Bayesian Estimation of Non-Linear Vector Error Correction Models: The Case of the Sugar-Ethanol-Oil Nexus in Brazil

Kelvin G. Balcombe and George Rapsomanikis*

Non-linear adjustment towards long-run equilibrium in the sugar-ethanol-oil nexus in Brazil is examined. We develop generalised bivariate error correction models that allow for cointegration between variables, where dynamic adjustments are potentially non-linear functions of the disequilibrium errors. Linear error correction models and threshold models are specific cases within this generalised structure. The models are estimated using Bayesian Monte Carlo Markov Chain algorithms and applied to oil, sugar and ethanol prices. Bayesian model selection methods suggest that the long-run drivers of Brazilian sugar prices are oil prices and that there are nonlinearities in the adjustment processes of sugar and ethanol to oil, but linear adjustment between ethanol and sugar. Threshold adjustment is supported between ethanol and oil and asymmetric adjustment between sugar and oil. Bayesian model averaging is used to estimate the elasticities of transmission from oil to sugar and ethanol.

* Kelvin G. Balcombe is a lecturer in the Department of Agricultural and Food Economics, University of Reading, UK and George Rapsomanikis is an economist at the Commodities and Trade Division of the Food and Agriculture Organization (FAO) of the United Nations
that developed an analogue to Granger’s representation theorem for nonlinear vector autoregressions (see Corradi et al., 2000; Escribano and Mira, 2002; Bec and Rahbec, 2004; Saikkonen, 2005).

In this article, we explore nonlinearities employing a Bayesian Monte Carlo Markov Chain (MCMC) approach to estimation and inference. Although Bayesian methods have been applied in the context of cointegrated threshold models in the past (see Balcombe, 2006; Balcombe et al., in press), applications are still few and there are a number of extensions that have yet to be investigated. We extend the literature in three ways. Firstly, we estimate the proportion of observations that lie within a symmetric threshold interval, rather than following earlier work in which the threshold location parameter itself is estimated and, therefore, identified indirectly. An additional advantage of this approach is that it facilitates Bayesian model selection. Second, we develop a number of models that embody functions of nonlinear adjustment towards long run equilibrium, including threshold effects, and construct point estimates and standard deviations of the parameters of interest. Third, the alternative non-linear functions of disequilibria are compared with the linear error correction model and the threshold error correction model by means of the marginal likelihood function, thus enabling model choice. Along with the tests for nonlinearity, we use the marginal likelihood to choose lag length, to test for weak and strong exogeneity and to select rank restrictions that characterise the behaviour of the variables as stationary, stochastically trending and/or cointegrated. The work in this article is limited to bivariate relationships, in line with much existing work that examines thresholds and nonlinearities. Naturally, the methods outlined herein might be usefully extended into a multivariate framework in the future.

We employ our methods to examine weekly price series for crude oil, ethanol and sugar in Brazil, the world’s leading producer and exporter of sugar and ethanol. The motivation for working with this data set is provided by the recent surges in the international crude oil prices that reached historically high levels during the period 2005-2006. Subsequent increases in the prices for ethanol and strong derived demand for sugarcane fuelled sugar prices that nearly doubled to a 23-year high level in the first half of 2006 (United Nations, Food and Agriculture Organization, 2006). Brazil’s dominant position in the world sugar economy implies that domestic sugar production affects the global sugar market balance and prices. Undeniably, the pressure exerted by increasing oil prices, in conjunction with growing environmental concerns, have encouraged a momentum for a major international switch from fossil fuels to renewable bioenergy. For many Kyoto Protocol signatory countries, biofuels consist of the main option for substituting fossil fuels and mitigating greenhouse emissions.

More than half of annual sugarcane production in Brazil is processed into ethanol. The growth of the Brazilian ethanol market has been realised due to a combination of factors, including government policies and technical change both in the processing of sugarcane into ethanol and in the manufacturing of vehicles that can use high level blends of ethanol with petrol. While the Brazilian ethanol market was effectively

---

1 The Brazilian alcohol programme began in 1975 and was characterised by production quotas and institutional setting of the price for ethanol at a level lower than that of petrol, combined with subsidies to ethanol distillers. The ethanol programme was gradually eliminated between 1996 and 2000.
liberalized between 1996 and 2000, the policies assisted the industry to achieve economies of scale which, in conjunction with technology improvements, have resulted in ethanol production costs being lower than the regional supply costs of petrol, in volume terms. Nevertheless, in energy terms, when prices per litre are converted into prices per energy unit, ethanol is, for most of the time, higher than those for petrol. However, substantial increases in the price of oil have resulted in increasing ethanol’s competitiveness not only during the period 2005-2006, but also for a number of months in the years 2002 and 2003 (International Energy Agency, 2004; Goldemberger et al., 2004; Hamelick and Faaj, in press). In a similar line, analysis undertaken by the Organization for Economic Cooperation and Development (2006) indicates that the current cost structure of the Brazilian ethanol industry results in ethanol being competitive with petrol over a price level of US$39 per barrel. Such analyses suggest that the prices may co-move above a certain level, as petrol can be substituted for ethanol, either partially in petrol-ethanol blends, or perfectly due to the introduction of flexible fuel vehicles that can use any combination of petrol-ethanol blends, but also pure ethanol only.2

In a manner similar to the existence of transaction and adjustment costs in spatial markets, price differentials between oil and ethanol, in conjunction with the extent to which the two fuel types can be substituted with one another, may result in nonlinearities. In essence, oil and ethanol markets, and by extent the sugar market, may be perceived as been separated when prices move within a certain threshold. Once prices move outside this threshold, substitution possibilities in consumption will be sufficient to induce the transmission of price signals from one market to another, thus giving rise to nonlinear behaviour and threshold effects. The Organisation for Economic Cooperation and Development (OECD) and the Food and Agriculture Organisation (FAO) of the United nations in their agricultural outlook (OECD and FAO, 2006) draw attention to the potential relationship between oil and sugar prices, suggesting that if oil prices are sustained at high levels in the future, the share of sugarcane use for ethanol is likely to continue rising steadily. Such developments are expected to adversely affect sugar output, triggering increases in sugar prices that may result in an expansion of the area under sugarcane in Brazil, but also in increases in the production of sugarcane and ethanol from other low-cost producers.

