 2009, p. 8); or "*The shock on the world ethanol market has a direct impact on the world ethanol price, as well as on the local ethanol markets in which the shock is initiated.*" (Fabiosa et al, 2009, p. 17)

Although those studies implicitly assume the existence of a well-developed multilateral ethanol market, anecdotal evidence suggests the contrary:

- Part of the Brazilian ethanol has been exported through long term contracts and arrangements that involve some degree of vertical integration; out of these contracts, exports exhibit a very irregular pattern;
- Ethanol futures contracts have limited liquidity, especially when compared to futures contracts of other commodities;
- Countries planning to adopt mandatory blending of ethanol in gasoline are faced with great uncertainty whether there is a sufficient supply of biofuel to meet their demand.

If that is the case, one should be suspicious about several predictions of those models, and an appropriate measure of price transmission should be somehow incorporated in their analysis. This article tackles this question based on the Brazilian and U.S. ethanol prices, and the international oil price, that allows for assessing how integrated are the main markets of ethanol.

The remaining of the paper is organized as follows. Next section details the empirical strategy to test the existence of a well-developed international ethanol market. Third, fourth and fifth sections present respectively the data, and unit root and cointegration tests. Section 6 presents the models and the hypothesis to be tested. Section 7, the core section, reports empirical results and robustness checks. Finally, Section 8 concludes with some implications and prospects for future research.

2. RESEARCH STRATEGY

This paper tests the implicit hypothesis adopted by the received literature that there is an international ethanol market able to coordinate efficiently production and

[1] It is fair to note that the authors make several caveats, for example, when they say that this is still a developing market, where supply and demand are stimulated by government interventions in the form of mandates or tax incentives, and when they use the price of the Brazilian anhydrous ethanol as a proxy for the international ethanol price. Actually, the need to use the Brazilian anhydrous ethanol price as a proxy is already an evidence that is difficult to affirm the existence of an international ethanol market.

consumption decisions. More specifically we test if unexpected shocks in one market are transmitted rapidly and strongly to the other regional markets.

In order to answer this question, this paper evaluates if the two most important ethanol regional markets, Brazil and United States, are connected by the market mechanism by means of which price signals transmit all relevant information regarding scarcity or excess of supply from one market to the other. In other words, it will be tested if the Brazilian and the American ethanol prices are cointegrated. Inasmuch as the ethanol is a substitute product of oil based fuels, it is possible that the connection between theses two regional markets is misled by the influence of the international crude oil market. Thus, the international crude oil price is also included in the system.

The paper's hypothesis was confronted with a sequential test procedure:

- Do the ethanol series have the same long-run stochastic trend, in a sense that short-run deviation is corrected by the system to a long-run equilibiurm? In other words, do the series are cointegrated? Therefore, the paper evaluated whether the series are cointegrated or not.
- For the specification in which was possible to accept the hypothesis of cointegration, we estimated the long-run price elasticities between the Brazilian and the American regional market. Are they consistent with a well developed international market? How fast is the short-term disequilibrium corrected to the long-run equilibiurm?

To assess whether an unexpected shock is transmitted strongly and quickly between the different regional ethanol market, we also estimated the impulse response functions and analyzed the variance decomposition.

3. DATA

To perform all those tests, we used four monthly price series from November of 2002 to August of 2011, comprising 106 observations of the:

- Brazilian anhydrous ethanol prices (US$/l): this series correspond to the ethanol price in São Paulo State in the trade between the sugar-cane producers and the distillers. The prices are collected daily and are transformed in a monthly series weighted by the volume traded in each transaction. The data is collected in Real (Brazilian currency) and is converted to Dolar by the weekly average of the exchange rate;
- American anhydrous ethanol price (US$/gallon): the ethanol rack prices per gallon FOB in Omaha are given for each month with an annual average for each year from 1982 to the present;
- International crude oil price: in this first approach, it was used as the monthly international crude oil price the Crude Oil (petroleum) Dated Brent (light blend 38 API) FOB price released by IMF (US$/per barrel).

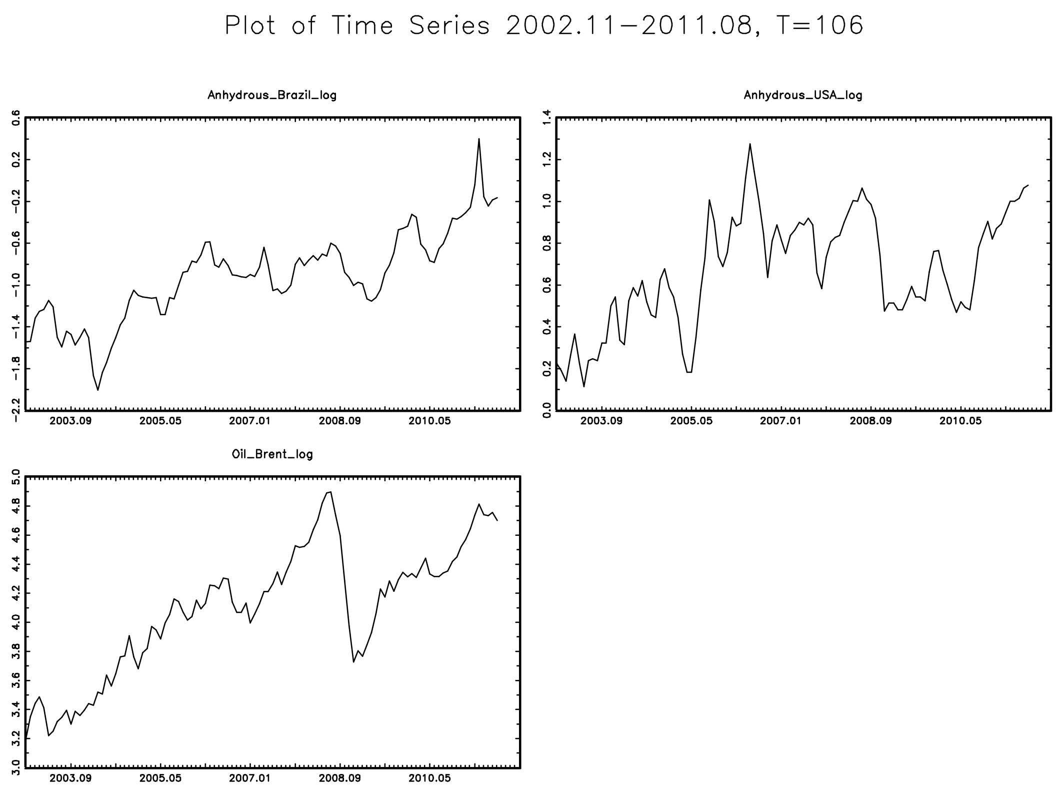
Table 1 summarizes the descriptive statistics of the series in level and in log. Figure 1 plots the graphs of the log of each series suggests the presence of outliers in the Brazilian ethanol prices and the presence of a structural break in the international crude oil price. These guesses are also evident when we plotted the series in first difference.

