Understanding the Behavior and Co-Movement in Commodity Prices using Incoterms

Hany Fahmy*
University of Waterloo

April 30, 2018

Abstract

The objective of this paper is to explain the nonlinear behavior and the observed co-movement in commodity prices. We propose a novel approach of capturing nonlinearity in groups of commodity prices that tend to move together. The approach rests on using the International Commercial Terms (Incoterms), also known as *border prices*, to classify commodities in groups that tend to have similar dynamics. In each group, we fit an appropriate regime switching model with exogenous transition variable that can capture the nonlinearity of the price processes. We show that the proposed border price classification is (1) the key to finding the suitable transition variable that is capable of capturing the regime switching in each group, and (2) capable of explaining the observed co-movement in commodity prices.

Keywords: Incoterms, border prices, commodity prices, smooth transition regression.

*hfahmy@uwaterloo.ca

1 Co-movement and Nonlinearity in Commodity Prices

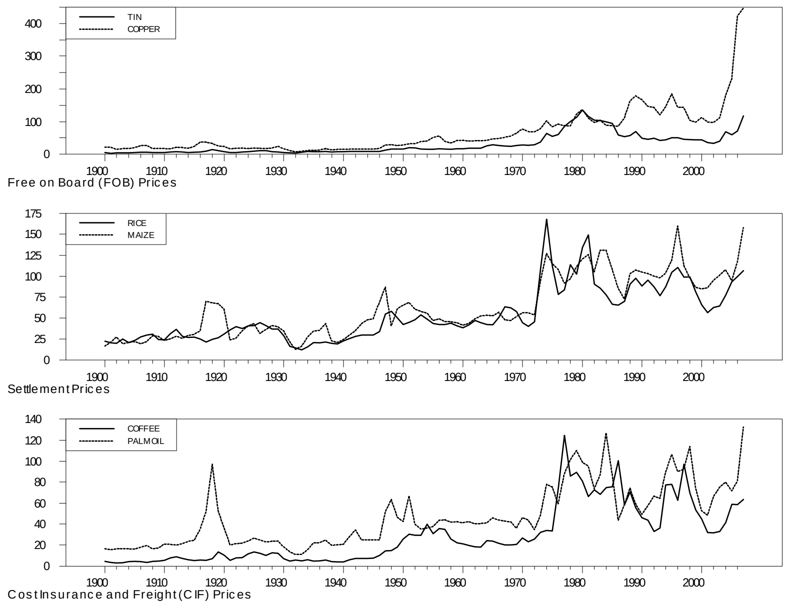

Figure 1: Co-movement in commodity prices.

Understanding the behavior of commodity prices is an issue that dominated economic discussions since the 1970's. Although the topic fell out of favor in the late 1980's and early 1990's as commodity prices generally declined, the recent observed surge in the price of oil and many other commodities brought back attention to the topic.

The dynamic behavior of commodity prices has been studied extensively since the late 1950's and early 1960's in the work of Gustafson (1958) on the theory of competitive storage and the work of Muth (1961), who introduced the rational expectations assumption in a model of commodity price formation. Many other extensions and contributions to the original theory of competitive storage have been introduced over the years.[1] The main focus of these contributions was to capture a number of stylized facts regarding the observed behavior of commodity prices such as skewness, kurtosis, high asymmetric volatility, lack of trend, and high degree of autocorrelation. The consensus is that commodity prices are *nonlinear* in general, and this nonlinearity is attributed to speculative behavior of agents for holding stocks or to unobserved demand and supply shocks.

A nonlinear time series that adhere to the previous stylized facts may be

[1] For a comprehensive survey of the literature on the subject, see Fahmy (2011 and 2014) and the references therein.

thought of as remaining in a given regime until pushed to a new regime by a shock, a series of shocks, or by a driving/transition variable. One can interpret such changing dynamic as switching between multiple equilibria. In other words, a commodity price time series that is inherently nonlinear can be best captured by a regime switching model that allows the existence of different equilibria via a mean reverting mechanism. This, in turn, means that the time series under consideration ought to be *stationary* so that it can fluctuate from one regime to another but always revert back to the mean.

Regime switching models are not new to the literature. The primary version of the switching regression model was due to Quandt (1958), who used maximum likelihood in estimating one switching point in a two regime regression system. Bacon and Watts (1971) considered two different distinct linear regression lines and developed a smooth transition technique from one linear regime to the other. Beach (1977) considered incorporating structural change in a regression model where the change occurs gradually over a known transition period. Recent accounts include Granger and Teräsvirta (1993) and Teräsvirta (1994), who combined the threshold autoregressive models and the exponential autoregressive models in a single family of models called the smooth transition regression (STR) models. The STR model is a nonlinear regression model that describes the changing dynamics of an autoregressive model from one regime to another such that the transition is smooth. The model can be considered as a generalization of the model devised by Bacon and Watts (1971). The model also nests Tong's (1978) threshold autoregressive (TAR) model, where the time series switch only between two regimes and the transition is swift.

The STR model pioneered by Granger and Teräsvirta (1993) and Teräsvirta (1994) has been used extensively in the regime switching literature. Most of the studies have been focusing on modelling nonlinearities in aggregate macroeconomic time series such as GDP, money demand functions, and industrial production. In addition to their popularity, STR models possess some appealing features. They are based on a three-stage modelling procedure starting from a specification stage, estimation, and finally an evaluation stage. Also, as opposed to other competing regime switching models, e.g., TAR models and Markov switching models, STR models have the flexibility of describing processes that can move from one regime to the other such that the transition is smooth.

Fahmy (2014) was able to capture the nonlinearity in a commodity price index using STR models. In the standard STR model of Granger and Teräsvirta (1993) and Teräsvirta (1994), the threshold/transition variable that is responsible for pushing the time series from one regime to another is one of its *autoregressive* lags. Hence, their model is usually referred to as smooth transition autoregressive or STAR for short. Fahmy (2014), however, showed that other exogenous transition variables are better than the autoregressive lags of the commodity index in capturing its dynamics. In particular, the author showed that inflation rate and oil price are the best transition variables that can capture the nonlinearity in an index of commodity prices. In this paper, we adopt Fahmy's approach in fitting STR models to individual commodity prices.

Another issue concerning nonlinearity is whether it enters the conditional mean or the conditional variance of the commodity price time series under consideration. Beck (2001), for instance, using a variation of Engle's (1982) Autoregressive Conditional Heteroskedasticity (ARCH) model found nonlinearity in storable and not in non-storable commodity prices. Naturally, the STR model, which is technically a nonlinear model in the conditional mean, can be extended to model nonlinearity in conditional variance. Hence, STR in variance models could also be considered. In fact, many variations of ARCH model could be used to capture nonlinearity in the conditional variance of the price process. Examples include, but are not limited to, Generalized ARCH (GARCH) models, smooth transition ARCH (ST-ARCH), and ST-GARCH. The existence of such variety of regime switching models poses an important question: which model should be selected to best capture the regime switching dynamic of individual commodity prices?

Another equally important concern regarding the behavior of commodity prices is the observed co-movement in some price processes. It is well known that commodity prices tend to move together in groups in response to a common transition variable. Figure 1, for instance, shows different groups of commodity prices that share the same behavior. Understanding the rationale behind this co-movement and the common transition variable that is responsible for such common behavior in each group is an issue of significant importance to international trade regulators and policy markers alike.

The main objective of this paper is to address the previous concerns. In particular, we propose a novel approach of capturing nonlinearity in groups of commodity prices that tend to move together. The approach rests on using the International Commercial Terms (Incoterms), also known as *border prices*, to classify commodities in groups that tend to have similar dynamics. In each group, we fit an appropriate regime switching model with exogenous transition variable that can capture the nonlinearity of the price processes. We demonstrate that the proposed border price classification is (1) the key to finding the suitable transition variable that is capable of capturing the regime switching in each group, and (2) capable of explaining the observed co-movement in commodity prices.

The paper is organized as follows: Section 2 explains the rationale behind the border price classification and how Incoterms can be used to select the best transition variables that are capable of capturing the dynamic behavior of commodity prices. Section 3 introduces the data set. Section 4 explains the empirical framework used to model nonlinearity in groups of commodity prices and summarizes the results. Finally, Section 5 concludes.

2 The Rationale of using Incoterms in Classifying Commodity Prices

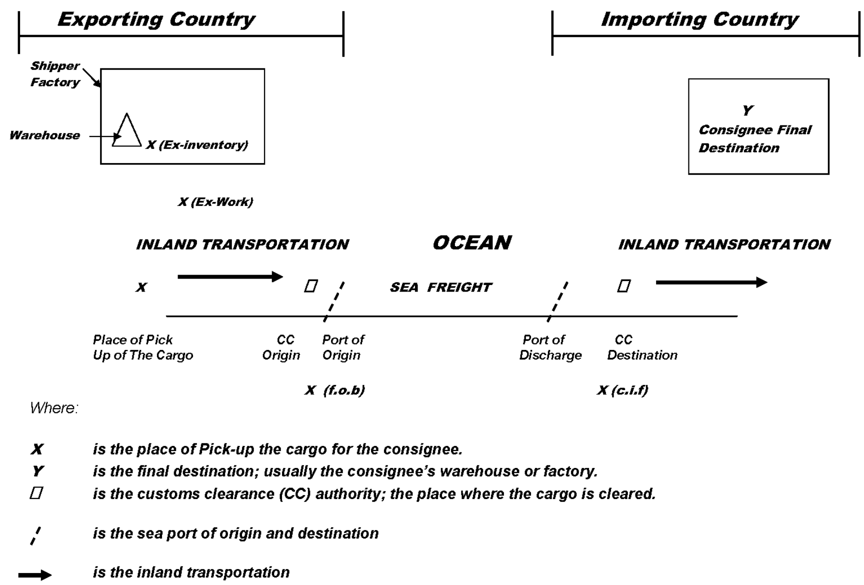

Figure 2: A typical outbound ocean freight operation.

Subject to a contractual agreement, usually a letter of credit, the shipper (owner of the goods) agrees to export a consignment (a shipment) to the consignee (importer) in exchange of the agreed selling price paid by the consignee. In addition to identifying the obligations of each party, the agreement states the type of the cargo shipped, the transport mode used, the agreed term of selling, and any other details considered by both parties.

The Incoterms, also known as border prices, are international sales terms. These terms, which are published by the International Chamber of Commerce, define the obligations of the trading parties, i.e., the obligations of the shipper and the consignee. There are many sales terms used in the shipping industry; the most frequently used are ex-works (EXW), free on board (FOB), and cost insurance and freight (CIF).

Under an EXW sale term, the shipper is obliged to deliver the goods outside his/her factory and it is the consignee's responsibility to pick up the cargo from that place and move it to its final destination. The consignee bears all the risks and shipping costs from the pick up point up to delivery at the final destination.

Ex-works terms feature many varieties; ex-inventory and ex-dock are examples.

A trading contract effected on a FOB basis implies that the exporter/shipper bears all the risks and costs of transporting the cargo from the point of origin, e.g., the exporter's factory, to the port of export in the country of origin (exit point of the exporting country). The importer/consignee bears all the risks and costs of the cargo from that point up to delivery to final destination. FOB price, thus, does not include freight, insurance, and other transportation costs needed to transfer the commodity from one country to another. A CIF price, on the other hand, is a FOB price plus insurance cost plus ocean freight cost. In other words, under a CIF contract, the exporter, in addition to the insurance, bears all risks and costs of transporting the cargo from the point of origin to the port of discharge (entry point of the importing country). Therefore, for simplicity, FOB prices are referred to as export prices at the exit point of an exporting country and CIF prices are referred to as import prices at the entry point of an importing country. Figure 2 represents a typical outbound freight operation from an exporting country to an importing country. The different points of bearing risks of transporting cargo according to a FOB and a CIF terms of sale are the exit point of the exporting country and the entry point of the importing country respectively.