The article proceeds by providing a review of nonlinear VECMs. In the following sections we introduce the linear and non-linear model specifications and discuss the estimation of these models along with priors used. We find that oil prices are the main driver of both sugar and ethanol prices in Brazil whilst our results support nonlinear transformations of the long-run disequilibria.

---

2 The sale of flexible fuel cars has increased dramatically since their introduction in 2003, assisted by some tax incentives. The share of flex-fuel cars in total car sales in Brazil reached 22 percent in 2004, 40 percent in 2005, and is expected to rise to 60 percent in 2006, allowing petrol and ethanol to be purchased on the basis of relative prices (Goldemberger et al., 2004; Tokgoz and Elobeid, 2006).
Nonlinear Adjustment in Vector Error Correction Models

A popular class of nonlinear models is characterised by transition functions, either discrete or continuous, that determine regime switches. The literature can largely be divided into two main methodologies, the Markov-switching VECMs (MSVECMs) and the threshold VECMs (TVECMs). Both these strains evolved from their corresponding univariate counterparts. MSVECMs models were originally applied by Hamilton (1989) in a business cycle context and were extended to a multivariate application in a form of error correction models by Krolzig (1997). MSVECMs are well suited for modelling the changing relationship between economic variables in the context of business cycles, mainly because regimes and their corresponding probabilities and duration, can be readily associated with periods of recession, or high growth (see Krozig et al., 2002 and Clements and Krolzig, 2002). Threshold cointegration was introduced by Balke and Forby (1997) and TVECMs have been applied by Hansen and Seo (2002) and Seo (2005). Arguably, TVECMs have their origin in the Self-Exciting Autoregressive model (SETAR) by Tong (1983) and Hansen (1999). MSVECMs are based on the general idea that model parameters are a stochastic function of an unobservable regime indicator variable, with the regimes being reflected by an \( m \)-state ergodic Markov chain. Regime switches are based on a probabilistic function for the incidence of different regimes, as compared to TVECMs where regimes are identified, with certainty, by means of a threshold parameter.

In practice, explicit estimation of threshold parameters is, in some cases, preferable to identifying the probability that a certain regime occurs at a particular point in time. Depending on the application, thresholds can be rationalised in a number of ways. For example, in testing the Law of One Price, deviations between prices of assets traded in different markets may occur due to many reasons including transaction costs, adjustment costs and capital constraints. Such factors may give rise to thresholds with arbitrage opportunities being realised when the deviation in prices is high. The estimation of threshold parameters enable the researcher to calculate the magnitude of the price spread above which prices adjust to their long run equilibrium. Other popular models, such as the Momentum Threshold Autoregressive models of Granger and Lee (1989), Enders and Granger (1998) and Escribano and Pfann (1998) allow for the variables to adjust differently depending on whether the disequilibria are negative, or positive. More general specifications can contain multiple thresholds and alternative points of attraction.

Threshold behaviour and asymmetric adjustment are not synonymous. Asymmetric adjustment can, in general, be characterised as a threshold approach. However, there are asymmetric forms of adjustment, which are modelled in this article, that do not strictly rely on the concept of a threshold at all. In addition, switches from one regime to another may not be sudden, but gradual, as economic agents may need time to adjust to changes in the economic environment or policies. A smooth TVECM (STVECM) (e.g. as in Kapetanios et al., 2006) is obtained by replacing the indicator discrete function by a smooth function, often adhering to exponential and logistic specifications, that give rise to smooth adjustments based on an inverted normal density function and a cumulative logistic distribution.
respectively. An additional benefit of such specifications is that smooth adjustment can be determined by parameters that reflect the speed of the adjustment as well as the average location, threshold, or asymmetric nature of that adjustment.

**Model Specification**

We examine three classes of non-linear models, in addition to standard linear models. The first model approximates unknown nonlinearity using a Taylor approximation. The second specifically allows for asymmetries in adjustment and the third is a threshold model. Defining a bivariate vector \( y_t' = (y_{1,t}, y_{2,t}) \) for \( t=1,\ldots,T \), all the models we employ are expressed as:

\[
\Delta y_t = \mu + F(y_{t-1}) + \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \nu_t
\]  

(1)

where \( \nu_t \) is assumed to be an independently and identically normally distributed vector with a mean of zero and variance covariance matrix \( \Omega \). The function \( F(y_{t-1}) \) may take a number of forms such as the standard Vector Autoregression (VAR):

\[
F^0(y_{t-1}) = A(y_{t-1})
\]  

(2)

where \( A \) is a square coefficient matrix. The absence of any rank restrictions on the matrix \( A \) implies that \( y_t \) is a stationary vector. The second model we employ is the standard VECM, whereby equation (1) defines the adjustment process, but \( F(y_{t-1}) \) is defined as:

\[
F^t(y_{t-1}) = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} (y_{y_{t-1}} - \beta_1 y_{2,t-1} - \beta_0 - \beta_2 t).
\]  

(3)

The coefficients \( \beta_1 \) and \( \beta_2 \) may be zero, whilst \( \alpha_1 \) and \( \alpha_2 \) are adjustment coefficients with respect to the disequilibria:

\[
e_{t-1} = y_{1,t-1} - \beta_1 y_{2,t-1} - \beta_0 - \beta_2 t
\]  

(4)

that are assumed to be stationary with zero mean. As generalisations of the linear structure, our nonlinear models take the form:

\[
F(y_{t-1}) = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} g(e_{t-1})
\]  

(5)

where disequilibria are defined as in equation (4), with the function \( g(e_{t-1}) \) taking a number of potential specifications. For model \( F(y_{t-1})^t \), we assume that the specification of the adjustment function is unknown, with the function been everywhere differentiable, with \( g(0)=0 \) and, since (5) is only identified up to a proportional constant, with \( \nabla g(0) = 1 \) (\( \nabla \) denoting the gradient). The associated third order Taylor expansion is:
\[ g^u (e_{t-1}) = e_{t-1} + \frac{1}{2} \nabla^2 g (0) e_{t-1}^2 + \frac{1}{6} \nabla^3 g (0) e_{t-1}^3 \]  

(6)

and \( \eta_1 = \nabla^2 g (0) \) and \( \eta_2 = \nabla^3 g (0) \) are parameters to be estimated. Therefore, the Taylor approximation of this unknown adjustment function modifies the impact of disequilibria on the variables in a manner that is determined by the parameters \( \eta_i \).