Table 1 – Descriptive statistics

Series	Mean	Median	Std. Dev.	Min	Max
Anhydrous_Brazil	0.44	0.41	0.20	0.13	1.50
Anhydrous_USA	2.02	1.95	0.53	1.12	3.58
Oil_Brent	65.39	63.28	27.07	24.50	133.90

Series (log)	Mean	Median	Std. Dev.	Min	Max
Anhydrous_Brazil	-0.92	-0.90	0.43	-2.01	0.40
Anhydrous_USA	0.67	0.67	0.27	0.11	1.28
Oil_Brent	4.09	4.15	0.44	3.20	4.90

Figure 1 – Graphs of the series



In this paper, we model a system that includes the Brazilian ethanol price, the American ethanol price, and the international price of crude oil price in order to evaluate how a price shock is transmitted between the Brazilian and the American regional ethanol markets. The first step is to evaluate if the series are non-stationary. The correlogram with the autocorrelation and partial autocorrelation of each series in level and in first difference suggests the presence of at least one unit root in all series, which indicates that there is evidence that all series are non-stationary. To confirm those guesses, we also performed unit root tests Table 2 and Table 3).

4. UNIT ROOT TESTS

According to the ADF test, it is not possible to reject the hypothesis that the American ethanol price and the international price of crude oil are non-stationary. However, with regard to the Brazilian ethanol price, this hypothesis is rejected if a deterministic trend is included in the model. Because both series seem to have a deterministic trend and due to the presence of outliers and a structural break that can be misleading these results, it is necessary to perform other tests.

Table 2 – ADF unit root test

Variable	Deterministc term	Lags	Test statistic	10% critical value	5% critical value	1% critical value
Brazilian anhydrous ethanol	constant and trend	1	-3.45	-3.13	-3.41	-3.96
	constant	2	-1.15	-2.57	-2.86	-3.43
		0	-1.41			
	none	2	-1.29	-1.62	-1.94	-2.56
		0	-1.53			
Brazilian anhydrous ethanol (1º dif)	constant	0	-8.69	-2.57	-2.86	-3.43
	none	1	-8.05	-1.62	-1.94	-2.56
		0	-8.65			
American anhydrous ethanol	constant and trend	2	-2.75	-3.13	-3.41	-3.96
	constant	2	-2.46	-2.57	-2.86	-3.43
	none	1	-0.03	-1.62	-1.94	-2.56
American anhydrous ethanol (1º dif)	constant	1	-8.19	-2.57	-2.86	-3.43
	none	1	-8.16	-1.62	-1.94	-2.56
Oil	constant and trend	2	-2.96	-3.13	-3.41	-3.96
	constant	6	-1.87	-2.57	-2.86	-3.43
	none	6	1.28	-1.62	-1.94	-2.56
Oil (1º dif)	constant	1	-5.41	-2.57	-2.86	-3.43
	none	5	-4.89	-1.62	-1.94	-2.56

Note: Critical values from Davidson & Macknnon (1993, Table 20.1)

Consistent with the ADF test , results obtained by the KPSS test also reject the hypothesis of stationarity to the American ethanol and to the international crude oil prices. Again, only without a deterministic trend, the KPSS test rejects the hypothesis of stationarity for the Brazilian ethanol price. Again, the presence of outliers and structural break can distort these results.

Because of the possible interference of outliers and structural brake in the last results, we applied the Lanne, Lütkepohl and Saikkonen (2002) unit root test with structural break. This test first estimates the deterministic terms using a generalized least-square (GLS) procedure and then subtract these terms from the original series. Finally an ADF-type test is performed on the adjusted series (Lütkepohl and Krätzig, 2004).

The graphs in Figure 1 and Figure 2 suggest the observed prices in April of 2011 and in February of 2004 in the Brazilian anhydrous series are possible outliers. August of 2008 seems to be the point of a structural break in the international crude oil series. Finally, there is no strong evidence of any outlier or structural break in the American ethanol series. To confirm these evidences suggested by the graphical analysis, the Lanne, Lüktepohl and Saikkonen (2003) test was applied. The latter test confirmed the graphical guesses.

Table 4 presents the results of the unit root test with structural break applied considering the dates above. Unfortunately, this test support just one outlier or structural break. Therefore, it was not possible to test the non-stationarity of the Brazilian anhydrous ethanol price considering both outliers. Nevertheless, interesting results were obtained: i) in controlling the structural break, it is possible to consider the international price of crude oil price stationary, contradicting the correlogram and the previous unit root tests; ii) it is possible to reject the hypothesis of stationarity of the Brazilian ethanol price only modeling the outlier in April of 2011; controlling only the outlier in, February of 2002 this stationarity hypothesis cannot be rejected. Unfortunately, it is not possible to test both outliers simultaneously.

Table 3 – KPSS unit root test

Variable	Test (stationarity)	Lags[1]	Test statistic	10% critical value	5% critical value	1% critical value
Brazilian anhydrous ethanol	level	4	1.61	0.35	0.46	0.74
		12	0.77			
	trend	4	0.11	0.12	0.15	0.22
		12	0.08			
Brazilian anhydrous ethanol (1º dif)	level	4	0.03	0.35	0.46	0.74
		12	0.07			
	trend	4	0.03	0.12	0.15	0.22
		12	0.07			
American anhydrous ethanol	level	4	0.83	0.35	0.46	0.74
		12	0.45			
	trend	4	0.27	0.12	0.15	0.22
		12	0.15			
American anhydrous ethanol (1º dif)	level	4	0.05	0.35	0.46	0.74
		12	0.08			
	trend	4	0.05	0.12	0.15	0.22
		12	0.08			
Oil	level	4	1.58	0.35	0.46	0.74
		12	0.74			
	trend	4	0.24	0.12	0.15	0.22
		12	0.14			
Oil (1º Dif)	level	4	0.06	0.35	0.46	0.74
		12	0.09			
	trend	4	0.04	0.12	0.15	0.22
		12	0.06			

Note: Critical values from Kwiatkowski et al. (1992)

[1] By Lütkepohl and Krätzig (2004), the lag truncation parameter can be chosed by $l_q = q(T/100)^{1/4}$, where T is the sample size.

Table 4 – Unit root test with structural break

Variable	Break	Shift Function	Lags	Test statistic	10% critical value	5% critical value	1% critical value
Brazilian anhydrous ethanol	2004 M02	Impulse dummy	3	-3.92	-2.76	-3.03	-3.55
			1	-3.64			
	2011 M04	Impulse dummy	3	-2.38	-2.76	-3.03	-3.55
			1	-3.48			
		Shift dummy	1	-2.83			
		Exponential shift	1	-2.63			
		Rational shift	3	-2.38			
			1	-2.33			
Oil	2008 M11	Shift dummy	2	-3.15	-2.76	-3.03	-3.55
		Exponential shift	2	-3.73			
		Rational shift	2	-3.84			

Note: Critical values from Lanne et al. (2002)

In short, it is not possible to reject the hypothesis of non-stationarity for the American and the Brazilian ethanol prices (with outliers). However the result for the international price of crude oil is not clear. As the latter test confront the conclusion of the first unit root test and the correlogram, all series will be treated as non-stationary. Consequently, since there are at least two clearly non-stationary variables that can produce a stationary system, we evaluated in which specification is possible to consider a cointegration relationship between those series.

5. COINTEGRATION TESTS

The cointegration tests were performed using the Johansen and the Saikkonen and Lüktepohl procedures, with and without specific treatment to outliers and structural break. We chosen these two tests because the former is the most conventional cointegration test in the literature, and the latter has asymptotically higher power than the Johansen test. (Saikkonen and Lütkepohl, 1999)

Before evaluating the three-dimensional systems, we performed the cointegration tests on all pairs of series. As argued by Lüktepohl and Krätzig (2004, p. 151-2), the reason is that "(…) *cointegration rank tests tend to have relatively low power – especially when applied to higher dimensional systems. Therefore, applying tests to bivariate systems first is a good check of the overall plausibility and consistency of the results obtained by looking at the data from different angles. Doing this can also help in finding a proper normalization of the cointegration matrix*" when the system with all the series is estimated.

