Threshold effects and unit roots of real commodity prices since the mid-nineteenth century

Pedro Clavijo-Cortes[1,2*] • Jacobo Campo-Robledo[2] • Henry Mendoza-Tolosa[2]

[1]University of Utah, Salt Lake City, Utah, US
[2]Facultad de Ciencias Económicas, Universidad Católica de Colombia, Bogotá, Colombia

Abstract

This document investigates whether a unit root process and nonlinearities can characterize real commodity prices for six major primary goods. An unconstrained two-regime threshold autoregressive model is used with an autoregressive unit root. Among the main results, it is found that terms of trade for agricultural, mineral, non-tropical, and non-oil goods are nonlinear processes that are characterized by a unit root process. The finding of nonlinearities explains why the deterioration of the terms of trade has been episodic. Additionally, we found there is no statistical evidence supporting the Prebisch-Singer hypothesis for agricultural, mineral, non-tropical, and non-oil goods.

Keywords: terms of trade; threshold autoregression; Prebisch-Singer hypothesis; unit root
JEL Classification Codes: Q02, O13

1. Introduction

One of the central concerns of the early development economists in Latin America was the understanding of the primary terms of trade long-run behavior. In general, for highly commodity-dependent developing countries, it becomes a significant policy concern to understand the underlying dynamic of the terms of trade properly. It is no surprise, then, that a considerable number of man-hours have been devoted to empirically analyze the behavior and the trajectory of primary terms of trade over time (see, e.g., Grilli and Yang 1988; Zanias 2005; Bidarkota and Crucini 2000; Ghoshray 2011; Erten 2011).

  Early studies, however, have focused mainly on studying the well-known Prebisch-Singer hypothesis, which predicts a secular decline in primary terms of trade over the long run. The bulk of the existing empirical literature has ignored the possibility that commodity prices may

---

contain threshold effects. Typically, this literature assumes that prices are characterized by a linear trend stationary process. Deaton (1999), and Wang and Tomek (2007), for instance, argue that primary terms of trade are better characterized by a trend stationary process. In this regard, Newbold and Vougas (1996) found that the Prebisch-Singer hypothesis requires the primary terms of trade to be stationary.

The lack of studies on nonlinearities of the terms of trade is intriguing since real prices for major categories of commodities exhibit periodic sharp spikes separated by prolonged slumps, which might presuppose that terms of trade move between a low and a high regime over time. Deaton and Laroque (1995) argue that because storage cannot be negative, commodity prices are inherently nonlinear. Holt and Craig (2006) document why primary goods might exhibit nonlinearities. They argue that discrepancies between supply and demand of primary goods caused by the interaction with related goods originate nonlinearities in commodity prices.

Indeed, it is precisely for those nonlinearities that asymmetric price adjustment is the rule rather than the exception in economic theory (Peltzman 2000). Those asymmetries are particularly relevant to study the interaction between developed and developing countries where terms of trade play a crucial role. In North-South trade models, which are widely employed to analyze this interaction, prices in the North, due to the higher degree of development, tend to be more rigid than prices in the South, causing uneven development. North-South trade models, however, have neglected nonlinearities in modeling terms of trade, even though asymmetric rigidity has profound implications for wealth distribution (see, Meyer and von Cramon-Taubadel 2004 for a survey on nonlinearities in prices).

Recently, Erten and Ocampo (2013) have provided evidence of nonlinearities in real commodities prices. Following the work of Cuddington and Jerrett (2008a, 2008b), Erten and Ocampo present evidence of super cycles and nonlinear trends that drive the behavior of terms of trade in the long run. They use a band-pass filter approach to decompose the series of six aggregate real primary prices and find four super cycles during 1865-2010. A couple of significant findings of their study are that super cycles are demand-driven and that the mean of each super cycle tends to be lower than that of the previous cycle. Thus, they find a stepwise deterioration over the period in support of the Prebisch-Singer hypothesis contradicting Thirlwall and Bergevin (1985).