A number of papers have posited the possibility of asymmetric adjustment (e.g. Granger and Lee, 1989). Therefore, model \( F(y_{t-1})^m \) has another interesting specification for the adjustment function as follows:

\[ g^m (e_{t-1}) = (e_{t-1} + \lambda |e_{t-1}|) \]  

(7)

The value of the threshold parameter \( \lambda \) in (7) determines adjustment towards the long run equilibrium, in a manner that it is potentially asymmetric. Additionally, \( \lambda \) can be bounded so that it lies between the range (-1,1). At these extremes the function reflects ‘one-sided’ adjustment. For example, if \( \lambda = 1 \):

\[ e_{t-1} + \lambda |e_{t-1}| = 2e_{t-1} \text{ for } e_{t-1} > 0, \]

\[ e_{t-1} + \lambda |e_{t-1}| = 0 \text{ for } e_{t-1} < 0. \]

(8)

Under these circumstances, adjustment would take place to positive, but not negative disequilibria. If \( \lambda = -1 \), the opposite would be true. In between these extremes we have asymmetric adjustment towards equilibrium depending on the sign of the disequilibrium term.

Finally, we have the standard threshold model, \( F(y_{t-1})^v \) in which the adjustment mechanism in equation (5) is expressed as:

\[ g^v (e_{t-1}) = \delta_i e_{t-1} + (1 - \delta_i) \gamma e_{t-1} \]

\[ \delta_i = 1 \text{ if } |e_{t-1}| > \lambda, \text{ and } \]

\[ \delta_i = 0 \text{ if } |e_{t-1}| \leq \lambda \]  

(9)

This model can be simplified, as in Hansen and Seo (2002), by setting \( \gamma = 0 \), in which case there is adjustment back towards equilibrium only if the disequilibria are outside of a threshold value \( \lambda \). It is this simplified model that we employ in this article. A problem associated with threshold models such as that reflected by (9), is the lack of identification of the threshold parameters, if no disequilibria lie inside, or alternatively outside, the threshold. This problem can be addressed by specifying that a minimum proportion of observations must lie inside and outside the thresholds. However, such an approach introduces difficulties when using a Bayesian framework, since it is desirable that all prior distributions integrate to one (i.e. have a proper prior, see discussion on priors and testing in the following section). If a prior \( f(\lambda) \) is defined over a subset of positive real numbers, identification of the model requires that this prior is multiplied by an indicator function as follows:
The indicator function, \( I(\lambda, \{e_{t-1}\}) \), takes the value of one, if a given proportion of the \( \{e_{t-1}\} \) lie inside and outside the threshold, otherwise, it takes the value of zero. The dependence of \( \{e_{t-1}\} \) on the data and the values of \( \beta \), means that the integrating constant for this prior cannot be easily obtained. Accordingly, we propose an alternative approach based on equation (9), but where the proportion of observations inside the threshold interval (\( \pi \)) becomes the parameter that is estimated. A symmetric interval around zero that contains \( 100(\pi)\% \) of the disequilibrium terms \( \{e_{t-1}\} \) is then constructed to give thresholds \( (\lambda_u(\pi), \lambda_l(\pi)) \), such that:

\[
Prob(e_{t-1} > \lambda_u(\pi)) = \frac{\pi}{2} \quad (11)
\]

\[
Prob(e_{t-1} < \lambda_l(\pi)) = \frac{\pi}{2}
\]

Such an approach is almost identical to estimating the thresholds directly. However, it has the advantage that a prior for the parameter of interest (\( \pi \)), that integrates to one, can be more easily specified, even when \( \pi \) is bounded between some sub-interval of (0,1). The values of the implied threshold can be calculated indirectly given the parameter estimates and the data.

**Estimation and Inference**

The estimation of nonlinear models can be problematic when using maximum likelihood. The likelihood function may no longer be unimodel and differentiable at all points in the parameter space, latent data may be present, and/or parameters may not be identified within certain regions of the parameter space. The classical literature has proposed a variety of solutions to these difficulties. MSVECMs are often estimated by means of the iterative Expectation-Maximisation algorithm (Krolzig, 1997) in conjunction with an adaptation of probability smoothers (Hamilton, 1989), a method that can be thought of as a discrete version of the Kalman filter. Testing the null of linear cointegration against the alternative of non-linearity is performed in a straightforward manner within the MSVECM framework. For example, Krolzig et al. (2002) tests for the significance of the intercept shifts by means of a Wald statistic. For the TVECM, Hansen and Seo (2002) proposed a quasi-MLE method based on a grid-search over the cointegrating vector and the threshold parameter. Bootstrapping provides the basis for parameter inference, along with supremum based tests that take into account the non-standard distributions of parameters that are not identified under the null hypothesis. Kapetanios et al. (2006) specify a STVECM in which the adjustment function follows the exponential smooth transition autoregressive functional form (ESTAR). They estimate the cointegrating regression, as a first step, before the estimation of a STVECM in which the exponential adjustment function is approximated by a first-order Taylor expansion. The specification of the STVECM in Kapetanios et al. (2006) allows them to develop a test for the null of no cointegration.
against the alternative of non linear ESTAR cointegration that is analogous to the Engel and Granger (1987) test for linear cointegration. However, no threshold parameters are estimated. Seo (2004) focuses on both exponential and logistic functional forms for the specification of the indicator function, and develops a LM test for smooth adjustment nonlinearity, the associated asymptotic theory and bootstrap inference. Likewise, tests are based on fixing the smooth adjustment parameters to their maximum likelihood value, based on a grid-search.

By contrast, on a computational level, Bayesian estimation of threshold and non-linear models requires only the fairly standard use of Gibbs Sampling with Metropolis steps, which are special cases of MCMC algorithms. Work employing MCMC algorithms is now becoming commonplace. Comprehensive descriptions of these algorithms can be found in textbooks such as Bauwens et al. (1999), whilst applications include the estimation of TVECMs (Balcombe, 2006; Balcombe et al., in press). Recently, Bayesian methods have been applied to Markov Switching Models (Paap and Van Dijk, 2003). The application of MCMC algorithms are no more demanding than classical bootstrapping methods and do not require the likelihood to be differentiable. As discussed in Bauwens et al. (1999), jagged likelihood functions do not create any particular difficulties when using MCMC algorithms. Non-identification of parameters, however, can create difficulties, but arguably the problems are less acute than in the classical case. Bayesian approaches allow inference to be based on moments of the posterior distributions that are simulated by the MCMC algorithms and by the calculation of the marginal likelihood, that is the expected value of the likelihood function under the prior distribution. Each model can be evaluated and compared according to the associated value of the marginal likelihood. The Bayesian approach naturally facilitates both nested and non-nested model comparison, an attribute that is particularly useful in the context of this work, since the non-linear models that we investigate are non-nested.