Because the unit root tests were inconclusive to identify whether the international price of crude oil can be considered stationary or non-stationary, we also applied cointegration tests to pair of variables including the oil prices. The Johansen cointegration tests suggest that it is not possible to reject the hypothesis of one cointegration vector with 5% level of confidence for all pairs of variables (Brazilian anhydrous x American anhydrous, Brazilian anhydrous x crude oil, and American anhydrous x crude oil), as long as the system models also the outliers and the structural break.

The results produced by the Saikkonen and Lütjepohl tests are fairly similar to the previous one. For the pairs of variables Brazilian anhydrous x crude oil and American anhydrous x crude oil, it is not possible to reject the hypothesis of one cointegration vector, with or without the outliers. But for the system with the Brazilian and American anhydrous, the hypothesis is not rejected only if it is considered a constant inside the cointegration vector and, even so, the hypothesis is accepted just marginally (10%).

After all, because there are evidences that all series share some stochastic long-run trend, it is reasonable to consider that there is a cointegration relationship between all the series. Moreover, it is possible to choose any variable to normalize the coefficients in the cointegration vector when all the three series are considered. As a final remark, it is important to highlight again that, despite after controlling the structural brake, the series of the international price of crude oil rejected the hypothesis of non-stationarity. Thus, the latter variable was considered in the system because (i) the cointegration test with pair of variables supported the hypothesis of a long-run relationship, and (ii) as argued by Harris and Sollis (2003, p. 112), "*it is quite common for there to be tests for cointegration even when the preceding unit root analysis suggests that the properties of the variables in the equation(s) are unbalanced (i.e., they cannot cointegrate down to a common lower order of integration"*. The same authors also stress that stationary variables may play a key role in establishing a sensible long-run relationship between non-stationary variables (p. 112).

Table 5 presents the results obtained by the cointegration test between the full system (the three series). Because there is no evidence of a long-run linear growth between the variables, we did not perform any test with a trend inside the cointegration vector. Among the tested specifications (with and without outlires and structural break), just three didn't reject the hypothesis of at least one cointegration vector in this system:

- Model 1: according to the Saikkonen and Lüktepohl test, a system with linear trend in the level of the data, dummies to control outliers and strucutural break, and one lagged difference should be modeled considering two cointegration vectors;
- Model 2: also according to the Saikkonen and Lüktepohl test, a system with no linear trends, but with dummies controlling outliers and strucutural break, and one lagged difference should be modeled with just one cointegration vector;

- Model 3: according to the Johansen test, the previous specification should be modeled considering two cointegration vectors.

Table 5 – Cointegration tests for the complete system

Johansen trace cointegration test

Variables	Deterministc term	Lagged differences	H_0: r = r_0	Test statistic	10% critical value	5% critical value	1% critical value
anhydrous_br, anhydrous_usa, oil	c, orth tr	1	r_0 = 0	29.74	27.16	29.8	35.21
			r_0 = 1	13.24	13.42	15.41	19.62
	c	1	r_0 = 0	30.99	32.25	35.07	40.78
			r_0 = 1	14.44	17.98	20.16	24.69
			r_0 = 2	3.29	7.6	9.14	12.53
	c, outliers	1	r_0 = 0	55.85	45.95	48.84	54.59
			r_0 = 1	27.7	27.84	30.26	35.15
			r_0 = 2	11.12	13.7	15.75	20.09
hydrous_br, anhydrous_usa, oil	c, orth tr	1	r_0 = 0	30.28	27.16	29.8	35.21
			r_0 = 1	13.89	13.42	15.41	19.62
	c	2	r_0 = 0	28.24	32.25	35.07	40.78
			r_0 = 1	12.66	17.98	20.16	24.69
			r_0 = 2	1.96	7.6	9.14	12.53
		1	r_0 = 0	31.61	32.25	35.07	40.78
			r_0 = 1	15.14	17.98	20.16	24.69
			r_0 = 2	3.22	7.6	9.14	12.53
	c, outliers	1	r_0 = 0	55.25	45.95	48.84	54.59
			r_0 = 1	28.41	27.84	30.26	35.15
			r_0 = 2	11.41	13.7	15.75	20.09

Note: Critical values from Johansen (1995)

Saikonen & Lütkepohl cointegration test

Variables	Deterministc term	Lagged differences	H_0: r = r_0	Test statistic	10% critical value	5% critical value	1% critical value
anhydrous_br, anhydrous_usa, oil	c, orth tr	1	r_0 = 0	26.82	18.67	20.96	25.71
			r_0 = 1	10.32	8.18	9.84	13.48
	c, orth tr, outliers	1	r_0 = 0	26.22	18.67	20.96	25.71
			r_0 = 1	12.88	8.18	9.84	13.48
	c	1	r_0 = 0	26.55	21.76	24.16	29.11
			r_0 = 1	10.02	10.47	12.26	16.1
			r_0 = 2	0.12	2.98	4.13	6.93
	c, outliers	2	r_0 = 0	23.74	21.76	24.16	29.11
			r_0 = 1	6.98	10.47	12.26	16.1
			r_0 = 2	0.07	2.98	4.13	6.93
		1	r_0 = 0	25.54	21.76	24.16	29.11
			r_0 = 1	12.4	10.47	12.26	16.1
			r_0 = 2	0.07	2.98	4.13	6.93
hydrous_br, anhydrous_usa, oil	c, orth tr	1	r_0 = 0	26.91	18.67	20.96	25.71
			r_0 = 1	10.49	8.18	9.84	13.48
	c, orth tr, outliers	1	r_0 = 0	25.98	18.67	20.96	25.71
			r_0 = 1	13.13	8.18	9.84	13.48
	c	2	r_0 = 0	23.86	21.76	24.16	29.11
			r_0 = 1	7.83	10.47	12.26	16.1
			r_0 = 2	0.11	2.98	4.13	6.93
		1	r_0 = 0	25.36	21.76	24.16	29.11
			r_0 = 1	12.69	10.47	12.26	16.1
			r_0 = 2	0.09	2.98	4.13	6.93
	c, outliers	2	r_0 = 0	23.86	21.76	24.16	29.11
			r_0 = 1	7.83	10.47	12.26	16.1
			r_0 = 2	0.11	2.98	4.13	6.93
		1	r_0 = 0	25.36	21.76	24.16	29.11
			r_0 = 1	12.69	10.47	12.26	16.1
			r_0 = 2	0.09	2.98	4.13	6.93

Note: Critical values from Lüktepohl and Saikkonen (2000, Table 1)

That is, among all the tested specification, just the three previous models can support the hypothesis of the existence of well developed international market for ethanol. All the other specifications, because they reject the existence of cointegration vectors, they don't endorse the existence of an international market for ethanol.

Finally, it is important to stress that all those three proposed models have interesting characteristics to the hypothesis analyzed in this paper. The existence of two cointegration vectors can be related to the assumption that there is no strong link between the Brazilian and the American ethanol regional markets. That is, under the latter assumption, probably both markets have a strong relationship with the crude oil market, but not between each other, and, therefore, it is necessary one cointegration vector to link each ethanol series to the international oil market. In this sense, it will also be interesting to evaluate the behavior of this system when all the series are present in the same cointegration vector (case with only one cointegration vector). We expect that at least one of the coefficients associated to one of the regional market for ethanol is non-significant. After all, to assess if the Brazilian and American ethanol regional markets are really connected and how strong this connection is, all those three models might be tested.