Despite the abundance of studies on the behavior of terms of trade's trend over time, nonlinearity has not been exploited to the fullest yet in the modeling of such essential variables. Recently, Geronimi and Taranco (2018) address this gap in the literature by employing a Markov regime-switching (MRS) model to determine how terms of trade for primary commodities have switched among regimes. The study reveals that there are three regimes over the 1900–2016 period, with the second regime (2005–) associated with a level of prices significantly higher. Geronimi and Taranco did not, however, formally test for nonstationarity, a potentially important component of any formal assessment of the Prebisch-Singer hypothesis.

Motivated by these considerations, this document tests for the presence of nonlinearities and unit root in terms of trade (TOTs) of agricultural (*A*), mineral (*M*), tropical (*T*), non-tropical (*NT*), oil (*O*), and non-oil (*P*) goods over the period 1865-2015. This document aims to test the hypothesis of asymmetric price adjustment. Additionally, to examine possible non-mean-reverting behavior in TOTs to assess the Prebisch-Singer hypothesis in line with the approach proposed by Newbold and Vougas (1996). This document is closely related to the work of Gil-Alana et al. 2012 who also have investigated possible nonstationarity in TOTs although without investigating the Prebisch-Singer hypothesis. In this sense, this document complements the related literature.

Several nonlinear modeling techniques have been suggested in the literature to capture nonlinearities in time series. In this document, we use the Threshold Autoregressive (TAR)

model, which has not been employed before to address these issues related to nonlinearity and nonstationarity in terms of trade. TAR model differs from conventional linear econometric models by assuming the existence of different segments or regimes. Within these regimes, the time series may exhibit different behavior such as stationarity or lack of it. Likewise, TAR model differs from other nonlinear models such as MRS models, not only by its simplicity but also because they allow describing the conditions that separate the regimes. We also prefer the TAR model over the Smooth Threshold Autoregressive (STAR) alternative since the discontinuous jump that characterizes TAR model better describes the speed of price adjustment in these commodities markets.

TAR was introduced by Tong (1978) to model regime switching in time series, and further developments were done by Tong and Lim (1980) and Tong (1983, 1995). TAR model features asymmetric cycles and jump phenomena between two or more branches governed by a threshold variable. This study relies upon the Caner and Hansen (2001) extension of a two-regime TAR model with an autoregressive unit root. Caner and Hansen loosen the assumption that series is stationary, ergodic, and has no unit root. This allows discriminating nonstationarity from nonlinearity. Caner and Hansen's (2001) econometric framework allows us to test for nonlinearity and nonstationarity to strengthen the previous literature and to assess our aim.

The rest of the document is organized as follows: Section 2 presents the methodology of Caner and Hansen to cope with nonstationarity and nonlinearity simultaneously; Section 3 shows data and results from Caner and Hansen methodology; and Section 4 concludes.

2. Methodology

This section borrows heavily on the work of Caner and Hansen (2001), who consider a two-regime TAR(k) model with an autoregressive unit root:

$$\Delta TOT_t^j = \theta_1' x_{t-1} 1_{\{Z_{t-1} < \lambda\}} + \theta_2' x_{t-1} 1_{\{Z_{t-1} \geq \lambda\}} + \varepsilon_t \tag{1}$$

where, $x_{t-1} = (TOT_{t-1}^j \, r_t' \Delta TOT_{t-1}^j \cdots \Delta TOT_{t-k}^j)'$, $1_{\{\cdot\}}$ is the indicator function, ε_t is an iid error, $Z_{t-1}^j = TOT_{t-1}^j - TOT_{t-m-1}^j$ for some $m \geq 1$ is the value of the threshold variable, r_t is a vector of deterministic components that includes an intercept and a linear time trend, $j = A, M, T, NT, O, P$, and λ is a threshold parameter, which takes values in the interval $\lambda \in \Lambda = [\lambda_1, \lambda_2]$, where λ_1, λ_2 are chosen such that $P(Z_t^j \leq \lambda_1) = \pi_1 > 0$ and $P(Z_t^j \leq \lambda_2) = \pi_2 < 1$. The components of θ_1' and θ_2' can be partitioned as follows:

$$\theta_1 = \begin{pmatrix} \rho_1 \\ \beta_1 \\ \alpha_1 \end{pmatrix}, \quad \theta_2 = \begin{pmatrix} \rho_2 \\ \beta_2 \\ \alpha_2 \end{pmatrix}$$