Another positive aspect of the Bayesian approach is that it is less susceptible to problems relating to sequential testing. For example, pretesting for unit roots is not required. Models that embody the assumption of unit roots and/or cointegration can be directly compared to those which embody alternative assumptions. Classical estimation of the MSVECM, TVECM and STVECM, usually consists of multiple steps with researchers often pretesting the data for unit roots in a first step, estimating unrestricted VARs in a second step to determine lag-lengths, testing for non-cointegration in a third step, before estimating nonlinear parameters and conducting further inference about these parameters. In most classical applications, non-cointegration is tested for, prior to testing for weak and strong exogeneity of other series. However, imposing exogeneity on one, or another of the series prior to testing may alter the outcome of tests. Using Bayesian methods, a comparison between cointegrated and non-cointegrated models with, or without exogeneity is

---

3 Some degree of pretesting may be required when using Bayesian methods simply due to the practicalities of reducing the number of models that are to be estimated and compared. For example, in this paper we also make a decision about the number of lags in the models using an unrestricted Vector-Autoregression.

4 In the case of strong exogeneity, this sequence is required because the existence of unit roots and cointegrating relationships can alter the distributions of classical tests for Granger Causality.
possible, without the necessity for making sequential decisions about cointegrating rank before testing for exogeneity and vice versa.

Nevertheless, the Bayesian approach has its own difficulties, as it requires that prior distributions are specified for each of the parameters in the model. There are some complex issues related to the specification of valid priors for each of the models we estimate. Each of the models (0, I, … IV) described in the previous section have the following associated parameter sets:

\[
\begin{align*}
\Phi^0 &= (\mu, \Phi, \{\phi_i\}, \Omega) \\
\Phi^I &= (\mu, \alpha, \{\beta_i\}, \{\phi_i\}, \Omega) \\
\Phi^{II} &= (\mu, \alpha, \{\beta_i\}, \{\phi_i\}, \Omega, \{\eta_i\}) \\
\Phi^{III} &= (\mu, \alpha, \{\beta_i\}, \{\phi_i\}, \Omega, \lambda) \\
\Phi^{IV} &= (\mu, \alpha, \{\beta_i\}, \{\phi_i\}, \Omega, \pi)
\end{align*}
\] (14)

Equation (1) is linear in its parameters \(A\), or \(\alpha\) for a given model specification (0, I, … IV), assuming the remaining parameters are known. Accordingly, endowing these parameters with the same priors that would be given to them in the standard treatment of the Bayesian linear model is advantageous in the respect that this will lead to conditional posterior distributions of a recognisable form. Therefore, across all models, the priors for parameters \(\mu, \Phi, \{\phi_i\}, \Omega\), should they not be assigned a zero value, are specified as independent multivariate normal with mean zero and variances equal to a large number (herein we use \(10^6\)). The priors for \(\Omega\) for all models, are specified as weakly informative inverse Wishart distributions, as would be commonly done in the standard Bayesian Linear regression framework. Essentially, this is the approach outlined in Geweke (1996). Our use of proper, rather than ‘flat’ priors facilitates the calculation of marginal likelihoods for each model.

The specification of priors for the cointegrating parameters warrants detailed discussion. Bayesian estimation of cointegrating relationships has been discussed in Bauwens et al. (1999), Chao and Phillips (2002), Kleibergen and Paap (2002), Strachan (2003) and Strachan and Inder (2004). A particularly useful review is provided by Koop et al. (2004). The first, and most obvious problem, is that if \(\alpha=0\), the cointegrating space is not identified. The specification of a non-diffuse prior for the cointegrating space partially solves this problem, but the prior can influence the value of the marginal likelihood of the model. Moreover, with \(p\) variables there is the potential to have up to \(p-1\) cointegrating vectors, and thus identification requires that there are as many identifying restrictions in each equation, including normalisation restrictions, as there are cointegrating relationships.

In principle, having made a decision on the number of cointegrating vectors and a given set of identifying restrictions, priors can be placed on the remaining free parameters within the cointegrating vectors. However, Strachan and Inder (2004) show that the posterior densities are likely to be highly dependent on the arbitrary choice of identifying restrictions and for this reason they examine priors placed on the cointegrating space, rather than on the free parameters. In the bivariate linear case, the choice of identifying restrictions is little more than a choice of how to
normalise the cointegrating vector. If one variable is assigned a coefficient equal to one, then it remains only to place a prior on the remaining parameter. Even so, Bauwens et al. (1999) show that a uniform prior on the remaining coefficient, results in a prior with undefined moments. Likewise, Chao and Phillips (2002) show that the use of the Jeffries prior for the cointegrated model delivers a posterior that does not have finite moments. Villani (2000) demonstrates that a uniform prior on the Grassman manifold, that is a representation of the cointegrating space, translates to a Cauchy distribution on the cointegrating parameters, whilst Bauwens et al. (1999) show that a t prior is sufficient to derive ‘poly-t’ posteriors for cointegrating parameters with finite moments.\(^5\)

We conducted a number of Monte Carlo trials using t-priors with one, two and three degrees of freedom. In general, the results were in close accordance, with the Monte Carlo trials suggesting that the priors were dominated by the data quickly, given the Monte Carlo data was generated from cointegrated data, in which case the cointegrating parameters were identified and the data informative about these parameters. The marginal likelihoods showed a clear preference for the models that were estimated in accordance with their data generation. However, the Cauchy distribution prior gave posteriors for the cointegrating coefficients with extremely ‘fat tails’ for data which were not cointegrated, or with extremely small adjustment coefficients (i.e. near non cointegrated series). When using positive degrees of freedom, the results more accurately reflected the data generating process for these series, with the data quickly dominating the priors.