6. MODEL AND HYPOTHESIS

The cointegration tests (Table 5) suggest models with the usual parameters of any other vector error correction model (VECM): an error correction term ($\beta' y_{t-1}$) representing the long-run relationship between the variables; a matrix (α) with the coefficients that represent how the short-run adjusts to the long-run equilibrium; a set of parameters ($\Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1}$) modeling the short-run dynamic; and the deterministic terms (CD_t) included in the model to capture linear trends, outliers, and structural breaks:

$$\Delta y_t = \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1} + CD_t + u_t \qquad (1)$$

where p is the number of lags in the levels of the data and $t = 1, .. T,$ with T representing the sample size. It is worth detailing that the α and β matrices to demonstrate which tests we performed to evaluate what the main characteristics of the link between the Brazilian and the American ethanol regional markets are. Below is presented the most unrestricted estimated model (two cointegration vectors, one lagged difference, linear trends, and dummies to capture outliers and a structural break - Model 1):

$$\begin{bmatrix} \Delta BR\_anh_t \\ \Delta US\_anh_t \\ \Delta Oil_t \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \end{bmatrix} \begin{bmatrix} \text{BR\_anh}_{\text{t-1}} \\ \text{US\_anh}_{\text{t-1}} \\ \text{Oil}_{t-1} \end{bmatrix} + \Gamma_1 \Delta y_{t-1} + CD_t + u_t \qquad (2)$$

The main objective of this paper is testing the hypothesis of strong connection between the Brazilian and American ethanol market. It is possible to reject this hypothesis:

- if ($\alpha_{11} \neq 0$ and $\beta_{12} \sim 0$) or ($\alpha_{12} \neq 0$ and $\beta_{22} \sim 0$): that is, if the Brazilian ethanol price moves toward a long-run path whose equilibrium has no influence (or only a weak influence) of the American ethanol price;
- if ($\beta_{12} \neq 0$ and $\alpha_{11} \sim 0$) or ($\beta_{22} \neq 0$ and $\alpha_{12} \sim 0$): that is, if the American ethanol price influences the long-run equilibrium of the system but the latter doesn't affect the Brazilian ethanol price;
- if ($\alpha_{21} \neq 0$ and $\beta_{11} \sim 0$) or ($\alpha_{22} \neq 0$ and $\beta_{21} \sim 0$): that is, if the American ethanol price moves toward a long-run path whose equilibrium that has no influence (or only a very weak influence) of the Brazilian ethanol price;
- if ($\beta_{11} \neq 0$ and $\alpha_{21} \sim 0$) or ($\beta_{21} \neq 0$ and $\alpha_{22} \sim 0$): that is, if the Brazilian ethanol price influences the long-run equilibrium but the latter doesn't affect the American ethanol price.

The following hypotheses are not the main goal of this paper, but they are also expected results:

- $\alpha_{31} = 0$ and $\alpha_{32} = 0$: that is, the price of crude oil is at least weakly exogenous in this system, but;
- $\beta_{13} \neq 0$ and $\beta_{23} \neq 0$: that is, the price of crude oil influences the long-run equilibrium and the short-run dynamics of the prices of both the Brazilian and the American ethanol markets.

The two hypotheses above are plausible because the international market of crude oil has been established for a long time and it is much wider than any ethanol market. Therefore, it is difficult to believe that the former is not at least weakly exogenous to the latter. However, because gasoline (a very important oil derivative) is a substitute (even though a imperfect substitute) to the Brazilian ethanol and because fertilizers (other important oil derivative) respond to a significant share of the corn (the main raw material of the American ethanol) total production costs, it is also reasonable to believe that the international price of crude oil exert a important influence over both ethanol markets and should be considered in the system.

6. MODEL AND HYPOTHESIS[2]

To estimate the suggested model, we used a two-step estimator (S2S). The first step is composed of two stages. The first stage estimates the matrix $\Pi = \alpha\beta'$ by OLS. The second stage eliminates the short-term parameters (Γ) by replacing them with their OLS estimator given Π obtained in the previous stage. Finally, the second step, given α, and Σ_u, estimates the final β by OLS[3].

Usually, the literature uses a restricted maximum likelihood (ML) estimator to estimate a VECM. Nevertheless, that two-step (S2S) estimator, besides has the same asymptotic distribution as the ML estimator (Reisen, 1993), can incorporate more easily restrictions on α and β matrices. Meanwhile, the ML estimator requires an iterative optimization whose results sometimes are not convergent. (Lütkepohl and Kräitzig, 2004)

Finally, before presenting the estimated models, it is important to choose which series will normalize each cointegration vector. As explained by Harris and Sollis (2003, p. 125-6), even if the VECM has only one cointegration vector ($\beta_{11}, \beta_{12}, ... \beta_{1n}$), it is not unique; any vector ($\lambda\beta_{11}, \lambda\beta_{12}, ... \lambda\beta_{1n}$), where $\lambda \neq 0$, is also a possible cointegration vector to this system. Hence, it is necessary to normalize this vector, fixing the coefficient of one variable at unity (e.g. $1, \beta_{12}/\beta_{11}, ... \beta_{1n}/\beta_{11}$).

Nevertheless, Lütkepohl and Kräitzig (2004) stress that the choice of which variable can be used to normalize the vector is not trivial; it cannot be a variable that is non-significant because it would mean dividing all the coefficients by something statistically equal to zero. Fortunately, all the series in that system are reasonable candidates because the cointegration tests between the pair of variables did not reject the hypothesis of one cointegration vector. Finally, with all those considerations, now it is possible to analyze the results.

7. RESULTS

The Box 1 shows the three estimated models (coefficients and their respective t-values). The first two models are both specified with two cointegration vectors, one lagged difference, and dummies to capture outliers in April and May of 2011 in the series of Brazilian ethanol price (*i11m04* and *i11m05*) and a structural break between September

[2] All the results were estimated using the open source software JMulTi®.
[3] For more details, Lütkepohl and Kräitzig (2004, p. 103-5).

and December of 2008 in the series of international price of crude oil (*d1*). In both specifications each cointegration vector represents a long-run equilibrium for each domestic ethanol market; the first cointegration vector represents the Brazilian regional ethanol market and the second cointegration vector depicts the American one. The difference between them is, while Model 1 contains linear trends for the variables in level, Model 2 is more restrictive and models just a constant in the cointegration vector.

Despite this difference, both models present quite similar results. The speed-of-adjustment coefficient of the American ethanol price associated to the error correction term that considers the Brazilian ethanol price (α_{12}) is not statistically significant, and, conversely, the speed-of-adjustment coefficient of the Brazilian ethanol price associated to the error correction model that considers the American ethanol price (α_{21}) is also not statistically different from zero. Those results suggest there is no strong connection between the Brazilian and the American ethanol markets.

It is also important to highlight that the price of crude oil is important to explain the long-run equilibrium of both domestic ethanol market, as long as β_{13} *and* β_{23} are statistically different from zero. More than that, in both models, the international price of crude oil can be considered exogenous because all the coefficients related to its speed of adjustment (α_{31} and α_{32}) and related to the first lagged difference cannot be considered statistically different from zero. That is, although the oil price is important to explain the long-run path of both regional ethanol markets, the latter do not cause any influence over the international oil market, confirming one of our expected results.