Market prices are distorted prices whereas border prices are pure prices (also known as shadow prices) that don't reflect market imperfections. For that reason, all international organizations, e.g., the World Bank, the IMF, and the OECD, use border prices in their databases and in the construction of price indexes. It is worth noting that although border prices outperform market distorted prices, cost and freight prices in particular are significantly influenced by transportation cost, which is mainly driven by oil prices.

When modelling the behavior of commodity prices empirically, either using nonlinear models or any other models, the statistical and econometric techniques are directly applied to the time series under consideration. Two significant issues are, consequently, overlooked in this context. First, the treatment of an index of commodity prices is different than the treatment of individual commodity price time series. Modelling individual price series requires a deeper look into the commodity's history, its main characteristics, its major exporters and importers, and its main shipping routes. Sometimes the commodity's history can explain some of the high swings perceived in the price process. Valuable information regarding major trading routes and sales terms (usually inferred from the data sources) of a commodity can assist in identifying the potential transition variable responsible for driving the commodity price time series from one regime to the other. For instance, if the recorded commodity price is a CIF price, one expects oil price to be the potential variable responsible for the transition of the commodity price time series from one regime to the other. This can be justified through the *oil price-commodity price* connection. The idea here is that fuel surcharge constitutes a significant portion of ocean freight, which, in turn, is included in the recorded cost and freight price. Therefore, the changing dynamic of the oil price causes the recorded border price to move from one regime to the other.

The second overlooked issue in modelling the behavior of commodity prices is understanding the way the commodity is traded. Commodities can be classified according to the type of trading into two groups: (1) imported or exported goods, where actual delivery of the merchandise is mandatory, and (2) commodities traded in exchange, where physical delivery is not a must; the trading is usually done over the counter for profit sake and the recorded prices are settlement prices.[2] The price behavior of the latter group resembles that of the financial assets, which can be captured by Engle's (1982) ARCH model or one of its variants. As for the former group, i.e., the imported or exported commodities, their border prices, e.g., FOB or CIF prices, are crucial in determining the best transition variable that is responsible for their switching behavior between multiple regimes; provided, of course, the existence of nonlinearity in the data generating process.

Before proceeding to any formal analysis, we propose to classify commodities according to their Incoterms. The rationale behind the border price classification rests on how the border price can provide insight regarding the macroeconomic variable that drives the individual price series in each group from one regime to the other. The way each price series is recorded and the history of the commodity's major exporter and importer is crucial in studying and modelling its dynamics.

The standard practice followed by major institutions when recording data on commodity prices is to select the largest trading route of a commodity and detect whether the trade volume is controlled by a major exporting or importing country. If the route is dominated by a major exporter, the export price at the exit point of the exporting country, i.e., the FOB price, will be recorded; if the route, on the other hand, is dominated by a major importer of the commodity, the import price at the entry point of the importing country, i.e., the CIF price, will then be recorded. Of course there are some minor exceptions, but this is usually the standard practice. This way of recording border prices suggests that the behavior of FOB prices can be best explained by macroeconomic news variables, e.g., inflation for instance, in the major exporting country; whereas, the behavior of CIF prices can be best explained by the price of oil or news variables in the major importing country.

The data set used in this paper consists of 24 annual primary commodity prices developed by Grilli and Yang (1988) between 1900 and 1986. Recently, Pfaffenzeller, Newbold, and Rayner (2007) developed an update of the series from 1986 to 2007. A quick look at the Grilli and Yang (1988) data sources, one can notice that some commodity prices are recorded based on a FOB and CIF border prices while others, mostly metals traded over the counter in London Metal Exchange (LME), are settlement prices. In particular, six commodities are traded on a CIF; these are bananas, palm oil, coffee, timber, cotton, and tobacco; seven commodities are traded on a FOB; these are wheat, jute, rice, hides, maize, sugar, and beef; five settlement prices for copper, zinc, lead, tin, and aluminum; two spot prices for rubber and wool; one wholesale price for

[2] This type of commodities includes metals, mainly, and some grains and agricultural items.

Series	Origin	Destination	Price	Top Exporter	Top Importer	s_t
Bananas	NA†	Gulf ports	CIF	India / Brazil	USA	oil price
Palm oil	Malaysia	Netherlands	CIF	Malaysia	Netherlands	oil price
Timber	NA†	UK	CIF	NA	NA	oil price
Coffee	Average‡	New York	CIF	Brazil	US/Germany	oil price
Cotton	Memphis	Europe	CIF	USA	China	oil price
Lamb	New Zealand	London	Wholesale	New Zealand	UK	
Tobacco	NA	USA	CIF	Brazil/USA	Russia/USA	oil price
Wheat	Canada	NA	FOB	USA	China/Japan	Π_t(USA)
Jute	Bangladesh	NA	FOB	India	Various	Π_t(India)
Rice	Bangkok	NA	FOB	Thailand	Philippines	Π_t(Thailand)
Hides	USA	NA	FOB			Π_t(USA)
Maize	Gulf Port	NA	FOB	USA	Japan	Π_t(USA)
Sugar	Caribbean Ports	Various	FOB	Brazil	Russia	Π_t(USA)
Beef	Argentina	NA	FOB	Australia	USA	Π_t(Argentina)
Copper	London Metal Exchange		Settlement			
Zinc	London Metal Exchange		Settlement			
Lead	London Metal Exchange		Settlement			
Tin	London Metal Exchange		Settlement			
Aluminum	London Metal Exchange		Settlement			
Cocoa	London and US Exchange		Option Price			
Rubber	Rubber Traders Association		Spot Price			
Tea	NA		Auction			
Wool	Australia Exchange		Spot quote			
Silver	Handy & Harry					

s_t is the transition variable suggested by the border price classification.

‡ The price is arithmetic average of El Salvador, Guatemala, and Mexico.

†Not Available; $\Pi_t(x)$ is inflation rate in country x.

Table 1: Trading route, border price, top importer and exporter, and the suggested transition variable for each individual commodity in the Grilli and Yang data set.

lamb; one auction price for tea; and one option price for cocoa. Finally the silver price time series is Hary & Harmer, New York, price. Table 1 summarizes each commodity's trading route, its top exporter and importer, the recorded border price, and the suggested transition variable entailed by the border price classification.

It turned out, as we will demonstrate shortly, that inflation rate in the country of origin is the best transition variable for nonlinear FOB commodities, whereas oil price is the best transition variable for commodities that are trade on a CIF basis. At this point, it is constructive to explain why inflation and oil price are best suited for capturing the nonlinearity in FOB and CIF commodity prices. We will do so by establishing two connections; namely, the commodity price-consumer price connection and the commodity price-oil price connection.

2.1 Commodity Price-Consumer Price Connection

The commodity price-consumer price connection is well documented in the literature on the subject. For instance, Bloomberg and Haris (1995) found that a surge in aggregate demand in response to a change in money supply causes commodity prices to initially overshoot their long run equilibrium. This so-called "overshooting hypothesis" has been used by Frankel (1986), Boughton and Branson (1991), and Fuhrer and Moore (1992) among others, to explain the connection between commodity price and consumer price (inflation). Since commodities are considered inputs into production of final products, the overshooting hypothesis explains one way causality from commodity prices to prices of final goods. Recently, however, Kyrtsou and Labys (2006) and Fahmy (2014) showed that changes in broad lag inflation are transmitted to current commodity prices. In particular, Kyrtsou and Labys (2006) found evidence of *nonlinear bidirectional* Granger causality between the growth rate (logarithmic differences) of the U.S. consumer price index and the U.S. commodity price index. Fahmy (2014) found similar bidirectional causality between lag inflation and a real commodity price index.

The rationale behind this commodity price-consumer price connection rests on the fact that commodity prices react to unanticipated macro economic shocks, which could be captured by broad inflation. A commodity price time series may be thought of as remaining in a given regime until pushed to a new regime by a shock or a series of shocks, which are captured by broad inflation. One can interpret such dynamics as switching behavior between multiple equilibria. The commodity price-consumer price connection suggests that the driving/transition variable is current or lag inflation. In practice, however, it is natural to consider, in addition to inflation, different macro economic variables in the export country as transition candidates. Examples of such variables include current and one period lag unemployment, money supply measures (M1 and M2), and current and one period lag interest rates.

2.2 Commodity Price-Oil Price Connection

There is no doubt that transportation cost is the most significant factor affecting the flow of trade from one nation to another. Exporters of FOB commodities care only about the trucking cost of transporting their commodities from the production facility to the port of origin, which is the exit point at export country as shown in Figure 2. Commodities are usually transported in standard $20'$ (20-feet) and $40'$ containers. The rates for inland transportation of $20'$ and $40'$ containers are pretty much standardized and do not fluctuate much over time. Airfreight and ocean freight rates, however, are highly volatile. This high volatility is mainly due to the fluctuations in oil price, which is considered one of the significant determinants of ocean freight rates. The connection between oil prices and commodity prices is established through this fluctuation in oil prices.

In this section, we seek to establish this connection by estimating, within a factor of ten, the relation between oil price and the price of any commodity recorded on a CIF basis. Although the data on the shipping industry exists, it is not freely available and, therefore, we are not looking for accuracy, but rather for a reliable estimate of the portion of CIF fluctuations that is due to oil price in order to justify the commodity price-oil price connection.

To illustrate the estimation method, we will consider the CIF price of bananas in the Grilli & Yang (1988) data set. The recorded price is CIF Gulf ports (central and south America) from the primary commodity data base. Since the port of origin was not mentioned in the commodity description, we will use Rio de Janeiro (Brazil); a frequent port of origin for bananas shipments going to central and south America. The relation that we are trying to estimate here is how much the bunker fuel cost (oil cost) represents out of the CIF price of bananas?

The distance between Rio de Janeiro and the Gulf ports is around 5097 nautical miles. It takes around 16 days to travel this distance with a medium-sized vessel cruising at an actual cruising speed, V, of 14 knots.[3] We were not sure about the nominal maximum cruising speed, V_N, for a medium-sized vessel, but it is definitely between 15 and 18 knots. So, we will choose the geometric mean of 16 knots. A medium-sized vessel burns around 40 tons of bunker fuel, F_N, per day for the main engines at nominal speed (Ronen (1982)). It is well known in ship engineering that the bunker fuel consumption of the main engines of a motor ship, F, is directly related to the third power of the speed (Manning (1956)); that is

$$F = \left(\frac{V}{V_N}\right)^3 F_N = \left(\frac{14\ knots}{16\ knots}\right)^3 \times 40\ tons \approx 27\ tons.$$

Therefore, the actual fuel consumption of a medium size vessel carrying bananas from Rio de Janeiro to Gulf ports is around 27 tons a day. It takes around 16 days (one way) to travel this distance, then the total fuel consumption of a round trip (two legs)[4] is

$$Total\ Fuel\ consumption = 2\ legs \times \frac{16\ days}{1\ leg} \times \frac{27\ tons}{1\ day} \approx 864\ tons.$$

Based on 2007 prices, the average oil price per barrel over the last three years ranges from a minimum of \$90 per barrel to a maximum of \$170 per barrel. We will take the geometric mean which is \$125 per barrel approximately. 1 barrel is equivalent to 0.17 cubic meters (cbm):

$$1\ barrel = 1\ barrel \times \frac{42\ gallons}{1\ barrel} \times \frac{0.004\ cbm}{1\ gallon} = 0.17\ cbm.$$

The cost of one cbm (1 ton) of bunker fuel is about

$$Fuel\ Cost\ per\ cbm\ (two\ legs) = 1\ cbm \times \frac{\frac{1}{0.17} barrel}{1\ cbm} \times \frac{\$125}{1\ barrel} \approx \$750.$$

[3] Source: www.searates.com

[4] In the returning trip, the vessel is usually empty yet the fuel cost is the same. Therefore, the carriers consider the two-leg bunker cost in their ocean freight calculations.