where ρ_1 and ρ_2 are scalar, β_1 and β_2 have the same dimension as r_t, and α_1 and α_2 are k-vectors. Thus (ρ_1, ρ_2) are the slope coefficients on TOT_{t-1}^j, (β_1, β_2) are the slopes on the deterministic components, and (α_1, α_2) are the slope coefficients on $(\Delta TOT_{t-1}^j, \ldots, \Delta TOT_{t-k}^j)$ in the two regimes. When estimating "Eq. 1", for each $\lambda \in \Lambda$, θ_1 and θ_2 are estimated by least squares. The least square estimate of the threshold λ and of the corresponding vectors θ_1 and θ_2 are those that minimize the residual sum of squares. Additionally, Caner and Hansen (2001) suggest that setting the bounds of the trimming region to $\pi_1 = 0.15$ and $\pi_2 = 0.85$ provides a reasonable trade-off between the power and size properties of the test for threshold effects.

Using "Eq. 1", we want to test two things: threshold effect and stationarity property. The threshold effect is based on the null hypothesis $H_0: \theta_1 = \theta_2$, which is tested using the Wald statistic: $W_T = W_T(\hat{\lambda}) = \sup_{\lambda \in \Lambda} W_T(\lambda)$. Caner and Hansen (2001) find that under the restriction of a unit root process, the asymptotic distribution of W_T depends on the data structure, which implies that critical values cannot be tabulated, and therefore the authors

suggest two bootstrap methods to approximate the asymptotic distribution of W_T. Caner and Hansen (2001) perform a Monte-Carlo experiment to explore the size and power properties of the bootstrap W_T test. Their results suggest that the test is free from size distortions and that the power of the test increases with the magnitude of the threshold effect. On the other hand, the stationarity of the process TOT_t^j can be established in two ways. First, when there is a unit root in both regimes, the null hypothesis is of the form $H_0: \rho_1 = \rho_2 = 0$. Second, when there is a unit root in only one of the regimes, a case of partial unit root, the alternative hypothesis is of the form:

$$H_2: \begin{cases} \rho_1 < 0 \text{ and } \rho_2 = 0 \\ \quad\quad or \\ \rho_1 = 0 \text{ and } \rho_2 < 0 \end{cases}$$

The null hypothesis is tested against the unrestricted alternative $\rho_1 \neq 0$ or $\rho_2 \neq 0$ using the Wald statistic: $R_{2T} = t_1^2 + t_2^2$. Here, t_1 and t_2 are the t ratios for $\hat\rho_1$ and $\hat\rho_2$ from the OLS estimation. However, Caner and Hansen (2001) claim that this two-sided Wald statistic may have less power than a one-sided version of the test. As a result, they propose the following one-sided Wald statistic:

$$R_{1T} = t_1^2 \mathbb{1}_{\{\hat\rho_1 < 0\}} + t_2^2 \mathbb{1}_{\{\hat\rho_2 < 0\}}$$

To distinguish between the stationary case given as H_1 and the partial unit root case given as H_2, Caner and Hansen (2001) suggest using individual t statistics t_1 and t_2. If only one of $-t_1$ and $-t_2$ is statistically significant, this will be consistent with the partial unit root case H_2. Thus, R_{1T} and R_{2T} allow us to test whether the series is stationary while t_1 and t_2 allow us to identify which of the regimes is stationary. In the context of this study, these tests statistics can be used to examine the validity of the Prebisch-Singer hypothesis.

To obtain maximum power from these tests, the relevant critical values are computed using bootstrap methods, as Caner and Hansen (2001) suggest. Using simulations, Caner and Hansen (2001) show that bootstrap unit root tests have better power than the conventional ADF unit root test when the correct process is nonlinear. For testing the unit root, it is possible to treat the threshold as unidentified, in which case the bootstrap is based on a linear AR model. The alternative is to treat the threshold as identified and to base the bootstrap on simulations from a unit root TAR model. The approach here is to perform an unidentified threshold bootstrap for TOTs, as Caner and Hansen (2001) recommend.