Although the third model contains just one cointegration vector, it supports the same conclusions suggested by the two previous models. Before analyzing their similarities, it is important to highlight that, because the coefficient related to the American ethanol price in the cointegration vector (US_anh_{t-1}) is not statistically significant and the values of the coefficient related to the international oil price (Oil_{t-1}) and of the constant are statistically the same as those presented in the first cointegration vector of the second model, it is possible to affirm that the only cointegration vector of the third model represents the long-run path of the Brazilian ethanol price. In other words, when the system is limited to just one cointegration vector, the long-run path of the American ethanol market is not anymore represented; the movements of the Brazilian ethanol price dominate the only available cointegration vector.

Even with just one cointegration vector, the conclusions are the same. Because US_ahn_{t-1} is not statistically significant, the American ethanol price does not influence the Brazilian ethanol price, and, as long as the speed of adjustment parameter linking the error correction term to the American ethanol price (α_{21}) also cannot be considered statistically different from zero, the Brazilian ethanol price also does not influence the American ethanol price. As the suggested by the previous models, the third model also rejects the hypothesis of a well developed international market for ethanol. Finally, the third model also supports the same conclusion about the exogeneity of the international oil price.

Table 6 shows the diagnostic test (residual autocorrelation, normality, and conditional heteroskedasticity) for all models. No one shows any problem of conditional heteroskedasticity or of joint non-normality. The first model also does not have any problem of residual non-normality in any individual equation, but Models 2 and 3 reject just marginally the hypothesis of non-normality for the third equation – that represents the international oil price. Similarly, the first model does not present any serious problem of residual autocorrelation, while the second model seems having some residual autocorrelation in latter lags. Because those problems are present just in latter lags and the results are the same for all models, it is not interesting to modifying the specification to try to overcome this undesired characteristic.

Box 1 – Estimared models

Model 1*

$$\begin{bmatrix}\Delta BR\_anh_t\\ \Delta US\_anh_t\\ \Delta Oil_t\end{bmatrix}=\begin{bmatrix}\underset{(3.707)}{-0.150} & \underset{(-0.678)}{-0.038}\\ \underset{(0.042)}{0.002} & \underset{(-4.220)}{-0.226}\\ \underset{(0.133)}{0.005} & \underset{(-0.611)}{-0.029}\end{bmatrix}\begin{bmatrix}\text{BR\_anh}_{t-1} & & \underset{(-6.705)}{-0.966}\,\text{Oil}_{t-1}\\ & \text{US\_anh}_{t-1} & \underset{(-4.685)}{-0.430}\,\text{Oil}_{t-1}\end{bmatrix}$$

$$+\begin{bmatrix}\underset{(3.412)}{0.257} & \underset{(2.075)}{0.191} & \underset{(-1.453)}{-0.176}\\ \underset{(-1.312)}{-0.095} & \underset{(3.451)}{0.306} & \underset{(2.596)}{0.302}\\ \underset{(-0.436)}{-0.028} & \underset{(-1.256)}{-0.099} & \underset{(0.662)}{0.068}\end{bmatrix}\begin{bmatrix}\Delta BR\_anh_{t-1}\\ \Delta US\_anh_{t-1}\\ \Delta Oil_{t-1}\end{bmatrix}$$

$$+\begin{bmatrix}\underset{(4.491)}{0.413} & \underset{(-6.028)}{-0.591} & \underset{(-3.519)}{-0.170} & \underset{(-3.619)}{-0.756}\\ \underset{(0.253)}{0.022} & \underset{(0.031)}{0.003} & \underset{(-0.427)}{-0.020} & \underset{(-1.162)}{-0.234}\\ \underset{(0.681)}{0.053} & \underset{(-1.035)}{-0.087} & \underset{(-5.461)}{-0.225} & \underset{(0.081)}{0.014}\end{bmatrix}\begin{bmatrix}d1_t\\ i11m04_t\\ i11m05_t\\ c\end{bmatrix}+\begin{bmatrix}\hat{u}_{1,t}\\ \hat{u}_{2,t}\\ \hat{u}_{3,t}\end{bmatrix}$$

Model 2*

$$\begin{bmatrix}\Delta BR\_anh_t\\ \Delta US\_anh_t\\ \Delta Oil_t\end{bmatrix}=\begin{bmatrix}\underset{(-4.062)}{-0.151} & \underset{(-0.681)}{-0.038}\\ \underset{(-0.058)}{0.002} & \underset{(-4.246)}{-0.227}\\ \underset{(-1.018)}{-0.033} & \underset{(-0.775)}{-0.038}\end{bmatrix}\begin{bmatrix}\text{BR\_anh}_{t-1} & & \underset{(-6.705)}{-0.966}\,\text{Oil}_{t-1} & \underset{(8.076)}{4.759}\,c\\ & \text{US\_anh}_{t-1} & \underset{(-4.685)}{-0.430}\,\text{Oil}_{t-1} & \underset{(2.876)}{1.078}\,c\end{bmatrix}$$

$$+\begin{bmatrix}\underset{(3.431)}{0.257} & \underset{(2.077)}{0.191} & \underset{(-1.453)}{-0.176}\\ \underset{(-1.294)}{-0.093} & \underset{(3.460)}{0.306} & \underset{(2.612)}{0.303}\\ \underset{(-0.165)}{-0.011} & \underset{(-1.131)}{-0.092} & \underset{(0.782)}{0.083}\end{bmatrix}\begin{bmatrix}\Delta BR\_anh_{t-1}\\ \Delta US\_anh_{t-1}\\ \Delta Oil_{t-1}\end{bmatrix}$$

$$+\begin{bmatrix}\underset{(-3.576)}{-0.170} & \underset{(4.522)}{0.413} & \underset{(-6.095)}{-0.591}\\ \underset{(-0.390)}{-0.018} & \underset{(0.281)}{0.025} & \underset{(0.068)}{0.006}\\ \underset{(-4.876)}{-0.204} & \underset{(0.955)}{0.077} & \underset{(-0.609)}{-0.052}\end{bmatrix}\begin{bmatrix}d1_t\\ i11m04_t\\ i11m05_t\end{bmatrix}+\begin{bmatrix}\hat{u}_{1,t}\\ \hat{u}_{2,t}\\ \hat{u}_{3,t}\end{bmatrix}$$

Model 3*

$$\begin{bmatrix}\Delta BR\_anh_t\\ \Delta US\_anh_t\\ \Delta Oil_t\end{bmatrix}=\begin{bmatrix}\underset{(-4.101)}{-0.152}\\ \underset{(-0.412)}{-0.016}\\ \underset{(-1.081)}{-0.035}\end{bmatrix}\begin{bmatrix}\text{BR\_anh}_{t-1} & \underset{(0.429)}{0.159}\,\text{US\_anh}_{t-1} & \underset{(-4.644)}{-1.034}\,\text{Oil}_{t-1} & \underset{(6.675)}{4.931}\,c\end{bmatrix}$$

$$+\begin{bmatrix}\underset{(3.427)}{0.257} & \underset{(2.096)}{0.185} & \underset{(-1.433)}{-0.169}\\ \underset{(-1.241)}{-0.097} & \underset{(2.147)}{0.197} & \underset{(3.296)}{0.405}\\ \underset{(-0.173)}{-0.011} & \underset{(-1.383)}{-0.108} & \underset{(0.939)}{0.098}\end{bmatrix}\begin{bmatrix}\Delta BR\_anh_{t-1}\\ \Delta US\_anh_{t-1}\\ \Delta Oil_{t-1}\end{bmatrix}$$

$$+\begin{bmatrix}\underset{(-3.600)}{-0.171} & \underset{(4.521)}{0.413} & \underset{(-6.091)}{-0.590}\\ \underset{(-0.623)}{-0.031} & \underset{(0.273)}{0.026} & \underset{(0.106)}{0.011}\\ \underset{(-4.921)}{-0.206} & \underset{(0.955)}{0.077} & \underset{(-0.600)}{-0.051}\end{bmatrix}\begin{bmatrix}d1_t\\ i11m04_t\\ i11m05_t\end{bmatrix}+\begin{bmatrix}\hat{u}_{1,t}\\ \hat{u}_{2,t}\\ \hat{u}_{3,t}\end{bmatrix}$$