Therefore, the total bunker fuel cost of a round (two legs) trip is about

$$Total\ Fuel\ Cost\ (two\ legs) = 2\ legs \times \frac{\$750}{ton} \times \frac{864\ tons}{2\ legs} = \$648,000.$$

Next, we estimate the cost per container. We will consider a standard $40'$ refrigerated container, also known as 'refer.' The internal dimensions of a standard 40-feet container are 12.022 m of length, 2.352 m of width, and 2.395 m of height.[5] Therefore,

$$1 \times 40' container = 1 \times 40' container \times \frac{(12.022 \times 2.352 \times 2.395)\ cbm}{1 \times 40' container} \approx 67\ cbm.$$

Only 65% to 70% of the container's space is filled with actual cargo; the rest is devoted to crating, boxes, and other packaging materials. This leaves us with an actual 45 cbm of cargo in a standard $40'$ container.

The capacity of our medium-sized vessel ranges from a minimum of 350 twenty foot equivalent unit (TEU) to a maximum of 400 TEU. We will take the geometric mean which is about 370 TEU. Dividing this number by 2 yields around 185 container ($40^{'}$ each), which can fit around 8300 cbm of cargo

$$\frac{45\ cbm}{1 \times 40' container} \times 185\ container = 8300\ cbm.$$

Therefore, the bunker fuel cost per cbm, F_{cbm}, is

$$F_{cbm} = \frac{Total\ Fuel\ Cost}{Total\ cbms} = \frac{\$648,000}{8300\ cbm} \approx \$80\ per\ cbm.$$

To account for the frequent fluctuations in the oil price, all the shipping lines adjusts the ocean freight per cbm by adding a fuel surcharge fee, also known as the Bunker Adjustment Factor (BAF). For a $40^{'}$ container, the BAF is about \$680, or \$15 per one cbm of actual cargo. Thus, the total cost of fuel (including BAF) for one cbm of cargo is around \$95.

From the primary commodity price data base, over the last three years, the CIF price of one cbm of bananas shipped from Rio de Janeiro to Gulf ports ranged from \$650 to \$800 per metric ton. Taking the geometric mean, we estimate \$720 per metric ton. Recall that this price includes the insurance cost, customs clearance at the port of origin and port of discharge, handling, other fixed surcharges (see Figure 2), and, of course, the cost of the vessel itself (capital cost or return on investment). The cost of insurance and the other charges ranges from a maximum of 40% of the price and a minium of 30%. We will take the geometric mean which is 35%. Then, the stripped price of bananas is around \$470 ($\$720 \times 0.65$). This implies that the bunker fuel cost represents roughly around 20% ($\frac{\$95}{\$470} \approx 0.20$) of the stripped CIF price of bananas. This is indeed a significant proportion.

[5] Source: www.geocities.com.

Commodity	Bananas	Palm oil	Timber	Coffee	Cotton
Origin	Brazil	Malaysia	Malaysia	N/A†	Memphis, USA
Port of loading	Rio de Janeiro	Kelang	Kelang		Memphis
Destination	USA	Netherlands	UK	USA	Northern Europe
Port of discharge	Gulf Ports	Rotterdam	London	NYC	Liverpool, UK
Distance (nautical miles)	5097	8083	7930	2845	5144
Time (days)	16	24	23.5	8.5	15.3
Fuel cost/cbm	$95	$250	$130	$65	$90
Price/cbm	$470	$890	$546	$650	$660
% of fuel out of the price	20%	28%	24%	10%	14%

†The port of origin was not mentioned explicitly. The price is arithmetic average of shipments from El Salvador, Guatemala, and Mexico. All calculations are done based on the arithmetic means of the available data.

Table 2: The percentage of fuel out of the CIF price for all CIF commodities in the Grilli and Yang (1988) data set.

The previous analysis was performed for all CIF commodities in the Grilli and Yang (1988) data set and the results were recorded in Table 2. The only exception was the tobacco price series as the data on the origin country was not available. The analysis reveals that the fuel cost represents roughly 20% of the price of commodities recorded on a CIF basis.

The previous simple estimation method suggests that the price of oil is a potential transition variable that is capable of capturing the regime switching in a nonlinear CIF commodity price time series. It is also worth noting that, guided by the rationale behind border price classification, one should also consider macro economic news variable or variables in the country of origin when searching for potential transition candidates for CIF prices.

3 The Data

The data set used in this paper consists of 24 annual primary commodity prices developed originally by Grilli and Yang (1988) between 1900 and 1986 and then updated by Pfaffenzeller, Newbold, and Rayner (2007). We will use the updated version, which includes annual observations on the 24 commodities from 1900 till 2007.

Aiming, among other things, at analyzing the long-run movement in the net barter terms of trade series, Grilli and Yang (1988) used these prices to develop an index, which is known in the literature as the Grilli and Yang Commodity Price Index (GYCPI). The authors deflated the GYCPI by an index of manufactured goods' unit values (MUV) between 1900 and 1986. The MUV is a trade-weighted index of the five major developed countries' (France, Germany, Japan, United Kingdom, and United States) exports of manufactured commodities to developing countries. The ratio GYCPI/MUV, or the *real* GY-CPI, measures the purchasing power of the primary commodities in terms of traded manufactures.

Our focus is not to study the behavior of the index itself; rather we seek to study the individual 24 primary commodity prices forming it. In particular, we wish to classify these commodities into groups and, in each group, we seek to find the best transition variable that is capable of capturing the nonlinear dynamic of the *real* price series. Thus, for every commodity i, for $i = 1, ..., 24$, we seek to study the regime switching behavior of

$$y_{it} = \log\left(\frac{P_{it}}{MUV_t}\right),$$

where P_{it} is commodity ith price at time, and to find the best exogenous transition variable s_{it}.

4 A Framework for Classifying and Modelling Regime Switching in Border Prices

In this section, we propose a framework of modelling regime switching in commodity prices. The approach is based on the rationale of border price classification that was introduced in Section 2. We will, therefore, only focus on modelling the dynamic of nonlinear FOB and CIF commodities in the data set. The framework can be summarized in two stages: (1) classification, specification, and nonlinearity testing, and (2) estimation and evaluation. Before we proceed with the analysis of the two stages, we give a brief overview of the regime switching model that will be adopted.

The standard STR model of order p for an individual real commodity price y_{it}, $i = 1, ..., 24$, is expressed as follows.

$$y_{it} = \mathbf{\Phi}' z_{it} + \mathbf{\Theta}' z_{it} G(s_t; \gamma, \mathbf{c}) + \varepsilon_{it}, \qquad t = 1, 2, ..., T, \tag{1}$$

where $y_{it} = \log\left(\frac{P_{it}}{MUV_t}\right)$, $z_{it} = (1, y_{it-1}, ..., y_{it-p})' = (1, \tilde{z}_{it}')'$, $\mathbf{\Phi} = (\phi_0, \phi_1, ..., \phi_p)' = \left(\phi_0, \tilde{\mathbf{\Phi}}'\right)'$, $\mathbf{\Theta} = (\theta_0, \theta_1, ..., \theta_p)' = \left(\theta_0, \tilde{\mathbf{\Theta}}'\right)'$, $\varepsilon_{it} \sim i.i.d.(0, \sigma^2)$, and $G(s_{it}; \gamma, \mathbf{c})$ is a transition function. It is a bounded function (between 0 and 1) of the continuous transition variable s_{it} and continuous everywhere in the parameter space for any value of s_{it}. In the standard STR model, the transition variable s_{it} is an element of the lags of the dependent variable, i.e., $s_{it} = y_{it-d}$ or $s_{it} = \triangle y_{it-d}$, where $d > 0$ is a delay parameter, or a linear time trend. In this paper, however, we follow Fahmy's (2014) suggestion and allow the transition variable s_{it} to be either a lagged endogenous variable or an exogenous variable that is entailed by the border price classification.

The definition of $G(s_{it}; \gamma, \mathbf{c})$ is the one governing the empirical applicability of (1). The function $G(\cdot)$ is a *logistic function* defined in general as

$$G(s_{it};\gamma,\mathbf{c}) = \left(1+\exp\{-\gamma \prod_{j=1}^{k}(s_{it}-c_j)\}\right)^{-1}, \qquad \gamma > 0, \tag{2}$$

where γ is the slope of the function, and $\mathbf{c} = (c_1, ..., c_k)'$ is a vector of location parameters such that $c_1 \leq \cdots \leq c_k$. Given such definition, the STR model defined in (1) is then referred to as the logistic smooth transition regression (LSTR) model.

The transition function in (2) is a monotonically increasing function of the transition variable s_t. The restrictions $\gamma > 0$ and $c_1 \leq \cdots \leq c_k$ are identifying restrictions. The choice of k is not only crucial in determining the behavior of the logistic transition function, but also has a significant implication in interpreting the time series under consideration. Two common choices for k are used in the literature: $k = 1$ and $k = 2$. In the LSTR model with $k = 1$ (LSTR1), the parameters vectors change monotonically as a function of s_{it} from $\mathbf{\Phi}$ to $\mathbf{\Phi} + \mathbf{\Theta}$. This implies that the LSTR1 model is capable of characterizing asymmetric time series behavior, i.e., processes whose dynamic properties are different in an upper regime from what they are in a lower regime such that the transition between the two regimes is smooth. The logistic function in the LSTR1 model takes the form

$$G(s_{it};\gamma,c) = (1+\exp\{-\gamma(s_{it}-c)\})^{-1}, \qquad \gamma > 0. \tag{3}$$

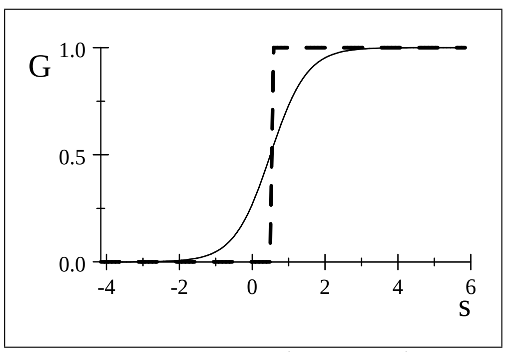

Figure 3: The smooth transition logistic function of order 1 with a moderate slope $\gamma = 2$ (the solid line) and with an extremely larger slope $\gamma = 1000$ (the dashed line). The threshold value $c = 0.5$.

Note that the logistic function of order 1 is bounded between 0 and 1; that is, when $s_{it} \longrightarrow -\infty$, $G(\cdot) = 0$; this defines the lower regime. On the other hand, when $s_{it} \longrightarrow +\infty$, $G(\cdot) = 1$ and the time series is said to be in an upper regime. The first-order logistic function in (3) is plotted in Figure 3, where the threshold $c = 0.5$ and the slope $\gamma = \{2, 1000\}$ for the solid and the dashed lines respectively.