For the analysis, we adopted two sets of priors, Cauchy, and t priors with 3 degrees of freedom and non-zero means. For the t-priors with 3 degrees of freedom, the mean for \(\beta_1\) was set to -1 and the mean for \(\beta_2\), for the models which included a time trend, was set to zero. In each model, \(\beta_0\) was set so as to give the disequilibrium term \((e_{t-1})\) a mean of zero. These informative priors were set on the grounds that, in price transmission studies using logged variables, there is a compelling basis for believing that the cointegrating vector may take the form \((1,-1)\). The t-priors that we have adopted reflect a prior belief that there is approximately 50-60 percent chance that the cointegrating coefficient lies between 0 and -2. We would argue that such priors are entirely reasonable, however, we also provide the marginal likelihood values for the linear models using Cauchy priors, which we take to be non-informative, in order to demonstrate that our results are not simply a function the priors selected.

As the normalization of the cointegrating vector can have a substantive impact on the results (Strachan and Inder, 2004), we experimented with exchanging the assignment of the unitary coefficient both in Monte Carlo trials and while using the price data employed in the empirical section. There were small differences in the results, depending on the choice of normalisation, albeit not large enough to alter conclusions, or preferences with regard to different models.\(^6\)

---

\(^5\) The proofs relating to this are for the linear case. Therefore, we cannot make a categorical claim that the same results extend to the non-linear case.

\(^6\) We provide comments on these differences in the Applications and Results section.
In model II, the priors for the parameters \( \{ \eta_i \} \) were set as standard normal. Again, these are informative and reflect a small bias towards the linear model.\(^7\) In model III, the prior for \( \lambda \) was standard normal, but truncated below -1 and above 1. In model IV, the prior for \( \pi \) was also uniform, but restricted to lie between the interval (0.2, 0.8), reflecting a restriction that a minimum of 20 percent of observations must lie outside the threshold, with a maximum of 80 percent of observations lying outside the threshold.\(^8\)

As in Geweke (1996), this framework implies that the conditional posteriors for \( \mu, \{ \text{Vec} (\phi) \}, \text{Vec} (A) \), or \( \alpha \), that is the posteriors given all other parameters in the model, are multivariate normal. Likewise, the conditional posterior for \( \Omega \) is inverse Wishart, as the variance covariance matrix is being treated exactly like parameters from a Bayesian linear model, given that the other parameters in the model are treated as known. Such a choice would be valid unless, for models I-IV, the function \( g(e_{i,t}) \) did not vary, in which case the parameters \( \alpha \) would not be identified. Our priors ensure that this is not the case. The estimation of the stationary VAR in equation (1) can be performed using Gibbs Sampling as it would be performed in any linear model. For models I-IV, the parameters \( \{ \eta_i \}, \{ \beta_i \}, \lambda, \pi \) cannot be mapped using Gibbs Sampling, as do not have conditional posterior distributions of a known form. Therefore, Gibbs Sampling is used in conjunction with random walk Metropolis-Hastings steps (Koop, 2003 and Bauwens et al., 1999 provide descriptions for Gibbs Sampling and Metropolis-Hastings algorithms).

The marginal likelihoods for each model were estimated using the Gelfand and Dey method (see Gelfand and Dey, 1994, or Koop, 2003, pp.104-106, for details). Care was taken so that each model had priors that integrated to one, regardless of whether the model had exogeneity restrictions imposed, or contained rank restrictions. Accordingly, the marginal likelihood of a given model can be compared to that of another model, with or without rank restrictions, and no explicit prior for the cointegrating rank needs to be specified. For example, comparing the fully differenced model to a cointegrated model would automatically impose the restrictions \( \alpha=0 \) and \( \beta_1=0 \), along with zero restrictions on any of the parameters characterising nonlinearity. Thus, testing for the cointegrating rank is treated as a test for exclusion restrictions for some of the parameters.

Strachan and Inder (2004) observe that the variance of the prior parameters on the adjustment vector \( \alpha \) can have an influence on the Bayes ratio for a model with and without restrictions. For models I, II, III, and IV, we imposed the condition that the values of \( \alpha \) are less than one in absolute value and are appropriately signed so that each variable would adjust downwards should it be above its long-run equilibrium state.\(^9\) This condition is not imposed when weak exogeneity, or non cointegration

---

\(^7\) Again Monte-Carlo experiments suggested that these priors were quickly dominated by the data where nonlinearity was present.

\(^8\) This restriction is similar to that used in other threshold work.

\(^9\) This can be implemented using a rejection step in the Gibbs Sampler. Support for this restriction can also be assessed through the calculation of the marginal likelihood. These restrictions are not a requirement of the VECM, however, from a structural modelling point of view, they make good sense.
condition is imposed in which case one, or both elements of $\alpha$ are set to zero. Our restriction that the adjustment coefficients lie between 0 and 1, in absolute value, effectively makes any normal prior with a large variance approximately flat over this interval, whilst the resulting Bayes ratios are nearly invariant to the level of the variance set for the normal priors on $\alpha$.

Finally, in all the empirical analysis that follows we used four separate MCMC chains. Each chain had burn-in of 2,000 iterations, followed by a further 20,000 from which we drew every 10th observation. The resulting 8,000 observations were then used to map the posterior distributions. Convergence was monitored visually, and in addition we conducted modified t-tests for subsets of the parameters in which the variance estimates of the parameters were replaced with the estimate of the long-run variance matrix (the spectral density matrix evaluated at frequency zero). All the models presented herein were found to have converged well.

**Application and Results**

We apply the models discussed above to explore the relationship between the logarithmic transformations of weekly prices for crude oil, ethanol and sugar in Brazil, expressed in Brazilian Real, for the period between July 2000 and May 2006. This period of time is characterised by technological change in terms of the introduction of the flexible fuel cars from 2003 onwards, as well as, during the last year of the sample, by high levels of oil prices.

While the high frequency of the series enables us to obtain a relatively large sample size over a short period, it also entails that adjustment towards equilibrium may appear small due to the high frequency of the series. Time plots of the logarithmic transformations of data series, subtracted from their means for comparison, are presented in Figure 1. The series are trending upwards in a way that is consistent with stochastic trends. The sugar and ethanol price series seem to be more closely correlated than the oil series, but visually, the plot suggests the possible existence of cointegrating relationships between the price series. The oil market is likely to be primary driver of the ethanol and sugar series. Consequently, *a priori*, we expect that, should cointegration between prices exist, oil prices may be weakly exogenous, not adjusting so as to preserve a long-run equilibrium, whilst sugar and ethanol prices may show a relatively small adjustment back towards equilibrium within one week.