** (t-values between parenthesis)*

In order to evaluate if the estimated results are consistent, we performed stability tests for all models. These tests consist in recursive estimates of each eigenvalue of the β matrix, each speed of adjustment parameter, and each lagged differenced coefficient. This stability analyses are important to evaluate if the model is time invariant in each possible subsample. As suggested by Lüktepohl and Krätzig (2004, p. 131-2), "*(...) the model is* *estimated on the basis of the first τ observations for τ = T_1,…, T, where T_1 is such that the degrees of freedom necessary for estimation are available. (…) Plotting the recursive estimates together with their confidence intervals for τ = T_1,…, T can give useful information on possible structural breaks.*".

Table 6 - Diagnostic tests

Test	Model 1		Model 2		Model 3	
	Test statistics	p-value	Test statistics	p-value	Test statistics	p-value
Autocorrelation						
Q_6	51.18	0.09	54.08	0.05	63.33	0.05
Q_{12}	109.18	0.12	110.27	0.11	121.60	0.07
Q_{18}	156.12	0.29	156.64	0.28	167.74	0.21
Q_{24}	216.67	0.21	215.34	0.23	223.71	0.22
LM_2	21.08	0.28	24.06	0.15	26.60	0.09
LM_4	38.85	0.34	49.05	0.07	47.10	0.10
LM_6	57.71	0.34	77.14	0.02	67.45	0.10
LM_8	74.33	0.40	97.07	0.03	83.23	0.17
Nonnormality						
Each series						
u_1	3.20	0.20	3.20	0.20	3.07	0.22
u_2	1.95	0.38	2.04	0.36	1.23	0.54
u_3	4.20	0.12	5.57	0.06	5.38	0.07
Joint						
LJB	9.86	0.13	11.15	0.08	8.94	0.18
LJB^L	7.62	0.27	8.82	0.18	7.57	0.27
Conditional heteroskedasticity						
$MARCH_{LM}(6)$	223.24	0.35	222.57	0.37	188.33	0.91

note: Q_n: Portmanteau test
LM_n: Breusch-Godfrey test
LJB: Lomnicki-Jarque-Bera test
LJB^L: LJB test with standardized residuals (Lütkepohl, 1991)
n: number of lags

The stability tests (Annex 1)[4] suggest that (i) the eigenvalues representing each cointegration vector, cannot be considered statistically equal to zero in any subsample, (ii) the speed of adjustment parameters α_{11} (for all models) and α_{22} (Models 1 and 2) cannot be considered statistically equal to zero in any subsample. It means that the conclusions about in the absent influence of each regional ethanol price over the other regional ethanol market did not change in any subsample. Adding the recursive estimates of each first lagged coefficient in this analysis, it is also possible to conclude that the system does not present any significant problem of structural break.

Finally, to evaluate how the dynamic of each variable in the system is, we estimated the impulse response function (Annex 2) and the forecast error variance decomposition (Annex 3) for each series for all models[5]. In all situations an unexpected shock in the Brazilian ethanol price does not cause a significant effect on the American ethanol price, nor a unexpected shock in the American ethanol price cause a significant effect over the Brazilian ethanol price. But an unexpected shock in the international crude oil price has a significant and permanent effect over the forecast error of both ethanol prices.

[4] Due to the limited space, it is showed just the stability tests for the first model.
[5] We showed here just the impulse response functions and the forecast error variance decomposition of the Model 1 because the results are the same for all models and the space here is limited.

8. CONCLUDING REMARKS

The main objective of this paper is to test the hypothesis that there is a well-developed international market for ethanol. One of the key characteristics of a well developed international market is the ability to transmit quickly and easily information between different domestic markets in order to eliminate arbitrage opportunities and to achieve a new equilibrium. To test this hypothesis this paper assessed how a price shock is transmitted between the Brazilian domestic market and the American domestic ethanol market, the two most important regional market for this biofuel. This exercise was motivated by a literature on bioenergy market simulations (Elobeid and Tokgoz, 2008; Fabiosa et al, 2009) that, with the aim to estimate the economic impact of s in the ethanol market (e.g. on price, land use, production etc.), assumes direct price transmission between the Brazilian and the American ethanol market. Apart from the U.S. import tariff, those models take price transmission for granted, and implicitly assume that the international ethanol market is able to coordinate production and price decisions as observed in other commodities markets, such as soybeans and wheat.

All tests reject the existence of a well-developed international market for ethanol. First, some cointegration tests with different specifications were performed. If it is possible to rejects the presence of at least one cointegration vector in one specification, then this specification rejects the paper's hypothesis. The reason is, if there is no cointegration vector between the Brazilian and the American ethanol prices, controlling by the movements of the oil price, it is impossible to affirm there is a strong connection between them. In short, the Brazilian and the American ethanol prices don't share the same long-run path. Among all the tested specifications (Table 5), only three did not reject the hypothesis of at least one cointegration vector. Thus, all the other specifications reject the integration hypothesis.

Second, we analyzed each one the three specifications that found at least one cointegration vector. The first model requires a system with two cointegration vectors and linear trends in the levels of the data, the second suggests a system with two cointegration vectors, but with no linear trends, and the last model asks for system also with no linear trends, but with only one cointegration vector.

After estimating the three models, both models with two cointegration vectors indicated that:

- Each cointegration vector represents the long-run relationship between each regional ethanol price and the international price of crude oil;
- Because the speed of adjustment parameters in the U.S. ethanol price equation associated with the Brazilian cointegration vector were not statistically different from zero, the cointegration vector that represents the Brazilian ethanol long-run relationship did not have any significant influence on the U.S. ethanol price;
- The opposite is also true. The cointegration vector that represents the U.S. ethanol long-run relationship also did not cause any significant influence on the Brazilian ethanol price because the Brazilian speed of adjustment parameters related to the U.S. cointegration vector were not statistically different from zero.

We also run a specification with only one cointegration vector because it represents the opportunity to see the behavior of the Brazilian and the U.S. ethanol prices when they can share the same long-run relationship. Again, results suggest that there is no strong link between those two ethanol markets:

- The coefficient associated to the U.S. ethanol price in the cointegration vector is not statistically different from zero. That is, the U.S. price does not affect the Brazilian market;

- Because the speed of adjustment parameter associated to the U.S. price equation is not also statistically different from zero, it also possible to affirm that the long-run path of the Brazilian ethanol price does not affect the American ethanol price.

The impulse response functions and the forecasts errors variance decomposition estimated by all the three models also suggest that there is no strong connection between the Brazilian and the American ethanol markets.