In section 3, we present the results from the unconstrained model. Although the unconstrained version has more parameters than the constrained model (when $\rho_1 = \rho_2 = 0$), Caner and Hansen show that for all cases, the size of both tests is excellent, and the two bootstrap procedures have near identical size in their simulations. Their simulations suggest that it might not matter much which method is used. Caner and Hansen recommend following a conservative approach, which is precisely the unconstrained case. Additionally, ρ_1 and ρ_2 measure the degree of stationarity of TOTs so they can be used for comparison purposes with other unit root tests in related works.

3. Data and results

The empirical analysis is based on annual data for the natural logarithm of TOTs of agricultural, mineral, tropical, non-tropical, oil, and non-oil goods. We develop our econometric estimations using the same observations for the period 1865-2015. The data is an extension of that used by Erten and Ocampo (2013) who for the non-oil commodity prices, use annual data composed of prices for 24 commodities up to 1961 and 32 since 1962, grouped into five indices: total, metals, total agriculture, tropical agriculture, and non-tropical (or temperate zone) agriculture. As for oil prices, Erten and Ocampo base their calculations on spliced series of West Texas

International using data from the World Economic Outlook and the Global Financial Data. Real price indices are computed based on world manufacturing prices for the earlier part of the sample and the Manufacturing Unit Value index developed by the United Nations and updated by the World Bank for the later part.

The nonlinearity and unit root tests are applied to the natural logarithm of each real commodity price index. Consequently, the transition variable (Z_t^j) can be interpreted as TOTs' growth rate. This treatment of the transition variable introduces a dynamic threshold, which is not often seen in empirical investigations on nonlinearities. We employ the econometric apparatus presented in section 2 to ask if there is any statistical evidence to reject the linear AR model in favor of a threshold model. As explained in section 2, the Wald test W_T can be used to examine whether we can reject the linear autoregressive model in favor of a threshold model. We begin with a linear AR(5) model for all six series, which was mostly the best model according to a sequential t-test. For some series, the t-test selected a different lag order, although the alternative model presented serial autocorrelation in the error term. The AR(5) guarantees no serial autocorrelation and the right balance between sample size and lags order to capture nonlinearities.

The results of the Wald test are reported in Table 1. We also report the bootstrap critical values generated at conventional levels of significance and bootstrap p-values for threshold variables of the form Z_t^j for delay parameters m from 1 to 5. Bootstrap critical values and bootstrap p-values were computed based on 10,000 replications.

Table 1 merely reports values for m, which maximize the Wald test statistic for each series. As can be noted, the statistic is significant for A, P, T, and M, and therefore, we reject the null hypothesis of linearity in favor of the threshold model at the 5% and 10% level of significance. For NT and O the statistic is highly insignificant. Hence, there is strong evidence for threshold effects, and therefore, nonlinearities for A, P, T, and M. Table 1 also reports the threshold estimate. Thus, in the case of M, if the change in the log terms of trade, that is, the growth rate of the terms of trade, rises above 0.21 in the past five years, M switches to the second regime. The first regime is when $Z_t^M < 0.21$, which occurs when terms of trade are growing at a slower rate over a five-year period.

For those series that exhibit nonlinearity, we use the number of observations that lies within each regime to analyze the threshold effect. Approximately 15% and 85% of observations of A, 86% and 14% of M, 15% and 85% of P, and 82% and 18% of T fall in regime 1 and 2 respectively. The distribution of observations among regimes suggests that regime 1 is dominant in M and T, and regime 2 is dominant in A and P. This distribution has important welfare implications. It implies a different distribution of trade gains because gains' size changes depending on whether regime 1 is more dominant than regime 2. Furthermore, the finding of nonlinearities supports the claim that North-South trade models should model TOTs in a nonlinear fashion to better capture how unequal development might arise and persist.

Table 1. Nonlinearities tests.