In the LSTR model with $k = 2$ (LSTR2), the parameter vectors change symmetrically around the midpoint $\bar{c} = \frac{c_1+c_2}{2}$, where the logistic function attains its minimum value. The LSTR2 model is a three-regime switching regression model in which the dynamics of the two outer regimes, associated with large and small values of s_t, are the same while the behavior in the transition period (the middle regime) is different. The second-order logistic function in the LSTR2 model takes the form

$$G(s_{it}; \gamma, c_1, c_2) = (1 + \exp\{-\gamma(s_{it} - c_1)(s_{it} - c_2)\})^{-1}, \qquad \gamma > 0,\ c_1 \leq c_2, \tag{4}$$

where $\gamma > 0$ and $c_1 \leq c_2$ are identifying restrictions. The function achieves its minimum value $G_{\min} = (1+\exp\{-\gamma\tilde{c}\}/\sigma^2)^{-1}$, where $\tilde{c} = c_1c_2 - \bar{c}^2$ and $\bar{c} = \frac{c_1+c_2}{2}$, when the transition variable s_t is equal to $\bar{c}$. The LSTR2 function in (4) is plotted in Figure 4, where $c_1 = -1, c_2 = 2$ and $\gamma = \{2, 1000\}$ for the solid and dashed lines respectively. Observe how the functions change symmetrically around $\bar{c} = 1/2$ (the midpoint). When $\gamma = 0$, the transition function $G(s_{it}; \gamma, c_1, c_2) = 1/2$, and the STR model in (1) nests the linear model. When $\gamma \to \infty$ in the LSTR1 model, the result approaches the pure threshold model; when $\gamma \to \infty$ in the LSTR2 model, the result is a pure threshold model with three regimes such that the two outer regimes are similar while the ground regime (the middle regime) is different and the transition between regimes is swift (see the dashed line in Figure 4).

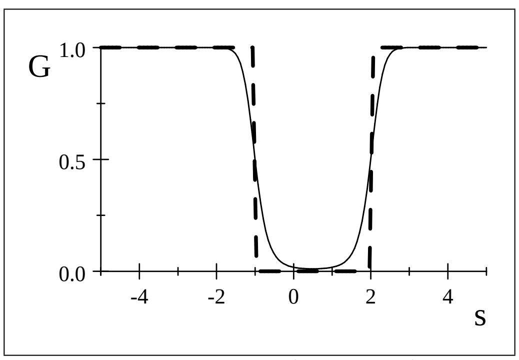

Figure 4: The smooth transition logistic function of order 2 with a moderate slope $\gamma = 2$ (the solid line) and with an extremely larger slope $\gamma = 1000$ (the dashed line). The threshold values are $c_1 = -1$ and $c_2 = 2$.

4.1 Stage 1: Border Classification, Specification, and Nonlinearity Testing

Border price classification and nonlinearity testing take place in the first stage; the *specification, classification, and non-linearity testing stage*. In this stage, we begin by specifying, for each individual real commodity price time series y_{it}, an adequate linear autoregressive model of order p, $AR(p)$, where p is the value that minimizes the Akaike (1974) information criterion (AIC) for every y_{it}. The resulting autoregressive model acts as the starting model for nonlinearity analysis. Before accepting the suggested $AR(p)$ model as the starting point of the analysis, preliminary diagnostic tests should be applied to the model in order to ensure its adequacy as a starting linear model. In particular, the Ljung-Box (1978) test of no serial correlation of order $q = 1$ up to $q = 8$ in the residuals and Engle's (1982) Lagrange multiplier (LM) test of no autoregressive conditional heteroskedasticity (ARCH) of order $v = 1$ up to $v = 4$ in the residuals are considered. The p-values of both tests are denoted by $Q(q)$ and $LM_{ARCH(v)}$ respectively. If the model fails these preliminary tests, especially the Engle's ARCH test, then we conclude that there exists nonlinearity in the conditional variance of the price series. (G)ARCH model is a suitable model for capturing the dynamic of these processes. On the other hand, if the model passes these preliminary tests, we move to nonlinearity testing.

The value of the lag order of the AIC and the p-values of the previous diagnostic tests are reported in Table 3. Judging by the Ljung-Box (1978) statistics, $Q(q)$, the null hypothesis of no serial correlation of order $q = 1$ up to $q = 8$ in the residuals series for all the 24 commodities is not rejected at the 5% level of significance. The null-hypothesis of normality of errors was rejected at the 5% level of significance for the majority of commodities as seen from the p-value of the Jarque and Bera (1980), JB, test statistic. This rejection is, however, due to the presence of outliers in the time series. Finally, the null hypotheses of no $ARCH(v)$, $v = 1, ..., 4$, were also not rejected at the 5% level of significance for the majority of the 24 commodities; notable exceptions are tobacco, silver, jute, lead, cotton, wool, aluminum, and tea. This is not surprising as the majority of these prices (with the exception of wheat and beef) are settlement or auction prices of commodities traded in exchanges and, therefore, tend to exhibit volatility clusters; a common feature of stock and option prices. Therefore, (G)ARCH or ST-(G)ARCH models are suitable models for this type of commodities. It is worth mentioning that commodities in this group are all storable commodities. This is consistent with Muth's (1961) hypothesis and with the results obtained by Beck (2001), who applied a variation of (G)ARCH techniques to commodity prices and found an ARCH process in storable but not in non-storable commodity data.

Since our focus is to investigate the nonlinearity in border prices (FOB and CIF prices), the previous 8 individual price series will be dropped from the analysis. The remaining 16 commodities, that passed the preliminary diagnostics, will be subjected to nonlinearity testing.

For the remaining 16 commodities, we proceed by fitting adequate linear $AR(p)$ for every y_{it}, where the order p is selected as suggested by the first column of Table 3. The resulting autoregressive models act as the starting model for nonlinearity analysis. Next, we test the existence of nonlinearity in the 16 commodities. If nonlinearity is present, we move to the *second stage* in the modelling framework, which is the *estimation and evaluation* stage. On the other hand, if nonlinearity is not present, then the time series is linear and doesn't display any regime switching.

The null hypothesis of linearity, denoted by H_{0L} in the text, is tested against the alternative of a nonlinear STR model. The test statistic is due to Teräsvirta (1994) and Luukkonen et al (1988), which is basically an LM test with an asymp-

y_{it}	$AIC(p)$	Residuals Analysis (AR Model) JB	K_3	K_4	$Q(1)$	$Q(8)$	$LM_{ARCH(1)}$	$LM_{ARCH(4)}$
Tobacco	$p=5$	†0.01	0.42	3.84	0.90	0.84	0.29	†0.04
Silver	$p=3$	†0.000	0.66	4.50	0.94	0.82	†6.2×10^{-5}	†0.003
Jute	$p=3$	0.13	-0.07	3.71	0.75	0.70	0.59	†0.03
Lead	$p=1$	†0.01	0.24	4.05	0.27	0.79	0.08	†0.01
Cotton	$p=4$	0.84	-0.14	2.84	0.87	0.89	†0.002	†0.04
Wool	$p=5$	0.42	0.22	3.23	0.92	0.99	†0.04	0.32
Aluminum	$p=3$	†0.000	0.62	6.11	0.98	0.91	†0.02	†0.02
Tea	$p=3$	†0.008	0.16	4.16	0.93	0.56	†6.4×10^{-4}	†0.02
Wheat	$p=9$	†0.003	0.42	4.10	0.99	0.99	0.83	0.89
Lamb	$p=5$	†0.000	0.21	4.89	0.69	0.96	0.37	0.87
Coffee	$p=1$	†0.02	0.51	3.55	0.52	0.80	0.33	0.78
Copper	$p=3$	0.62	0.19	3.05	0.77	0.68	0.58	0.14
Cocoa	$p=3$	†0.004	0.65	3.59	0.92	0.99	0.93	0.05
Timber	$p=1$	0.50	0.10	3.31	0.60	0.29	0.39	0.11
Tin	$p=3$	†0.01	-0.33	3.95	0.86	0.97	0.87	0.52
Zinc	$p=2$	†0.000	1.64	10.2	0.99	0.96	0.63	0.98
Maize	$p=5$	†0.000	-0.38	4.70	0.75	0.95	0.32	0.56
Beef	$p=1$	†0.000	0.57	5.77	0.80	0.72	0.71	0.78
Rice	$p=5$	†0.01	-0.07	4.18	0.84	0.86	0.41	0.39
Bananas	$p=3$	0.08	-0.36	3.52	0.95	0.97	0.76	0.21
Palm oil	$p=3$	†0.000	-0.54	4.25	0.99	0.63	0.12	0.44
Rubber	$p=1$	†0.000	0.75	5.28	0.39	0.47	0.63	0.10
Hides	$p=3$	0.20	-0.33	3.30	0.76	0.64	0.35	0.66
Sugar	$p=3$	†0.000	0.67	5.35	0.93	0.30	0.45	0.89

†The null-hypothesis is rejected at the 5% level of significance.

Table 3: The lag order of the AIC and the p-values of the diagnostic tests of the linear AR models' residuals applied to the 24 commodities in the Grilli and Yang (1988) data set. K3 and K4 are skewness and kurtosis respectively.

totic F distribution when H_{0L} is valid. The problem with nonlinearity testing is that the nonlinear STR model under consideration is only identified under the alternative hypothesis. Luukkonen, Saikkonen, and Teräsvirta (1988), based on a paper by Davies (1977), suggested a solution to this problem. Their solution, which was adopted by Teräsvirta (1994), was simply to replace the transition function (2) in (1) by a Taylor approximation about the null hypothesis $\gamma = 0$. In particular they assumed a first order logistic function, i.e., $k = 1$ in (2), and performed a third order Taylor approximation about the null hypothesis $\gamma = 0$. The approximation yields, for every commodity i,

$$y_{it} = \mathbf{\Phi}^{'} z_{it} + \frac{1}{4}\gamma \mathbf{\Theta}^{'} z_{it}(s_{it} - c) - \frac{1}{48}\gamma^3 \mathbf{\Theta}^{'} z_{it}(s_{it} - c)^3 + \varepsilon_{it}. \tag{5}$$

Using $z_{it} = (1, \tilde{z}_{it}^{'})^{'}, \mathbf{\Phi} = \left(\phi_0, \tilde{\mathbf{\Phi}}^{'}\right)^{'}$, $\mathbf{\Theta} = \left(\theta_0, \tilde{\mathbf{\Theta}}^{'}\right)^{'}$, and reparameterizing, equation (5) can be expressed as

$$y_{it} = \delta_0 + \boldsymbol{\delta}_1^{'} \tilde{z}_{it} + \boldsymbol{\pi}_1^{'} \tilde{z}_{it} s_{it} + \boldsymbol{\pi}_2^{'} \tilde{z}_{it} s_{it}^2 + \boldsymbol{\pi}_3^{'} \tilde{z}_{it} s_{it}^3 + \varepsilon_{it}^*, \tag{6}$$

where $\varepsilon_{it}^* = \varepsilon_{it} + R(\gamma, c, s_{it})$, $R(\cdot)$ being the remainder, and $\boldsymbol{\pi}_j$, $j = 1, 2, 3$, is of the form $\gamma \tilde{\boldsymbol{\pi}}_j$, where $\tilde{\boldsymbol{\pi}}_j \neq 0$ is a function of $\tilde{\mathbf{\Theta}}$. The null hypothesis of linearity is then $H_{0L} : \boldsymbol{\pi}_1 = \boldsymbol{\pi}_2 = \boldsymbol{\pi}_3 = 0$. Also note that because $\varepsilon_{it}^* = \varepsilon_{it}$ under the null hypothesis, the asymptotic theory will not be affected if an LM test is used. Following Luukkonen et al. (1988) and Teräsvirta (1994), a convenient procedure for computing the LM statistic by OLS is to estimate (6) under the null hypothesis and compute the sum of squares of the residuals (SSR_0), then estimate (6) and compute SSR_1. The LM statistic is computed as $\frac{T(SSR_0 - SSR_1)}{SSR_1}$. The test statistic has an asymptotic chi-square distribution with $3p$ degrees of freedom when the null hypothesis is valid. However, the F statistic is recommended because the chi-square statistic can be size-distorted in small and even moderate samples. In this paper, we shall use the F distribution with $3p$ and $T - 4p - 1$ when the null hypothesis H_{0L} is valid. The test is repeated for each transition candidate in the transition set. If the null hypothesis of linearity, H_{0L}, using the F test (F_L) is rejected for at least one of the models, the model against which the rejection, measured in the p-value, is strongest is chosen to be the STR model to be estimated.