The estimated coefficient do not differ among the different models analyzed (Annex 4). In the same way, recursive estimates indicated that those results were valid for almost any subsample. After all, the three specifications supported the same conclusion: it is not possible to affirm there is a strong link between the Brazilian and the U.S. ethanol markets. Considering all the results obtained, we did not find any evidence to support the hypothesis of a well-developed international market for ethanol.

The estimated models indicate also other interesting results. First, in all the three models the international price of crude oil was weakly exogenous to the system. This result is plausible, inasmuch as it is expected the crude oil price to affect ethanol prices but not the reverse. The long-run price elasticity between crude oil and the Brazilian ethanol can be considered statistically unitary, whereas the long-run price elasticity between crude oil and the U.S. ethanol is almost half than the Brazilian one (Annex 4). This result is probably due to the significant flex-fuel fleet in Brazil, whereas the ethanol market in United States is driven by mandates. However, as so high values for price elasticities were unexpected, further detailed studies are welcoming for checking robustness.

Finally, it is important to highlight that, during all the period analyzed by this research, the U.S. applied import tariffs on ethanol, which could also drive the results. It is plausible that the U.S. price is not high enough, or the Brazilian price is not lower enough, to compensate this trade barrier, what would be a possible reason for the absence of cointegration for the whole time series. Only when arbitrage gains are large enough one should expect the two regional ethanol prices to share the same stochastic trend. Therefore, future research shall consider a threshold cointegration model so as to incorporate the effect of trade barriers. Moreover, as the U.S. dropped the import tariff on January 2012, future research will be able to rule out the trade barriers hypothesis.

REFERENCES

Elobeid, A., and Tokgoz, S. (2008). Removing distortions in the U.S. ethanol market: what does imply for the United States and Brazil? American Agricultural Economics Association. 2007 Annual Meeting, July 29-August 1 Portland, Oregon Nº 9808.

Fabiosa, Jacinto F. et alii (2009). Land allocation effects of the global ethanol surge: prediction from the international FAPRI Model. Center for Agricultural and Rural Development. Iowa State University. Working Paper 09-WP 488.

Harris, R., and Sollis, R. (2003). Applied Time Series Modelling and Forecasting. Wiley, Chichester.

Lanne, M., Lütkepohl, H. and Saikkonen, P. (2002). Comparison of unit root tests for time series with level shifts, Journal of Time Series Analysis 23: 667–685.

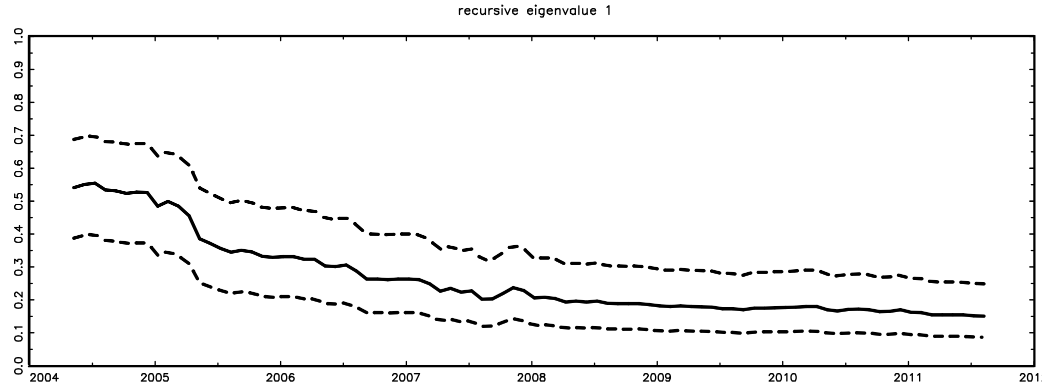
Lanne, M., Lütkepohl, H. and Saikkonen, P. (2003). Test procedures for unit roots in time series with level shifts at unknown time, Oxford Bulletin of Economics and Statistics 65: 91–115.

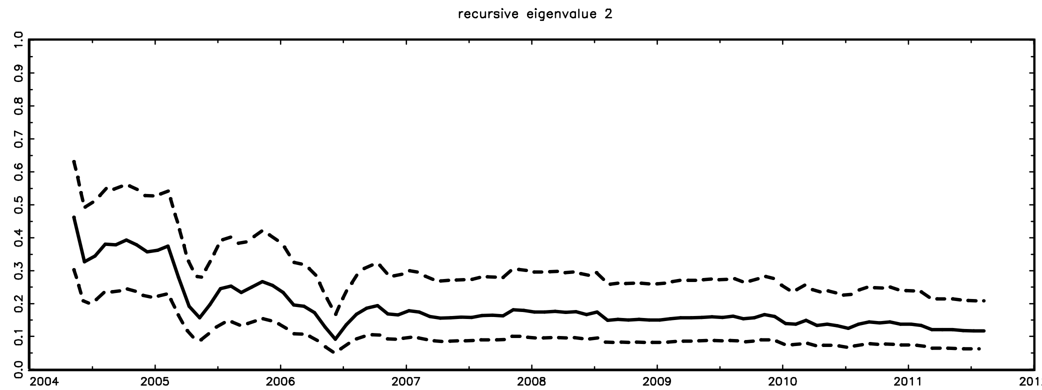
Lütkepohl, H., Krätzig, M. (2004) Applied Time Series Econometrics. Cambridge University Press.

Reinsel, G. C. (1993). Elements of Multivariate Time Series Analysis, Springer-Verlag, New York.

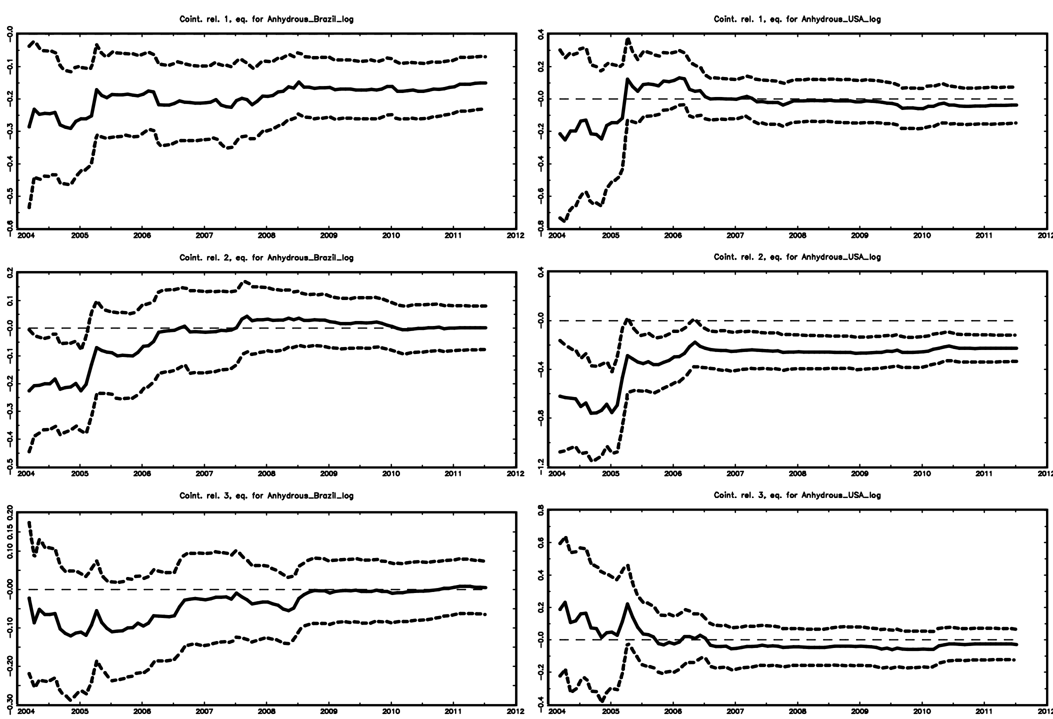
Saikkonen, P. and Lütkepohl, H. (1999). Local power of likelihood ratio tests for the cointegrating rank of a VAR process, Econometric Theory 15: 50–78.