TOT	m	W_T	10% Critical Value	5% Critical Value	Bootstrap p-value	Threshold Estimate
A	1	23.8	22.0	24.9	0.065	-0.07
NT	5	18.0	22.4	25.4	0.257	0.14
O	5	20.6	21.8	24.6	0.139	0.18
P	1	24.3	22.1	25.0	0.059	-0.07
T	4	25.4	21.9	24.8	0.043	0.22
M	5	42.7	22.1	24.9	0.001	0.21

Table 2. R_{1T} and R_{2T} unit root tests.

| | | | R_{1T} | | | | R_{2T} | | |
| | | | Bootstrap critical and p values | | | | Bootstrap critical and p values | | |
TOT	m	W_T	10%	5%	p-value	W_T	10%	5%	p-value
A	1	5.6	13.29	15.92	0.57	5.67	13.58	16.16	0.58
P	1	6.23	13.27	16.1	0.52	6.23	13.54	16.39	0.54
T	4	8.59	14.22	17.17	0.36	8.59	14.5	17.35	0.37
M	5	4.98	14.36	17.28	0.67	4.98	14.54	17.47	0.69

Table 3. t_1 and t_2 unit root tests.

| | | | t_1 | | | | t_2 | | |
| | | | Bootstrap critical and p values | | | | Bootstrap critical and p values | | |
TOT	m	t-stat	10%	5%	p-value	t-stat	10%	5%	p-value
A	1	-0.26	2.98	3.37	0.91	2.367	3.017	3.4368	0.2419
P	1	-0.026	2.978	3.359	0.875	2.495	3.056	3.4492	0.2198
T	4	2.46	3.11	3.52	0.24	1.597	3.095	3.4968	0.5014
M	5	2.02	3.07	3.46	0.36	0.941	3.139	3.5453	0.7104

Next, by using the unit root tests described in section 2, we analyze the behavior of these two regimes exclusively in those TOTs that exhibit nonlinearities. We have found evidence that terms of trade of agricultural, mineral, tropical, and non-oil goods are nonlinear processes. In what follows, we explore the unit root properties of these four terms of trade. As explained above, to investigate the unit root properties, we calculate the threshold unit root test statistics R_{1T}, R_{2T}, t_1, and t_2 for each delay parameter m, ranging from 1 to 5. However, the bootstrap critical and p values correspond to the m that is calculated in Table 1. We begin by examining the R_{1T} and R_{2T} tests results, which, together with the bootstrap critical values at the conventional levels of significance and the bootstrap p-values, are reported in Table 2. For instance, in the case of M when $m=5$, the W_T statistic obtained from the R_{1T} and R_{2T} tests are smaller than the bootstrap critical value at the 10% and 5% levels of significance. As a result, the null hypothesis of a unit root for M cannot be rejected. In general, the null of a unit root cannot be rejected for any series, implying that TOTs are nonstationary processes.

We examine further evidence on the unit root hypothesis by examining the individual t statistics, t_1 and t_2. The results are reported in Table 3. The t_1 test statistic is smaller than the bootstrap critical values for all series at 10% and 5% levels of significance. Likewise, t_2 is smaller than the bootstrap critical values at the same significance levels. This leads us to the conclusion that the unit root null cannot be rejected.

On the question of whether there is a tendency for primary prices to decline, as the Prebisch-Singer hypothesis suggests, the unit root tests results suggest there is no statistical evidence in favor of this hypothesis, according to Newbold and Vougas (1996).

In general, our results provide a different approach to understand the discontinuous deterioration of the terms of trade (see, Ocampo and Parra, 2004). The observed discontinuity in the deterioration of the terms of trade is indeed illustrating movements of the terms of trade between two different regimes due to the nonlinear nature of these variables. This conclusion can be discerned at least for A, P, T, and M.

4. Final remarks

Motivated by the possibility that terms of trade of commodities may be characterized by a nonlinear data generating process, we test for nonlinearities for six major commodities prices. We aimed to extend the analysis further by testing the presence of a unit root to contrast the

theory that real commodities prices should be stationary. Thus, in this document, we use an asymptotic theory of inference for a two-regime TAR model with an autoregressive unit root developed by Caner and Hansen. The main finding is that terms of trade for agricultural, mineral, non-tropical, and non-oil goods, are nonlinear and nonstationary. Nonlinearity justifies the use of nonlinear techniques (such as TAR models) to characterize terms of trade in future research on the matter. Likewise, future research should employ nonlinear unit root tests, since the typical test assumes that the data generating process is linear. Finally, we did not find evidence supporting the Prebisch-Singer hypothesis for agricultural, mineral, non-tropical, and non-oil goods.