Another purpose of conducting the linearity test is to use the test results for model selection. If linearity is rejected and a transition variable is selected, the next step is to choose a model type, i.e., to choose between LSTR1 or LSTR2 models. The choice between the two models can be based, again, on the auxiliary regression (6). Teräsvirta (1994) showed that when $c = 0$ then $\boldsymbol{\pi}_2 = 0$ when the model is an LSTR1, whereas $\boldsymbol{\pi}_1 = \boldsymbol{\pi}_3 = 0$ when the model is an LSTR2. The following F tests sequence[6] was then suggested based on the auxiliary regression in (6):

[6]For more details on the selection criterion for STR models, see Teräsvirta (1994) and the references therein.

1. Test the null hypothesis: $H_{04} : \boldsymbol{\pi}_3 = 0$ with an ordinary F test (F_4). A rejection of H_{04} can be interpreted as a rejection of the LSTR2.

2. Test the null hypothesis that $\boldsymbol{\pi}_2 = 0$ given that $\boldsymbol{\pi}_3 = 0$, $H_{03} : \boldsymbol{\pi}_2 = 0|\boldsymbol{\pi}_3 = 0$, using another F test (F_3). Failure to reject H_{03} indicates that the model is an LSTR1.

3. The last F test (F_2) in the sequence is to test the null hypothesis that $\boldsymbol{\pi}_1 = 0$ given that $\boldsymbol{\pi}_2 = \boldsymbol{\pi}_3 = 0$ as $H_{02} : \boldsymbol{\pi}_1 = 0|\boldsymbol{\pi}_2 = \boldsymbol{\pi}_3 = 0$. Rejecting H_{02} after accepting H_{03} supports the choice of the LSTR1 model. Accepting H_{02} after rejecting H_{03} points to the LSTR2 model.

After carrying out the three F tests and noting which hypotheses are rejected, if the test H_{03} yields the strongest rejection measured in the p-value, choose the LSTR2 model; otherwise select the LSTR1 model.

The previous nonlinearity tests and the model selection procedure were applied to the remaining 16 commodities in the data set. In the model selection procedure, we followed the rationale of border price classification and used exogenous transition variables (in addition to the autoregressive lags of the time series) to choose the best LSTR model. In particular, the transition set for the remaining 16 commodities, Ω, consists of the autoregressive lags of the dependent variable, the current and one period lag inflation rates, Π_t and Π_{t-1} respectively, and the current and one period lag of the growth rate of the real price of oil defined as the first difference of the logarithm of real oil and its one period lag respectively; that is, $r_t = \triangle \log \left(\frac{Oil\ Price}{CPI}\right)_t$ and r_{t-1}.

$$\Omega = \{y_{t-1}, y_{t-2}, ..., y_{t-p}, \Pi_t, \Pi_{t-1}, r_t, r_{t-1}\}.$$

Applying the nonlinearity tests sequence above to each transition candidate in Ω, linearity was not rejected for 8 commodities. In particular, linearity was not rejected for beef, cocoa, lamb, wheat, tin, copper, zinc, and rubber as shown in the top portion of Table 4. The behavior of these commodities is best described by a linear model, and therefore, they were all dropped from our analysis. It is worth noting that in the case of zinc, the transition variable that showed the highest rejection of linearity tests was the current growth rate of oil, r_t, and the associated suggested model was the LSTR1. This, however, is inapplicable in this case since the zinc price is a settlement price, and consequently, oil doesn't play any role in its fluctuations. Another noticeable exception is the case of rubber. The one period lag U.S. inflation rate showed the highest rejection of linearity tests and the associated model was the LSTR1. But, again, the price of rubber is a spot price, and consequently, U.S. inflation should not be used as transition variable (even if the nonlinearity tests suggest so). One should exercise caution when using exogenous transition variables for the statistical procedure does not provide intuition or rationale for the choice of the right variable. For that reason, the proposed border price classification is useful in suggesting transition candidates that make sense.

Nonlinearity was confirmed for the remaining 8 commodity prices; namely, maize, rice, sugar, hides (FOB prices), and bananas, palm oil, timber, and coffee

(CIF prices). As shown from Table 4, the nonlinearity and model selection tests results confirm the proposed border price classification; inflation in the exporting country was the best transition variable that showed the highest rejection of linearity tests in FOB prices whereas oil price showed the highest rejection in CIF prices. The selected model (either LSTR1 or LSTR2) in each group is the one tagged with '$*$.' A noticeable exception in the FOB group is the price of rice, where the transition variable that showed the highest rejection of linearity was the fifth order autoregressive lag, y_{t-5}. This is not surprising, however, since the border price of rice is FOB Thailand (see Table 1), and thus, according to our border price classification, inflation or any macro economic news variable in Thailand should be used as transition variable in this case. However, due to unavailability of data, an autoregressive transition variable was used instead. The nonlinearity tests were executed for all 5 lags and the one that showed the highest rejection was the fifth and the corresponding model was the LSTR2 as shown in Table 4. Notice here that U.S. inflation or the price of oil are not applicable in this case.[7]

4.2 Stages 2: Estimation and Evaluation

After determining the transition variable and the type of the STR model, the next step is estimation. The parameters of the STR model in (1) are estimated using conditional maximum likelihood. Assuming normality of the error term, the log-likelihood function of the STR model, for every commodity i, is

$$l(\mathbf{\Phi}, \mathbf{\Theta}, \sigma, \gamma, \mathbf{c}) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma^2) - \frac{1}{2}\frac{\{y_{it} - (\mathbf{\Phi}^{'} + \mathbf{\Theta}^{'} G(z_{it}; \gamma, \mathbf{c}))z_{it}\}^2}{\sigma^2}. \tag{7}$$

Conditional on starting values of the parameters, the log-likelihood function in (7) is maximized using the iterative Broyden-Fletcher-Goldfarb-shanno (BFGS) algorithm. Finding good starting values is important for the algorithm to work properly.

Starting values are obtained by constructing a grid in γ and $\mathbf{c}$, estimating the parameter vectors $\mathbf{\Phi}$ and $\mathbf{\Theta}$ conditionally on (γ, c) for $k = 1$ (LSTR1 model) or (γ, c_1, c_2) for $k = 2$ (LSTR2 model), and computing the sum of squared residuals. The parameter values that correspond to the minimum of that sum are taken as the starting values.[8]

The final step in this stage is evaluation, where the adequacy of the fitted model is considered. The misspecification tests for the STR models that have been considered in Eitrheim and Teräsvirta (1996) and Teräsvirta (1998) will

[7] For the commodities in the FOB group, in addition to inflation, we have also experimented with other macro economic news variables in the exporting country (USA) such as unemployment, money supply measures, and intereste rate. Inflation rate was the variable that showed the highest rejection of linearity tests.

[8] To facilitate the construction of an effective grid, weI follow Teräsvirta's (1998) suggestion to standardize the exponent of the transition function $G(s_t; \gamma, \mathbf{c})$ by dividing it by the k^{th} power of the sample standard deviation of the transition variable $\widehat{\sigma}_s^k$. This is done mainly to render the parameter γ scale-free.

P_{it}	$AIC(p)$	Border Price	The Transition Variable and the Suggested Model $y_{t-j};\ j \in p$	Π_t	Π_{t-1}	r_t	r_{t-1}
Linear commodity prices							
Beef	$p=1$	FOB & CIF	Linear	Linear	Linear	Linear	Linear
Cocoa	$p=3$	Option price	Linear	Linear	Linear	Linear	Linear
Lamb	$p=5$	CIF	Linear	Linear	Linear	Linear	Linear
Wheat	$p=9$	FOB	Linear	Linear	Linear	Linear	Linear
Tin	$p=3$	Settlement	Linear	Linear	Linear	Linear	Linear
Copper	$p=3$	Settlement	Linear	Linear	Linear	Linear	Linear
Zinc	$p=2$	Settlement	Linear	Linear	Linear	LSTR1†	Linear
Rubber	$p=1$	Spot price	Linear	LSTR2	LSTR1‡	Linear	Linear
Nonlinear FOB Group: STR Models with inflation rate as transition variable							
Maize	$p=5$	FOB	y_{t-2}; LSTR2	LSTR1*	Linear	NA	NA
Hides	$p=3$	FOB	Linear	LSTR2	LSTR1*	NA	NA
Rice	$p=5$	FOB	y_{t-5}; LSTR2*	NA	NA	NA	NA
Sugar	$p=3$	FOB	y_{t-1}; LSTR1	LSTR1*	LSTR2	NA	NA
Nonlinear CIF Group: STR Models with oil price as transition variable							
Bananas	$p=3$	CIF	Linear	Linear	Linear	Linear	LSTR1*
Palm oil	$p=3$	CIF	y_{t-3}; LSTR2	LSTR1	Linear	LSTR2*	LSTR1
Timber	$p=1$	CIF	Linear	LSTR1	Linear	LSTR1*	Linear
Coffee	$p=1$	CIF	Linear	LSTR2	Linear	Linear	LSTR2*

NA≡Not Applicable.

r_t and $r_{t-1}$are current and one period lag of growth rate in real oil prices respectively.

Π_t and $\Pi_{t-1}$are current and one period lag inflation respectively.

‡Transition variable exhibits the highest rejection of linearity tests, but the coefficients of the fitted model were insignificant.

†Although this variable showed, statistically speaking, the highest rejection of linearity tests, it is not applicable in this case.

**Both models were close; but the one tagged with two stars outperforms the one star model.

Table 4: Nonlinearity and model selection tests for Settlement, FOB, and CIF prices.

be considered in this paper. In particular, three tests will be considered. The first test is an LM-type test of no error autocorrelation of order q. The p-value of the test is denoted by $LM_{AUTO(q)}$. This test can be viewed as a special case of a general test that was first suggested by Godfrey (1988). Note that the usual Ljung-Box (1978) test of no serial correlation is inapplicable here because the asymptotic distribution of the test statistic is unknown when the residuals from the STR model are used.[9] The second diagnostic test is an LM-type test of no remaining nonlinearity in the fitted STR model. The p-value of this test is denoted by $NRNL$. The test is carried on against the transition variable selected in the specification stage. The last diagnostic test is a parameter constancy, where the null hypothesis of parameter constancy in the STR model is tested against non-monotonic change, non-monotonic symmetrical change, and non-monotonic and non-symmetrical change with p-values denoted by $PC1$, $PC2$, and $PC3$ respectively.