ANNEX 1: RECURSIVE ESTIMATES OF THE EIGENVALUES AND THE SPEED OF ADJUSTMENT PARAMETERS

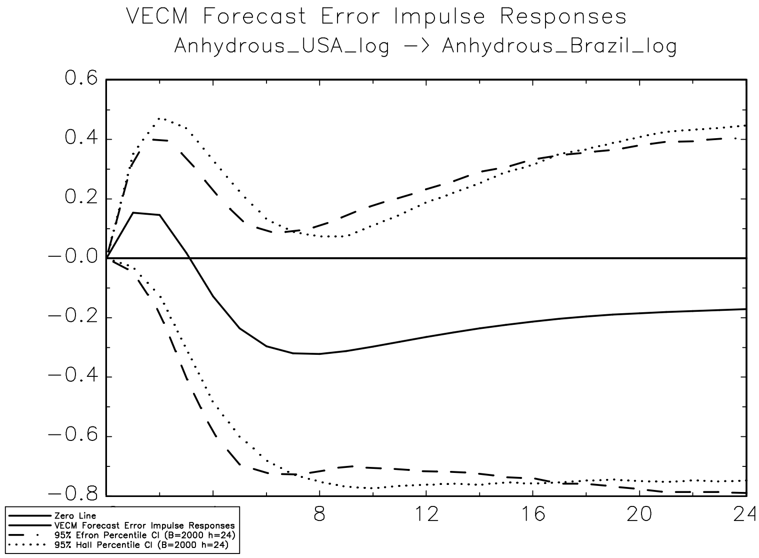


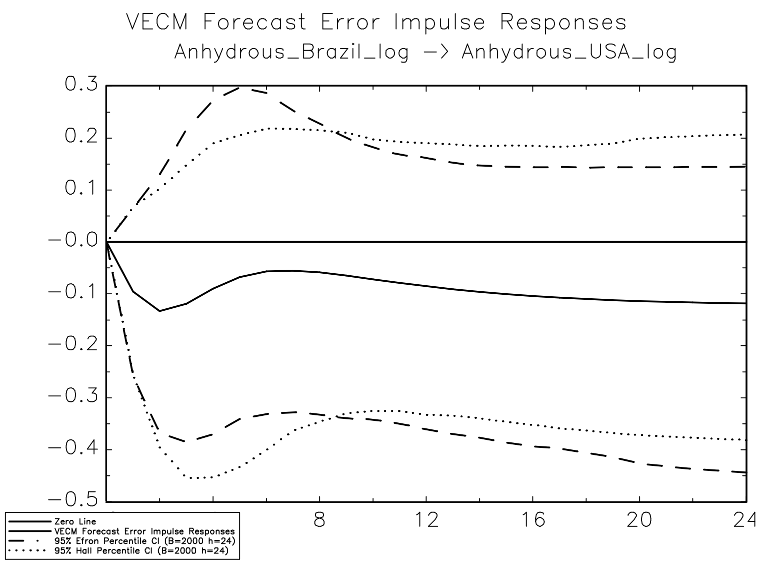


Alpha vector

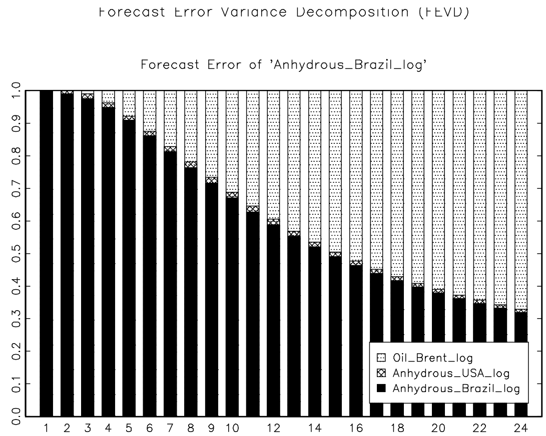


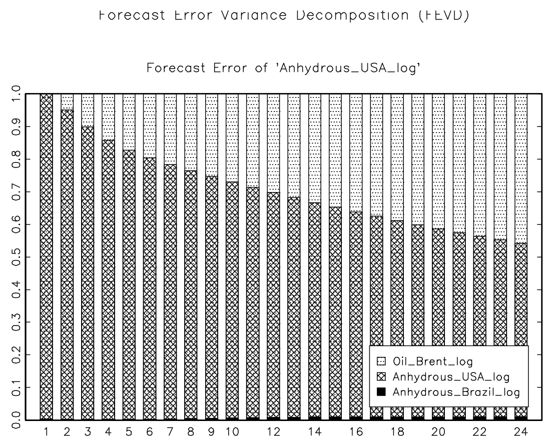
ANNEX 2: IMPULSE RESPONSE FUNCTIONS

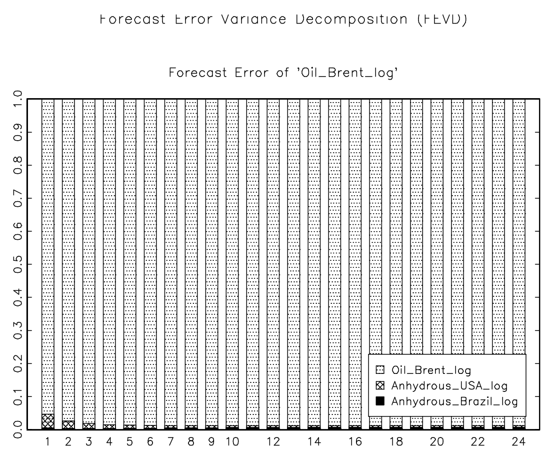




ANNEX 3: FORECAST ERROR VARIANCE DECOMPOSITION







ANNEX 4: COMPARISON OF ESTIMATED REGRESSORS

Cointegration vectors

Price elasticity Brazilian ethanol - crude oil	
Specification 1	-0.966
	-6.705
Specification 2	-0.966
	-6.705
Specification 3	-1.034
	-4.644

(t-values)

Price elasticity American ethanol - crude oil	
Specification 1	-0.430
	-4.685
Specification 2	-0.430
	-4.685
Specification 3	0.159
	0.429

Alpha matrix

Speed of adjustment parameter: Brazilian ethanol		
	α_{11}	α_{12}
Specification 1	-0.15	-0.038
	-3.707	-0.678
Specification 2	-0.151	-0.038
	-4.062	-0.681
Specification 3	-0.152	-
	-4.101	-

Speed of adjustment parameter: American ethanol		
	α_{21}	α_{22}
Specification 1	0.002	-0.226
	0.042	-4.220
Specification 2	-0.002	-0.227
	-0.058	-4.246
Specification 3	-0.016	-
	-0.412	-

Speed of adjustment parameter: American ethanol		
	α_{31}	α_{32}
Specification 1	0.005	-0.029
	0.133	-0.611
Specification 2	-0.033	-0.038
	-1.018	-0.775
Specification 3	-0.035	-
	-1.081	-

(t-values)