---

10 Data on sugar and ethanol have been collected by UNICA (www.unica.br), data on world oil prices have been collected from the Energy Information Administration of the United States government (www.iea.doe.gov) and information on exchange rates from the Federal Reserve Bank of St Louis (http://research.stlouisfed.org).
Our empirical analysis entailed the estimation of over 70 variants of the models outlined in the previous sections. Herein, we only present a subset of these results. For the results pertaining to oil-sugar and oil-ethanol price pairs, the marginal likelihood consistently indicated a preference for specification with only one lag of the differenced variables in the error correction model (two lags in an unrestricted VAR). For the ethanol-sugar pair, two lags were preferred in the error correction model (three lags in an unrestricted VAR). The inclusion of time trends in the unrestricted VARs and cointegrating relationships also decreased the marginal likelihood values. Consequently, throughout the empirical analysis presented here, the lag-lengths were kept be fixed and no deterministic time trends were included. Our key results, however, are not dependent on these restrictions.

As our estimation strategy does not require pretesting for unit roots and then for cointegration, the absence of unit roots in the data should lead to a larger marginal likelihood for the unrestricted VAR model, relative to a cointegrated model, or a fully differenced model. Table 1 presents the logged marginal likelihoods for linear models, for each of the pair-wise relationships between the three series. Marginal likelihoods are presented for unrestricted, fully differenced and cointegrating VARs. Fully differenced models imply no cointegrating relationships between the variables, that is \( A=0 \) in equation (1), but are nested within the cointegrated models. The latter impose a rank restriction on the matrix \( A \), thus becoming variants of model I. Weak exogeneity, in the context of VECMs, imposes the condition that either \( \alpha_1 \), or \( \alpha_2 \) are equal to zero in equation (3), but not both. The strong exogeneity of a given variable requires that the variable is weakly exogenous and, in addition, is not Granger-caused by the other variables.
The unrestricted VAR is characterised by the lowest marginal likelihood in each pair-wise case, thus supporting the existence of unit roots in the behaviour of the prices. A comparison between the marginal likelihoods of fully differenced and cointegrating VARs is used to provide support for cointegration. Without imposing exogeneity, the results are mixed, nevertheless, using the t-priors with 3 degrees of freedom, the cointegrated models are slightly preferred to the fully differenced models, for the oil-sugar and sugar-ethanol price pairs, whilst for the oil-ethanol price pair, the marginal likelihood is lower than that of the fully differenced model. The Cauchy priors are even less supportive of cointegration, since the marginal likelihoods are lower across all price pairs. As noted above, alternative identification restrictions, amounting to the choice of which variable to normalise in the cointegrating vector, can have an impact on the results when using informative priors. For the cointegrated models, the cointegrating vector was normalised so that the coefficient of variables labelled as \( Y \) is, in each model, equal to one. We also experimented by changing the normalisation to assess the differences in the logged marginal likelihoods, using the 3 degrees of freedom t priors. The normalisation change resulted in marginal likelihoods amounting to 1417.99, as compared to previously estimated 1418.51 for oil-sugar, 1532.96 as compared to 1532.94 for sugar-ethanol, and 1302.80 as compared to 1303.24 for oil-ethanol. Such differences would be insufficient to change our conclusions using the t-priors.

Support for cointegration improved dramatically for all price pairs when weak exogeneity conditions were imposed, regardless of the priors. The results unambiguously support pair-wise cointegration between each of the variables, once oil prices are treated as weakly exogenous with respect to the other two prices and sugar prices are treated as weakly exogenous to ethanol prices. On the other hand, the marginal likelihood decreases if weak exogeneity is imposed on each of the variables labelled as \( Y \). Given the support for weak exogeneity for each of the \( X \) variables, strong exogeneity is imposed on oil prices in the oil-sugar and oil-ethanol relationships and on sugar prices in the sugar-ethanol relationship. The imposition of strong exogeneity on these variables resulted in an increase in the marginal likelihoods, for both the
informative t-priors and the Cauchy prior. We again experimented with changing the normalisation for some of the cointegrated models, where exogeneity was imposed for the informative t-prior. The marginal likelihoods for the cointegrated models declined slightly, in most cases up to -0.5, but left our overall conclusions unchanged.

In the context of a VECM, weak exogeneity implies that the stochastic trends in the system are composed of the cumulative shocks to the weakly exogenous variable. Therefore, the evidence is consistent with our a priori expectations that oil prices are the main long run drivers of the other two prices. Moreover, as oil prices appear to be strongly exogenous to the other two prices, there is no price transmission from the other two markets into the oil market. Less expected, perhaps, is the result that it is the shocks in the sugar market that are strongly exogenous to ethanol shocks, suggesting that the causal chain runs directly from oil prices to sugar, rather than through the ethanol market. This result may indicate that sugar markets anticipate price rises following oil price hikes, rather than sluggishly responding to changes in the ethanol price.

The results related to the analysis of nonlinear behaviour are presented in Table 2, where we adopt the t-priors with 3 degrees of freedom for the cointegrating parameters and impose strong exogeneity conditions on all models.\textsuperscript{11} We calculate Bayes ratios that compare the performance of each model to the best performing linear model, that is the model with the highest marginal likelihood (see also Table 1).\textsuperscript{12} The Bayes ratios suggest that all three non-linear models (II, III, IV) are preferred to the linear model for the oil-sugar pair and the oil-ethanol pair. However, linear adjustment is supported for the sugar-ethanol pair. For the oil-sugar pair, the asymmetric adjustment model (model III) is preferred, whilst the threshold model (model IV) is preferred for the oil-ethanol pair. For the first two price pairs, the results of the Taylor expansion models (model II) are in accordance with the asymmetric model (model III) since the quadratic coefficients, \( \eta_1 \), have a negative sign, as do the asymmetric coefficients, \( \lambda \). However, the cubic coefficients, \( \eta_2 \), for model II are very small, suggesting, in conjunction with model III, that adjustment in response to disequilibria is stronger if the sugar and ethanol prices are below their long-run equilibriums, as this is determined by oil prices, than if they are above their long-run equilibriums. The sugar-ethanol pair have very small estimates for the parameters characterising nonlinearity within \( g(e_{t-1}) \), which is consistent with the small Bayes ratios associated with these models. The oil-sugar and oil-ethanol price pairs have similar parameter estimates for threshold behaviour (model IV). The value for \( \pi \) is around 0.4 for both pairs, indicating that around 40 percent of the sample values lie outside the thresholds. The implied thresholds are estimated around 0.22 for oil-sugar and 0.18 for oil-ethanol. Since the variables are logged, the threshold models suggest that sugar and ethanol can deviate up to 22 percent and 18 percent from their equilibrium values respectively, before adjusting in response to disequilibria.