The best fitted models that we managed to obtain for the FOB group after dropping the insignificant parameters and the results of Eitrheim and Teräsvirta (1996) and Teräsvirta's (1988) misspecification tests are reported in Table 5. The figures in parentheses beneath the parameter values are standard deviations and the figures recorded for the misspecification tests are p-values. All parameters are significant, and judged by the p-values of the misspecification tests, the models are adequate. Perhaps the most noticeable detail of Table 5 is the large standard deviation of the estimated slope of the logistic function γ in the case of hides, maize, and sugar. It is common for LSTR models that the estimated standard deviation of γ tends to be large for large values of γ. This is not crucial, however, as it will not affect either the shape of the logistic function $G(\cdot)$ or the other estimates of the model. The large slope indicates a swift transition from one regime to another as shown from the transition functions in Figures 5, 6, and 8 respectively, or from their dot plots in Figure 9. The slope in the case of rice was small, which indicates a smooth transition as shown from Figure 7 or the dot plot in Figure 9. Other than the large slope of the transition function, all models passed the misspecification tests. One exception is the rejection of the null-hypothesis of normality of errors at the 5% level of significance as seen from the p-value of the JB test statistic for maize, rice, and sugar. But, this is due to the existence of outliers in the time series. A quick look at the standardized residuals time series plotted in Figure 10, one can notice the very large outliers in 1947 for maize, the spike in 1973 for rice, and the spike in the late 1950s and beginning of 1960s for sugar, where the absolute value of the standardized residuals is greater than three. These outliers correspond to the post WWII, the post oil shock of 1973, and the high swings in the U.S. sugar price during the 1960s respectively.

[9] For more details, see Eitrheim and Teräsvirta (1996).

FOB prices: LSTR models with inflation as exogenous threshold variable†

$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + \phi_5 y_{t-5};$

$+(\theta_0 + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \theta_3 y_{t-3} + \theta_4 y_{t-4} + \theta_5 y_{t-5}) G(\gamma, c_1, c_2; s_t) + \varepsilon_t$

$G(\cdot) = (1 + \exp\{-\gamma(s_t - c_1)\}/\sigma_s)^{-1}$ or $G(\cdot) = (1 + \exp\{-\gamma(s_t - c_1)(s_t - c_2)\}/\sigma_s^2)^{-1}$

Commodity Price	Hides	Maize	Rice	Sugar
$AIC(p)$	$p = 3$	$p = 5$	$p = 5$	$p = 3$
s_t; Model	Π_{t-1}; LSTR1	Π_t; LSTR1	y_{t-5}; LSTR2	Π_t; LSTR1
ϕ_0		1.21 (0.31)		
ϕ_1	0.65 (0.08)		1.06 (0.09)	0.89 (0.09)
ϕ_2			−0.30 (0.09)	−0.32 (0.12)
ϕ_3	0.25 (0.08)	−1.24 (0.42)		0.22 (0.09)
ϕ_4				
ϕ_5			0.35 (0.08)	
θ_0	−0.20 (0.09)	−1.19 (0.32)	−0.24 (0.09)	0.42 (0.11)
θ_1		0.83 (0.07)		
θ_2				
θ_3		1.38 (0.43)		
θ_4			−0.5 (0.15)	
θ_5				
γ	30.6 (132)	16.7 (22.5)	2.42 (1.15)	32 (40)
c_1	0.10 (0.01)	−0.02 (0.003)	−0.62 (0.17)	0.10 (0.002)
c_2			0.73 (0.05)	
$\widehat{\sigma}_s$	0.047	0.047	0.45	0.047
$\overline{R}^2$	0.65	0.84	0.91	0.70
$LM_{AUTO(1)}; LM_{AUTO(8)}$	0.40; 0.65	0.43; 0.59	0.27; 0.36	0.29; 0.30
$LM_{ARCH(1)}; LM_{ARCH(4)}$	0.62; 0.45	0.09; 0.49	0.58; 0.38	0.63; 0.52
$NRNL$	0.31	0.28	0.97	0.51
$PC(1); PC(2); PC(3)$	0.34; 0.19; 0.3	0.3; 0.21; 0.15	0.38; 0.41; 0.48	0.05; 0.6; 0.5
$JB; K_3; K_4$	0.15; −0.23; 3.8	0.002; −0.39; 5.01	0.001; −0.25; 4.7	0.02; 0.2; 4.2

†$\widehat{\sigma}_s$ is the sample standard deviation of the transition variable s_t

Table 5: Estimation and evaluation tests results for FOB prices. Figures in parentheses beneath the parameters are standard deviations. Figures pertaining to the evaluation tests are p-values.

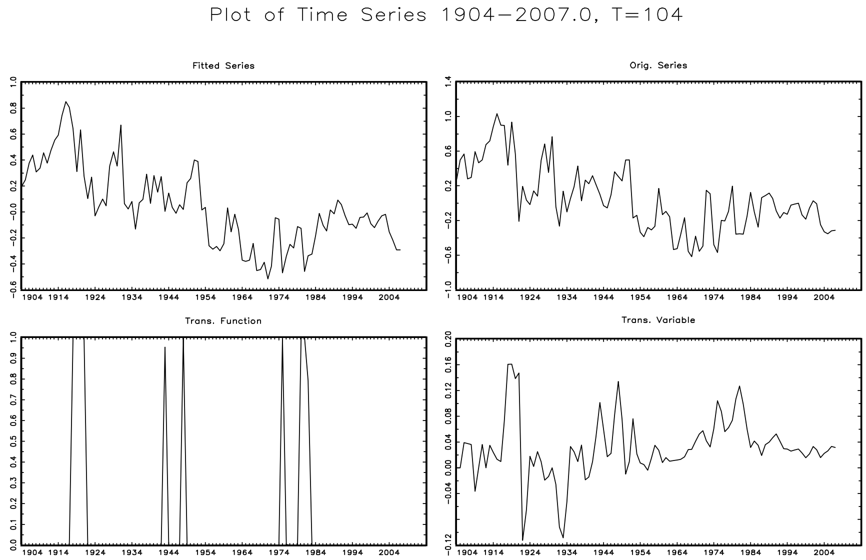

Figure 5: The figures, from left to right row-wise, show the fitted series $\widehat{y}$, the original series y_t, the transition function, and the transition variable Π_{t-1} for hides.

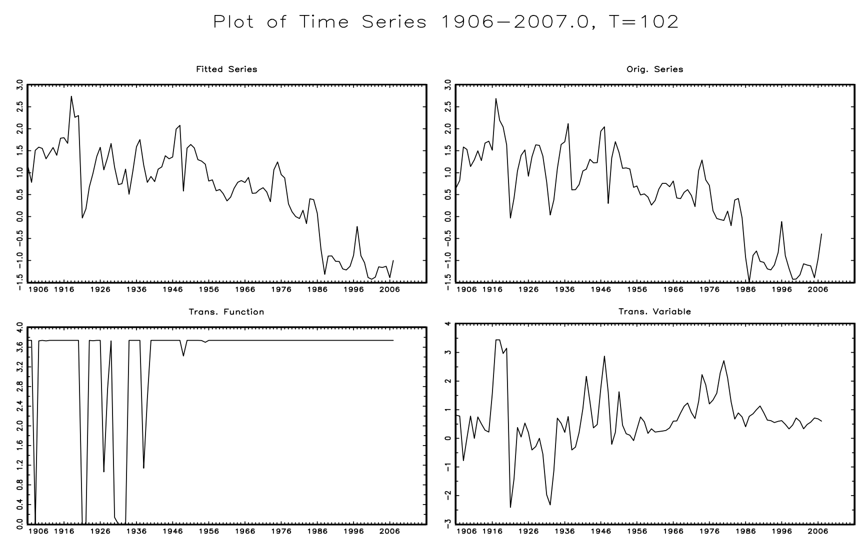

Figure 6: The figures, from left to right row-wise, show the fitted series $\widehat{y}$, the original series y_t, the transition function, and the transition variable Π_t for maize.

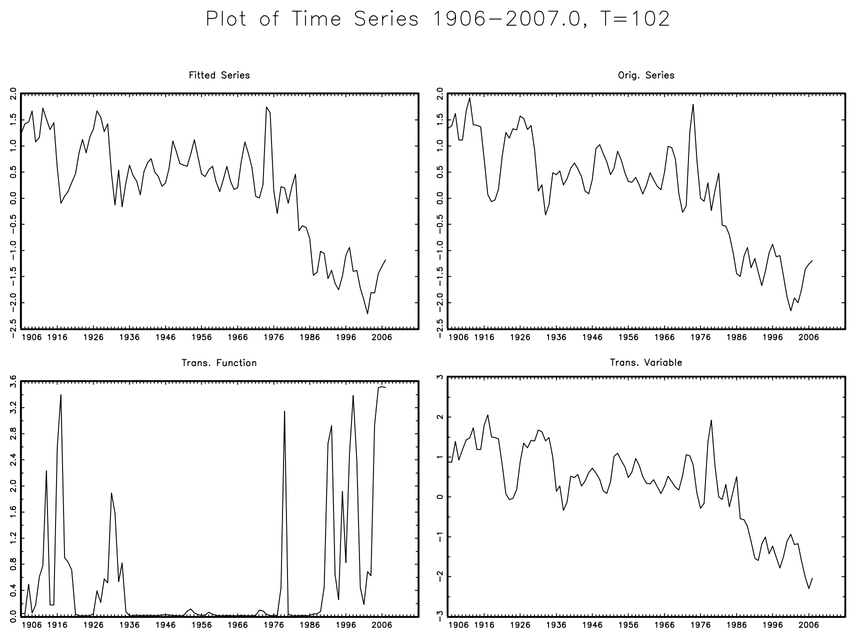

Figure 7: The figures, from left to right row-wise, show the fitted series $\widehat{y}$, the original series y_t, the transition function, and the transition variable y_{t-5} for rice.

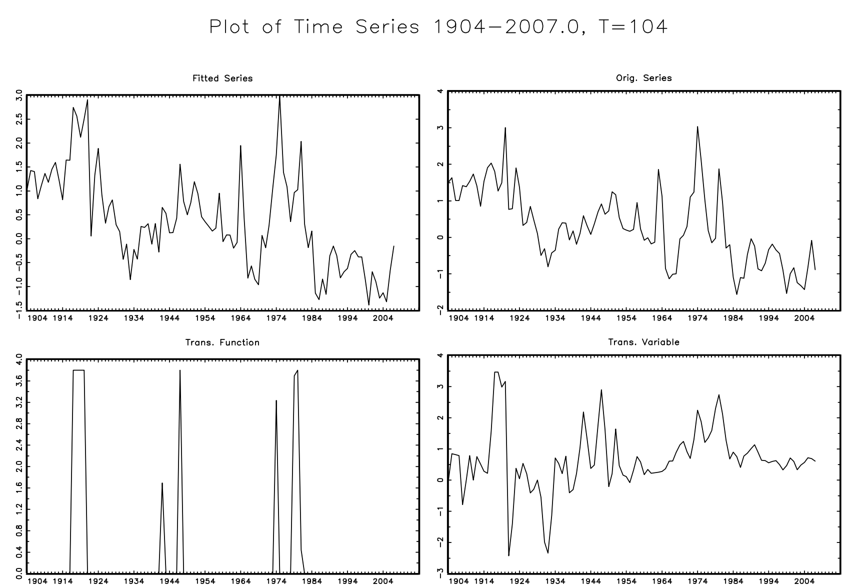

Figure 8: The figures, from left to right row-wise, show the fitted series $\widehat{y}$, the original series y_t, the transition function, and the transition variable H_t for sugar.