Acknowledgments

References

Bidarkota, P., and Crucini, M. J. (2000) Commodity Prices and the Terms of Trade, *Review of International Economics,* 8, 647-666

Caner, M., and Hansen, B. (2001) Threshold autoregression with a unit root, *Econometrica,* 69(6), 1555-1596

Cuddington, J. T., and Jerrett, D. (2008a) Super-cycles in real metal prices?, *IMF Staff Papers,* 55(4), 541-565

Cuddington, J. T., and Jerrett, D. (2008b) Broadening the statistical search for metal price super cycles to steels and related metals, *Resources Policy,* 33(4).

Deaton, A. (1999) Commodity Prices and Growth in Africa, *Journal of Economic Perspectives,* 13, 23-40

Deaton, A., and Laroque, G. (1995) Estimating a Nonlinear Rational Expectations Commodity Price Model with Unobservable State Variables, *Journal of Applied Econometrics,* 10, S9–S40.

Erten, B. (2011) North-South terms-of-trade trends from 1960 to 2006, *International Review of Applied Economics,* 25, 171-184

Erten, B., and Ocampo, J.A. (2013) Super Cycles of Commodity Prices Since the Mid-Nineteenth Century, *World Development,* 44, 14-30

Geronimi, V., and Taranco, A. (2018) Revisiting the Prebisch-Singer hypothesis of a secular decline in the terms of trade of primary commodities (1900–2016). A dynamic regime approach, *Resources Policy,* 59, 329-339

Gil-Alana, L., Cunado, J., and Pérez de Garcia, F. (2012) Persistence, Long Memory, and Unit Roots in Commodity Prices, *Canadian Journal of Agricultural Economics*, 60(4), 451-468

Ghoshray, A. (2011) A reexamination of trends in primary commodity prices, *Journal of Development Economics,* 95, 242-251

Grilli, E. R., and Yang, M. C. (1988) Primary Commodity Prices, Manufactured Goods Prices, and the Terms of Trade of Developing Countries: What the Long Run Shows, *The World Bank Economic Review,* 2, 1-47

Holt, M. T., and Craig, L. A. (2006) Nonlinear Dynamics and Structural Change in the U.S. Hog–Corn Cycle: A Time-Varying STAR Approach, *American Journal of Agricultural Economics,* 88, 215–33.

Meyer, J., and von Cramon-Taubadel, S. (2004) Asymmetric price transmission: A survey, *Journal of Agricultural Economics*, 55, 581–611.

Newbold, P., and Vougas, D. (1996) Drift in the relative price of primary commodities: a case where we care about unit roots, *Applied Economics*, 28, 653–661.

Ocampo, J. A., and Parra, M. A. (2004) *The Terms of Trade for Commodities in the Twentieth Century*, Econ Working Paper Archive, International Trade Series No. 0402006.

Peltzman, S. (2000) Prices Rise Faster than they fall, *Journal of Political Economy*, 108(3), 466-502.

Thirlwall, A. P., and Bergevin, J. (1985) Trends, Cycles, and Asymmetries in the Terms of Trade of Primary Commodities from Developed and Less Developed Countries, *World Development,* 13, 805-817

Tong, H. (1978) *On a Threshold Model.* In Chen, C.H. (Ed): Pattern Recognition and Signal Processing. Sijthoff & Noordhoff: Amsterdam, 574-586

Tong, H. (1983) *Threshold Models in Nonlinear Time Series Analysis* (Lecture Notes in Statistics No. 21), Springer-Verlag: New York.

Tong, H. (1995) A personal overview of non-linear time series analysis from a chaos perspective, *Scandinavian Journal of Statistics,* 22, 399-445

Tong, H., and Lim, K.,S. (1980) Threshold Autoregression, Limit Cycles and Cyclical Data (with discussion), *Journal of the Royal Statistical Society,* 42, 245-292.

Wang, D., and Tomek, W. G. (2007) Commodity Prices and Unit Root Tests, *American Journal of Agricultural Economics,* 89, 873-689

Zanias, G. P. (2005) Testing for trends in the terms of trade between primary commodities and manufactured goods, *Journal of Development Economics,* 78, 49-59