\textsuperscript{11}The results for Cauchy priors are not reported and do differ to some extent, but not in a way that would substantively change our conclusions.
\textsuperscript{12}A Bayes ratio for any two models can be calculated by taking the exponential of the difference between the two logged marginal likelihoods, reflecting the posterior odds that the model with the higher marginal likelihood is preferred over the other, providing the prior odds are equal. The higher the Bayes ratio, the more support a given model has relative to the linear model.
Table 2: Comparison of Non-Linear Models

<table>
<thead>
<tr>
<th>Variable</th>
<th>MODEL I</th>
<th>MODEL II</th>
<th>MODEL III</th>
<th>MODEL IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cubic</td>
<td>Asymmetric</td>
<td>Threshold</td>
<td></td>
</tr>
<tr>
<td>Oil (X)-Sugar (Y)</td>
<td>$\beta_1 = -0.56$</td>
<td>$\beta_1 = -0.51$</td>
<td>Bayes Ratio (III/I)=4.55</td>
<td>Bayes Ratio (IV/I)=3.07</td>
</tr>
<tr>
<td></td>
<td>$\text{Stdv}(\beta_1) = 0.16$</td>
<td>$\text{Stdv}(\beta_1) = 0.16$</td>
<td>$\eta_1 = -1.07$</td>
<td>$\lambda = -0.61$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\eta_2 = 0.015$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$\text{Stdv}(\eta_2) = 1.00$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\text{Bayes Ratio (II/I)} = 2.01$</td>
<td></td>
</tr>
<tr>
<td>Oil (X)-Ethanol (Y)</td>
<td>$\beta_1 = -0.59$</td>
<td>$\beta_1 = -0.67$</td>
<td>Bayes Ratio (III/I)=1.79</td>
<td>Bayes Ratio (IV/I)=3.9</td>
</tr>
<tr>
<td></td>
<td>$\text{Stdv}(\beta_1) = 0.20$</td>
<td>$\text{Stdv}(\beta_1) = 0.24$</td>
<td>$\lambda = -0.51$</td>
<td>$\eta_1 = 0.164$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sugar (X)-Ethanol (Y)</td>
<td>$\beta_1 = -0.88$</td>
<td>$\beta_1 = -0.89$</td>
<td>Bayes Ratio (III/I)=0.57</td>
<td>Bayes Ratio (IV/I)=0.35</td>
</tr>
<tr>
<td></td>
<td>$\text{Stdv}(\beta_1) = 0.14$</td>
<td>$\text{Stdv}(\beta_1) = 0.16$</td>
<td>$\lambda = -0.047$</td>
<td>$\eta_1 = 0.002$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: All Bayes Ratios are relative to the last models in Table 1. Strong Exogeneity for X is imposed on all models. Parameter values are the means of the posterior distributions. $\beta_1$ should be interpreted as the cointegrating coefficient in the normalisation \( \ln Y + \beta_1 \ln X \). \text{Stdv}(.) of parameter values are the standard deviation of the posterior distributions.

The estimates for the cointegrating coefficients for the oil-sugar price pair vary between -0.48 to -0.64, indicating that, in the long run, a 10 percent rise in oil prices will translate into approximately a 5 percent increase in the price of sugar. The estimates for the oil-ethanol pair are slightly higher, ranging from -0.59 to -0.71. The estimates for the sugar-ethanol pair are also higher, being between -0.88 and -0.90. In general, users of econometric results are likely to prefer estimates that are not dependent on the outcome of a single model, particularly if different models produce quite different estimates of the parameters of interest. Bayesian analysis has a coherent rationale for a weighted averaging of parameters, and other quantities, that are produced by different models. In Bayesian model averaging, the weights are constructed from the relative marginal likelihoods for the models over which averaging.
is to take place. Models that have a higher marginal likelihood receive greater weight. When model averaging is used to produce cointegrating coefficient estimates from the estimates in Table 2, the cointegrating coefficients are -0.55 for oil-sugar, -0.62 for oil-ethanol, and -0.89 for sugar-ethanol.\(^{13}\) In addition, for each price pair, a test for a unitary cointegrating coefficient was conducted by imposing the condition that \(\beta_1 = -1\), and once again calculating the marginal likelihood that model. Imposition of this restriction on each of the top performing models in Table 2, resulted in a dramatic reduction in the marginal likelihoods of the oil-sugar and oil-ethanol price pairs with the Bayes ratio amounting to 0.059 for the oil-sugar price pair and to 0.069 for the oil-ethanol pair. Therefore, there is powerful evidence against the hypothesis of unitary cointegration for these two pairs. However, for the sugar-ethanol pair, the Bayes ratio amounted to 1.71, supporting the hypothesis of \(\beta_1 = -1\).