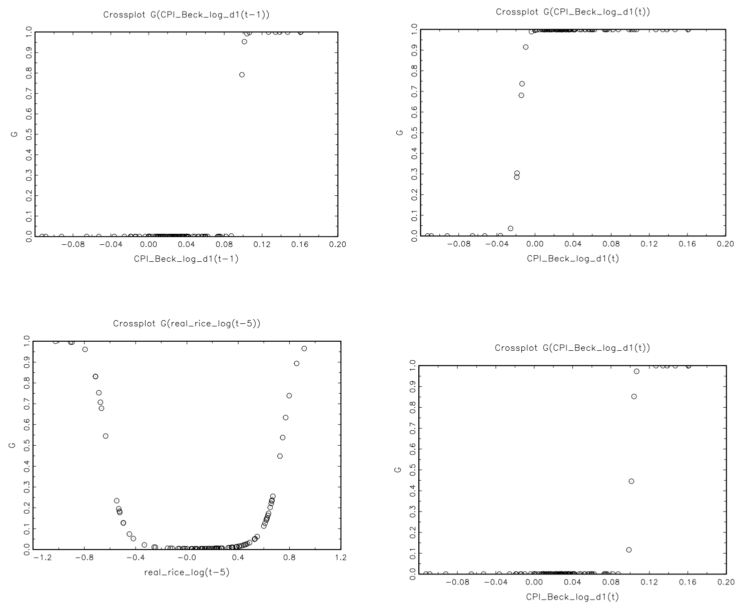

Figure 9: Transition functions of hides (top left), maize (top right), rice (bottom left), and sugar (bottom right). Each dot corresponds to one observation.

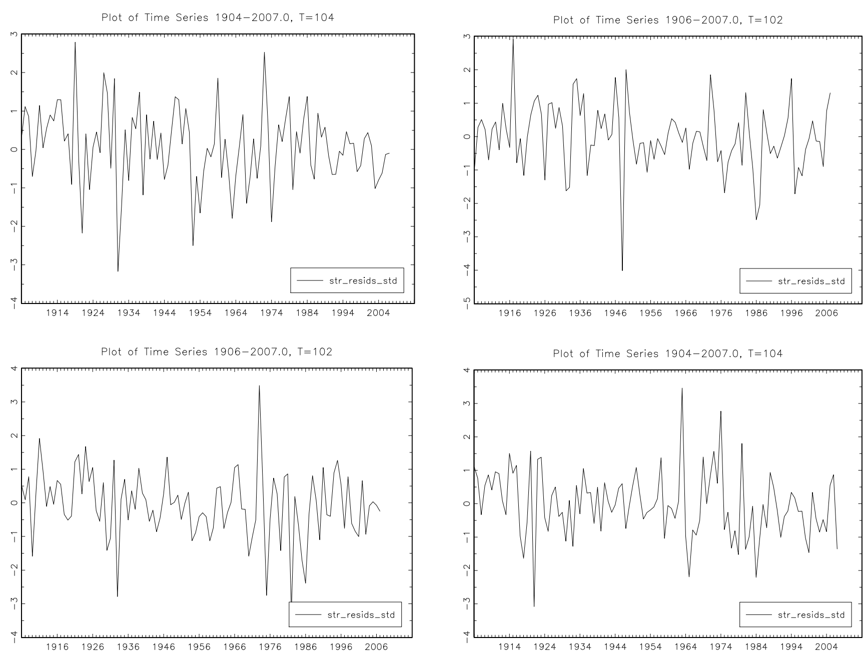

Figure 10: Standardized residuals of hides (top left), maize (top right), rice (bottom left), and sugar (bottom right).

CIF Prices: LSTR models with oil as exogenous threshold variable $y_t = \phi_0+\phi_1 y_{t-1}+\phi_2 y_{t-2}+\phi_3 y_{t-3}$ $+(\theta_0+\theta_1 y_{t-1}+\theta_2 y_{t-2}+\theta_3 y_{t-3})G(\gamma, c_1, c_2; s_t)+\varepsilon_t$, where $G(\cdot)=(1+\exp\{-\gamma(s_t-c_1)\}/\sigma_s)^{-1}$ or $G(\cdot)=(1+\exp\{-\gamma(s_t-c_1)(s_t-c_2)\}/\sigma_s^2)^{-1}$.				
Commodity Price	Bananas	Palm oil	Timber	Coffee
$AIC(p)$	$p=3$	$p=3$	$p=1$	$p=1$
s_t; Model	r_{t-1}; LSTR1	r_t; LSTR2	r_t; LSTR1	r_{t-1}: LSTR2
ϕ_0	$\phi_0 = 0$	$\phi_0 = 0$	-0.23 (0.06)	-0.11 (0.06)
ϕ_1	0.96 (0.09)	0.77 (0.05)	0.93 (0.03)	0.83 (0.05)
ϕ_2	-0.31 (0.13)	$\phi_2 = 0$		
ϕ_3	0.28 (0.09)	$\phi_3 = 0$		
θ_0	0.08 (0.03)	0.08 (0.04)	0.22 (0.06)	-0.42 (0.19)
θ_1	$\theta_1 = 0$	0.61 (0.13)	$\theta_1 = 0$	$\theta_1 = 0$
θ_2	$\theta_2 = 0$	-1.18 (0.22)		
θ_3	$\theta_3 = 0$	0.91 (0.18)		
γ	9.8 (18.43)	14.7 (21.48)	251 (1420)	0.5 (0.55)
c_1	0.14 (0.06)	-0.64 (0.20)	-0.37 (0.32)	-0.71 (0.23)
c_2		0.07 (0.01)		0.56 (0.23)
$\widehat{\sigma}_s$†	0.20	0.20	0.20	0.20
$\overline{R}^2$	0.80	0.86	0.88	0.73
$LM_{AUTO(1)}$;$LM_{AUTO(8)}$	0.39; 0.42	0.24; 0.62	‡NA	0.38; 0.82
$LM_{ARCH(1)}$;$LM_{ARCH(4)}$	0.94; 0.67	0.69; 0.92	0.65; 0.16	0.12; 0.38
$NRNL$	0.25	0.20	‡NA	0.32
$PC(1); PC(2); PC(3)$	0.05; 0.11; 0.24	0.69; 0.17; 0.24	‡NA	0.93; 0.67; 0.82
$JB; K_3; K_4$	0.36; -0.23; 3.5	0.36; -0.22; 3.4	0.37; 0.22; 3.51	0.007; 0.62; 3.9
Figures in parentheses beneath the parameter values are standard deviations and those after the test statistics are p-values. ‡Matrix inversion problem. † σ_s is the sample standard deviation of the transition variable s_t.				

Table 6: Estimation and evaluation tests results for CIF prices. Figures in parentheses beneath the parameters are standard deviations. Figures pertaining to the evaluation tests are p-values.

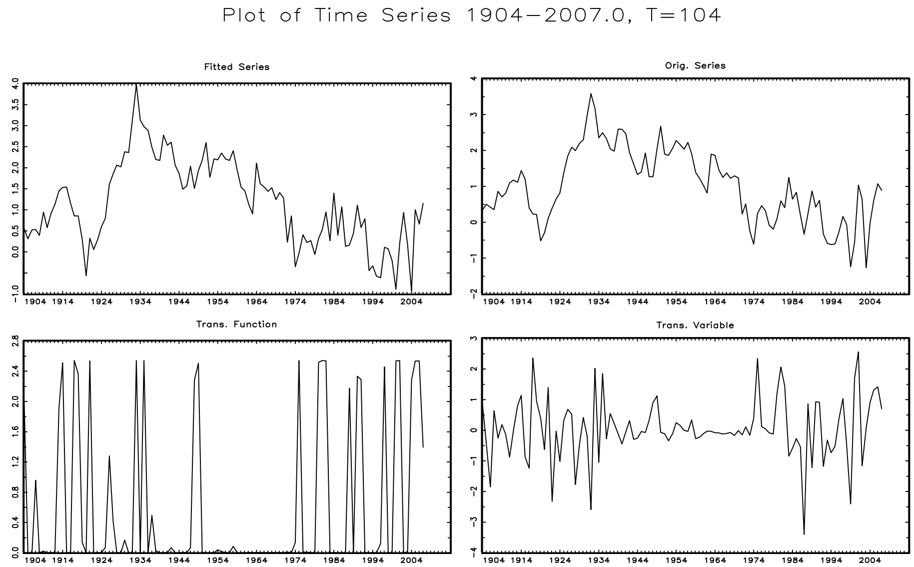

Figure 11: The figures, from left to right row-wise, show the fitted series $\widehat{y}$, the original series y_t, the transition function, and the transition variable r_{t-1} for bananas.

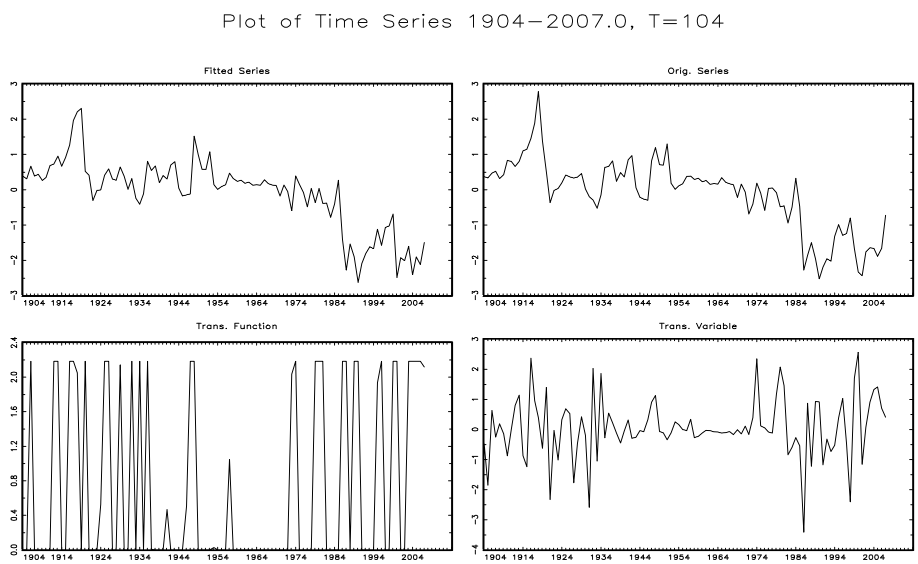

Figure 12: The figures, from left to right row-wise, show the fitted series $\widehat{y}$, the original series y_t, the transition function, and the transition variable r_t for palmoil.

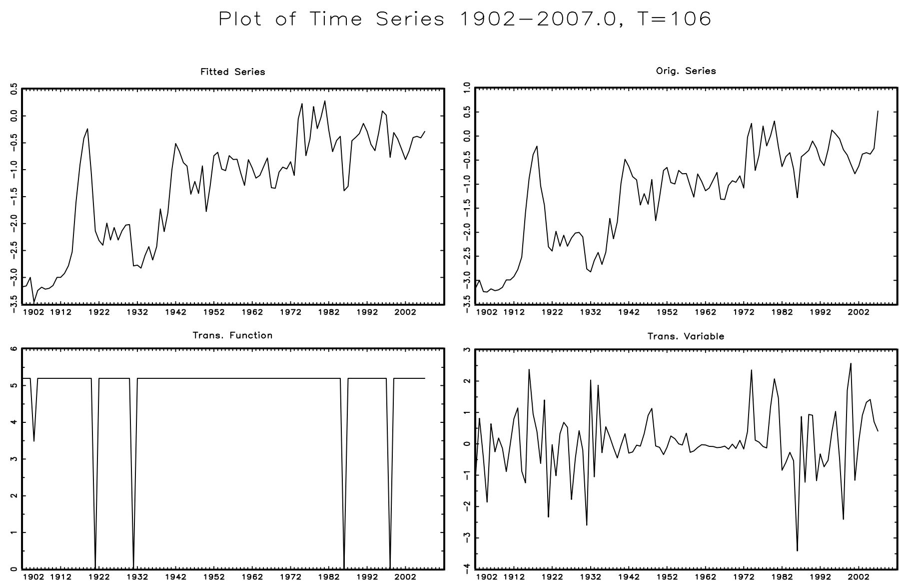

Figure 13: The figures, from left to right row-wise, show the fitted series $\widehat{y}$, the original series y_t, the transition function, and the transition variable r_t for timber.

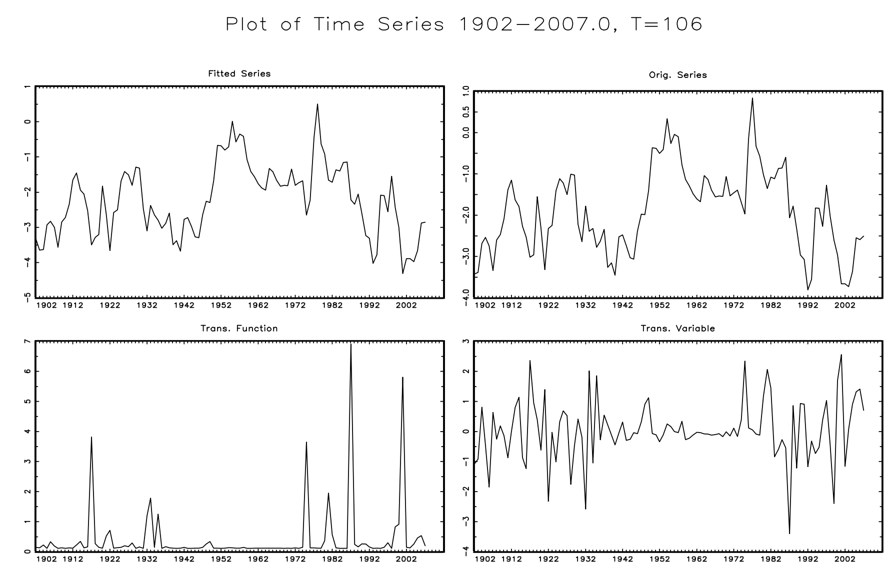

Figure 14: The figures, from left to right row-wise, show the fitted series $\widehat{y}$, the original series y_t, the transition function, and the transition variable r_{t-1} for coffee.

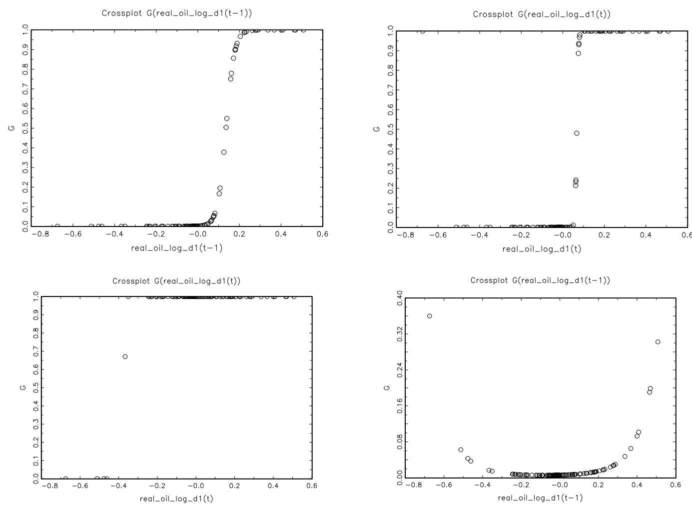

Figure 15: Transition functions of bananas (top left), palmoil (top right), timber (bottom left), and coffee (bottom right). Each dot corresponds to one observation.

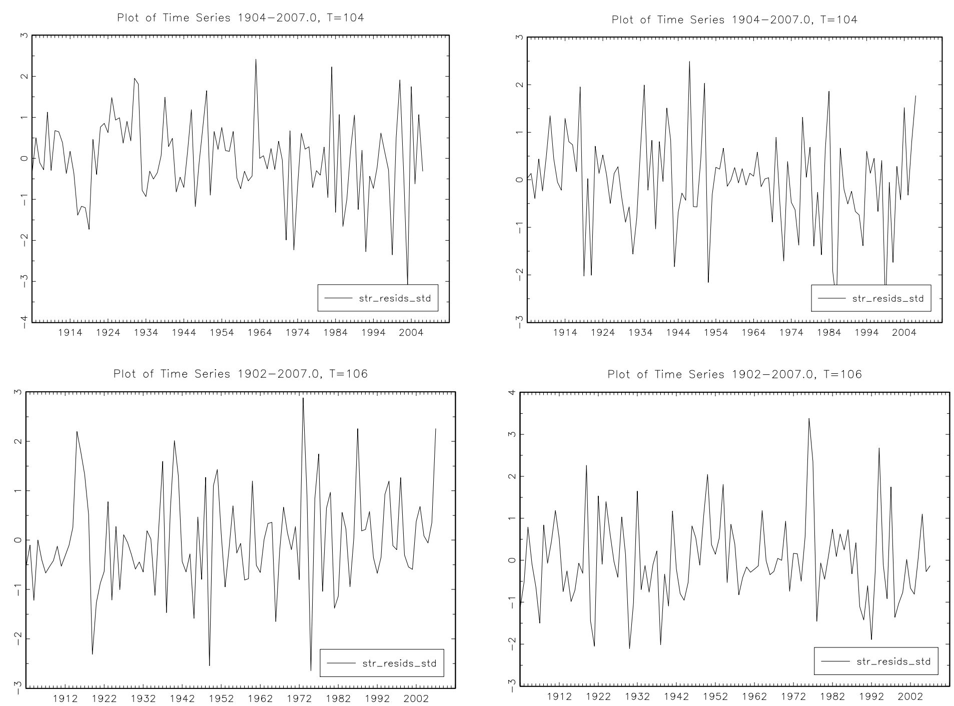

Figure 16: Standardized residuals for bananas (top left), palmoil (top right), timber (bottom left), and coffee (bottom right). Each dot corresponds to one observation.

The estimation and misspecification tests results for CIF group are reported in Table 6. The transition variable that showed the highest rejections of linearity tests was the growth rate of real price of oil (or its one period lag) and the selected models where LSTR1($p = 3$) for bananas, LSTR2($p = 3$) for palm oil, LSTR1($p = 1$) for timber, and LSTR2($p = 1$) for coffee. This is consistent with our claim that oil prices play a major role in the behavior of the commodities recorded on a CIF basis. This also confirms the connection between commodity prices, oil, and consumer prices that was discussed in Section 2. All estimated coefficients are significant at the 5% percent level of significance and the models pass all misspecifications tests. The transition from one regime to the other was smooth in case of bananas and coffee. This is confirmed from the moderate slope of the transition functions in both cases (see Figures 11 and 14) or the dot plots of the transition functions in Figure 15.

5 Conclusion

In this paper, we proposed a novel framework for modelling and understanding the behavior of individual commodity prices. In particular, prior to any econometric analysis, we suggested to closely examine the term of sale (Incoterm) pertaining to each commodity, its trading route, and its major exporter/importer. By doing so, we can classify nonlinear commodities according to their border prices. In each group, we can then proceed to find the best transition variable that can capture this nonlinearity.

The advantage of this proposed border price classification is twofold: (1) It allows us to determine the best potential transition variable that is capable of capturing the nonlinear dynamic in each group. In particular, the classification suggests that inflation rate (or any other macro economic news variable) in the exporting country is capable of capturing the nonlinearity in FOB prices whereas oil price is the best transition variable for CIF prices. (2) The border price classification is capable of explaining the observed co-movement in commodity prices; all CIF and FOB prices, being pushed from one regime to another by the classification's suggested transition variable, exhibit similar dynamics.

Our results confirm these predictions; current and one period lag U.S. inflation rate captured nonlinearity in all nonlinear U.S. FOB prices of the Grilli and Yang (1998) data set whereas current and one period lag of the growth rate of real oil price captured the dynamics in all nonlinear CIF prices. Both transition variables were capable of capturing the regime switching dynamics in each group via smooth transition regression models. This classification, thus, is successful in explaining and understanding the observed co-movement in commodity prices.

References

[1] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19: 716-723.

[2] Bacon, D. W. and Watts, D. G. (1971). Estimating the transition between two intersecting lines, *Biometrika,* 58: 525–534.

[3] Beach, C. (1977). An alternative approach to the specification of structural transition functions, *Canadian Journal of Economics*, 10: 132-141

[4] Beck, S. (2001). Autoregressive conditional heteroscedasticity in commodity spot prices, *Journal of Applied Econometrics*, 16: 115-132.

[5] Blomberg, B. S. and Harris, E. S. (1995). The commodity-consumer prices connection: fact or fable? *Federal Reserve Bank of New York Economic Policy Review,* October, 21-38.

[6] Boughton, J. M. and Branson, W. H. (1991). Commodity prices as a leading indicator of inflation. In *Leading Economic Indicators: New Approaches and Forecasting Records* (Lahiri, K. and Moore, G. H., eds.). Cambridge University Press, 305-338.

[7] Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 64: 247-254.

[8] Eitrheim, ∅. and Teräsvirta, T. (1996). Testing the adequacy of smooth transition autoregressive models. *Journal of Econometrics*, 74: 59-75.

[9] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50: 987-1007.

[10] Fahmy, H. (2011). *Regime switching in commodity prices* (thesis). Concordia University, Montreal, Canada.

[11] Fahmy, H. (2014), Modelling nonlinearities in commodity prices using smooth transition regression models with exogenous transition variables, *Journal of Statistical Methods and Applications* 23: 577-600.

[12] Frankel, J. A. (1986). Expectations and commodity price dynamics: the overshooting model. *American Journal of Agricultural Economics*, 68: 344–348.

[13] Fuhrer, J. and Moore, G. (1992). Monetary policy rules and the indicator properties of asset prices. *Journal of Monetary Economics*, 29(2): 303-336.

[14] Godfrey, L. G. (1988). *Misspecification tests in econometrics.* Cambridge, U.K.: Cambridge University Press.

[15] Granger, C. W. J. and Teräsvirta, T. (1993). *Modelling nonlinear economic relationships*. Oxford University Press: Oxford.

[16] Grilli, E. R. and Yang, C. (1988). Primary commodity prices, manufactured goods prices, and the terms of trade of developing countries: What the long run shows. *World Bank Economic Review,* 2: 1-47.

[17] Gustafson, R. L. (1958). Carryover levels for grains. *U.S.D.A. Technical Bulletin*, 1178.

[18] Jarque, C. M. and Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55: 163-172.

[19] Kyrtsou, C. and Labys, W. C. (2006). Evidence for chaotic dependence between US inflation and commodity prices. *Journal of Macroeconomics,* 28: 256-266.

[20] Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time-series models. *Biometrika,* 65: 297-303.

[21] Luukkonen, R., Saikkonen, P., and Teräsvirta, T. (1988). Testing linearity against smooth transition autoregressive models. *Biometrika,* 75: 491-499.

[22] Manning, G. (1956). *The Theory and Technique of Ship Design*, MIT Press, Cambridge, MA

[23] Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica,* 29: 315-335.

[24] Pfaffenzeller, S., Newbold, P., and Rayner, A. (2007). A short note on updating the Grilli-Yang commodity price index. *World Bank Economic Review,* 21(1): 151-163.

[25] Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes, *Journal of the American Statistical Association* 53: 873-880.

[26] Ronen, D. (1982). The effect of oil price on the optimal speed of ships, *Journal of Operational Research Society* 33: 1035-1040.

[27] Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of American Statistical Association,* 208-218.

[28] Teräsvirta, T. (1998). Modeling economic relationships with smooth transition regressions. In *Handbook of Applied Economic Statistics*, Marcel Dekker: New York, 507-552.

[29] Tong, H. (1978), 'On a Threshold Model', in *Pattern Recognition and Signal Processing*, ed. C.H. Chen, Sijhoff & Noordhoff: Amesterdam.