### Table 3: Dynamics of Adjustment

<table>
<thead>
<tr>
<th></th>
<th>(\Delta\text{(Oil)})(_t)</th>
<th>(\Delta\text{(Sugar)})(_t)</th>
<th>(\Delta\text{(Ethanol)})(_t)</th>
<th>(\Delta\text{(Ethanol)})(_{t-1})</th>
<th>(\Delta\text{(Sugar)})(_{t-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model III</td>
<td>0.0026</td>
<td>-0.0026</td>
<td>0.0002</td>
<td>0.0004</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>(0.0026)</td>
<td>(0.0019)</td>
<td>(0.0019)</td>
<td>(0.0019)</td>
<td>(0.0012)</td>
</tr>
<tr>
<td>(g(e_{t-1}))</td>
<td>.</td>
<td>-0.029</td>
<td>-0.042</td>
<td>-0.056</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.011)</td>
<td>(0.015)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta\text{(Oil)})(_{t-1})</td>
<td>0.226</td>
<td>0.011</td>
<td>(0.0012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.029)</td>
<td>(0.042)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta\text{(Sugar)})(_{t-1})</td>
<td>.</td>
<td>0.781</td>
<td>0.143</td>
<td>1.119</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.084)</td>
<td>(0.052)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta\text{(Ethanol)})(_{t-1})</td>
<td>.</td>
<td>.</td>
<td>0.559</td>
<td>0.588</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.047)</td>
<td>(0.056)</td>
<td></td>
</tr>
<tr>
<td>(\Delta\text{(Sugar)})(_{t-2})</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-0.048</td>
<td>-0.441</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.085)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>(\Delta\text{(Ethanol)})(_{t-2})</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-0.192</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.056)</td>
</tr>
</tbody>
</table>

Note: Sugar and ethanol prices each have two equations, because they differ depending on the variables that they are paired with. As oil prices were found to be strongly exogenous when paired with sugar and ethanol, they are therefore the dependent variable in only one equation. Standard deviations are shown in parenthesis.

Finally, the remaining coefficients of the VECMs are presented in Table 3. Here, the parameters of most interest are the adjustment parameters with respect to \(g(e_{t-1})\).

\(^{13}\) We could average over a larger set of models including the linear models without exogeneity etc. However, given the large difference in the marginal likelihoods, these models would be given near zero weight.
The small values of the adjustment parameters are consistent with the high frequency of the data. In addition, lagged differenced oil prices do not appear to have any influence on either sugar or ethanol, with the main source of adjustment to oil prices being in response to disequilibria.

![Figure 2: Asymmetric vs Symmetric Responses](image)

Disequilibria of Sugar Prices in Response to an Oil Price Shock

Utilising the standard linear and asymmetric sugar-oil price models, we compared the response of sugar to a one standard deviation shock in oil prices. A plot of the simulated disequilibria in sugar prices, as dictated by oil prices, is depicted in Figure 2. The positive and negative disequilibria induced by an oil shock fade very slowly in a symmetric fashion for the linear model. Under the asymmetric model (Model III), the negative disequilibria dissipate more quickly whilst the positive disequilibria more slowly, relative to those of the linear model. A similar picture would characterise the responses of ethanol to oil price shocks, using the results of Model III, since this also supported asymmetric behaviour relative to linear adjustment, and the parameters values were similar. However, as already noted, this price pair supported threshold adjustment. In the case of threshold adjustment, the impact of an oil shock would have flat impact profile on the disequilibria, since only shocks that would drive the ethanol outside its threshold would be gradually eroded. For this reason we do not present an analogous graph to Figure 2 for the ethanol-oil price pair, since it would not be very illuminating.

---

14 A similar picture would illustrate the responses of ethanol to oil price shocks, using the results of model III, since this also supported asymmetric behaviour relative to linear adjustment, and the parameters values were similar. However, as already noted, this price pair supported threshold adjustment. In the case of threshold adjustment, the impact of an oil shock would have flat impact profile on the disequilibria, since only shocks that would drive the ethanol outside its threshold would be gradually eroded. For this reason we do not present an analogous graph to Figure 2 for the ethanol-oil price pair, since it would not be very illuminating.
supported threshold adjustment. In the case of threshold adjustment, the impact of an oil shock would have flat impact profile on the disequilibria, since only shocks that would drive the ethanol outside its threshold would be gradually eroded.

Concluding remarks

The article examined the relationships between Brazilian sugar and ethanol prices and international oil prices using a Bayesian methodology. Existing VECM models were extended to allow for adjustment with respect to a nonlinear transformation of disequilibria. The nonlinear functions included a third order Taylor approximation of an unknown function, an asymmetric model, and a threshold model. The Bayesian concept of the marginal likelihood proved a key instrument in conducting this analysis and allowed a unified approach to testing for cointegration, exogeneity and nonlinear behaviour. We maintain that this method diminishes the extent to which sequential testing imparts bias further down the testing chain and also provides the basis for the model averaging for the cointegrating coefficients.

Imposing exogeneity conditions greatly improved the support for cointegration between the price series. Not surprising, oil prices were found to be the main driver of both sugar and ethanol prices in Brazil. However, that oil prices determine the long run equilibriums of these variables is a significant finding. It was further found that sugar prices Granger-caused ethanol prices, but not the other way around, suggesting a causal hierarchy from oil to sugar to ethanol, rather than the ordering of oil to ethanol to sugar which we initially expected. It is uncertain whether the current high oil prices will be sustained in the medium term. However, the long run prognosis for oil prices must be that they will continue to have a robust trend upwards. The analysis, therefore, leads to the conclusion that sugar prices will also trend upwards at around 55 percent of the rate of oil prices, as will ethanol prices (at around 60 percent of the price of oil). Consequently, ethanol will become increasingly competitive to oil, and there will be a growing incentive for a larger component of the sugar production will be used for conversion into ethanol.

Finally, nonlinear transformations of the long-run disequilibria were supported in two out of three price pairs. A threshold model had the strongest support for the oil-ethanol price pair, and a transformation that allowed for asymmetric adjustment was supported for the oil-sugar price pair. A standard linear VECM was the preferred specification for sugar-ethanol prices. Such results support our a priori expectations that high oil prices and increased competitiveness of ethanol in Brazil is characterised by threshold effects with the trend in sugar prices being determined by that of oil. Such price relationship may not be unique to sugar only, but also to other commodities that can be utilised for the production of biofuels, such as wheat, maize rapeseed, soya and oil-palm. With a number of countries, including the US and Malaysia investing in biofuel production, it is likely that high oil prices may trigger increases the price for food. Although, such developments will undoubtedly benefit the environment, as well as producers, such as the approximately 1.5 million

15 The rate of growth for oil prices implied by the parameters for the oil price equation in Table 3 is around 20% per year.
Brazilian farmers that produce cane for ethanol, they may result in increased pressure for food importing least developed countries.
References


Hansen, B. E. 1996. “Inference when a Nuisance Parameter is not Identified under the Null Hypothesis.” *Econometrica* 64:413-430.


