UNITED NATIONS UNIVERSITY

UNU-WIDER

World Institute for Development Economics Research

Working Paper No. 2011/71

The Long-Run Behaviour of the Terms of Trade between Primary Commodities and Manufactures

A Panel Data Approach

Jesús Otero[1] and Ana María Iregui[2]

November 2011

Abstract

This study examines the Prebisch and Singer hypothesis using a panel of 24 commodity prices from 1900 to 2010. The modelling approach stems from the need to meet two key concerns: (1) the presence of cross-sectional dependence among commodity prices; and (2) the identification of potential structural breaks. To address these concerns, the Hadri and Rao test (2008) is employed. The findings suggest that all commodity prices exhibit a structural break at different locations across series, and that support for the Prebisch and Singer hypothesis is mixed. Once the breaks are removed from the underlying series, the persistence of commodity price shocks is shorter than that obtained in other studies using alternative methodologies.

Keywords: Prebisch and Singer hypothesis, panel stationarity.
JEL classification: O13, C33

[1]Universidad del Rosario, Bogotá, Colombia; [2]Unidad de Investigaciones, Banco de la República, Bogotá, Colombia.

This study has been prepared within the UNU-WIDER project ‘New Directions in Development Economics’.

UNU-WIDER gratefully acknowledges the financial contributions to the research programme by the governments of Denmark (Ministry of Foreign Affairs), Finland (Ministry for Foreign Affairs), Sweden (Swedish International Development Cooperation Agency—Sida) and the United Kingdom (Department for International Development).

ISSN 1798-7237 ISBN 978-92-9230-438-6

Acknowledgements

Work on this study began while Jesús Otero was a Visiting Scholar at the World Institute for Development Economics Research (WIDER) of the United Nations University (UNU) in Helsinki. Jesús Otero would like to express his gratitude to UNU-WIDER for providing a welcoming and supportive environment for this research. Its financial support is also gratefully acknowledged. We thank Stephan Pfaffenzeller not only for providing us with the dataset used in Pfaffenzeller et al. (2007), but also for updating the dataset until 2010. We also thank Imed Drine, Monica Giulietti, James Thurlow and seminar participants at UNU-WIDER for their constructive comments and suggestions that helped to improve our study. The views expressed here are the responsibility of the authors. Any remaining errors are ours.

Tables and figures appear at the end.

The World Institute for Development Economics Research (WIDER) was established by the United Nations University (UNU) as its first research and training centre and started work in Helsinki, Finland in 1985. The Institute undertakes applied research and policy analysis on structural changes affecting the developing and transitional economies, provides a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and promotes capacity strengthening and training in the field of economic and social policy making. Work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.

www.wider.unu.edu *publications@wider.unu.edu*

UNU World Institute for Development Economics Research (UNU-WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Typescript prepared by Rosaleen McDonnell.

1 Introduction

Ever since Prebisch (1950) and Singer (1950) put forward the hypothesis that there is a secular deterioration in the price of primary commodities relative to that of manufactures, the study of the long-run behaviour of the terms of trade of developing countries has received a great deal of attention. This is undoubtedly because the Prebisch and Singer (PS) hypothesis challenges the conventional classical view, according to which, rapid technical progress in the production of manufactures, the operation of the law of diminishing returns in the production of primary goods and a growing population would actually cause a long-run increase in the relative price of primary commodities. Bowley (1903), for instance, finds mixed supporting evidence for this bias against industrial countries, which tend to rely heavily on primary product imports. Indeed, Bowley reveals that British imports became more expensive relative to exports between 1873 and 1881, but that this trend is reversed in the years 1881-1901, when imports became less expensive than exports. Keynes (1912), invoking the operation of the law of diminishing returns, indicates that this *abnormal* tendency detected by Bowley during the last two decades of the nineteenth century, i.e. of British imports being less expensive than exports, comes to an end during the first decade of the twentieth century.[1]

PS attribute the secular deterioration of the terms of trade of developing countries to two main factors. First, in developed countries technical progress results in higher wages and improvements in the standard of living of the workers, due to the enhanced market power of trade unions, but not in lower prices of their products, some of which are exported to developing countries. By contrast, in developing countries technical progress does not result in higher wages, because of the presence of a Lewis (1954) type excess supply of labour, but in lower prices of their products. Thus, the benefits from technical progress are transferred from developing to developed countries or, in the terminology of Prebisch, from the periphery to the centre. Second, there is the combination of low price and income elasticities of demand for primary commodities relative to those of manufactures. Indeed, primary commodities (and most especially some agricultural products) can be regarded as necessities rather than luxuries, so that their income elasticity of demand is less than one. Thus, other things being equal, if increases in income shift the demand curve for primary commodities to the right by less than the corresponding shift in the demand curve for manufactures, the price of primary commodities relative to manufactures would tend to decline as time passes.

The importance of the PS hypothesis clearly relies on its main policy recommendation, that developing countries should avoid specialization according to their Ricardian comparative advantage. Balassa (1989: 1645-89) reviews the main proposals that have been made for joint action on the part of developing countries to compensate the secular deterioration of their terms of trade. These include the exploitation of the monopoly power that some developing countries may possess in the production of particular primary products, especially those that face an inelastic demand in developed countries.

[1] For a discussion of the early literature on the development of the terms of trade in Great Britain see Rostow (1950). More recently, Sarkar (1986) finds evidence of deterioration of the terms of trade in Great Britain during the first half of the nineteenth century.

Another policy proposal is for developing countries to follow an inward-oriented development strategy that would allow them to reduce their dependence on imports from developed countries. Thus, industrialization is to be promoted by means of government intervention for the protection of specific activities, which usually involves providing imported inputs at artificially low prices (typically accomplished either through overvalued exchange rates or through the presence of multiple exchange rates), as well as direct credit at subsidised interest rates; see Agénor and Montiel (1996: Chapter 1). The objective of these policy measures is to change the structure of production of developing countries, by limiting the role of prices as a signalling mechanism in the resource allocation process.

Because the validity of the PS hypothesis raises an empirical question, early criticisms focused on the inappropriateness and quality of the data. Then, attention turned to the fact that tests of the hypothesis were not based on a formal statistical procedure, but rather on informal approaches, such as visual inspection of the time-series data, and year-to-year comparisons. During the last three decades or so, however, these shortcomings have been addressed in two main ways. On the one hand, a significant amount of effort has been put on the creation of a consistent dataset for a relatively large number of commodity prices. An important contribution in this area is Grilli and Yang (1988), who constructed a US dollar commodity price index from 1900 to 1986, which consists of 24 internationally traded non-fuel commodities. Generally speaking, studies on the PS hypothesis can be classified into those that have analysed aggregate price indices of commodities, those that have looked at major commodity groupings, and those that have focused on individual commodities. On the other hand, tests of the hypothesis are now based on the application of recent developments of modern time-series econometrics. These developments include fitting regressions against time, estimating structural (also referred to as unobserved components) time-series models, and applying unit root tests (also allowing for structural breaks at known and unknown dates).

This study aims to further our understanding of the PS hypothesis by looking at two aspects that, to the best of our knowledge, have not been explored thus far in the existing literature. First, we will use a panel data framework to analyse the prices of the 24 primary commodities that make up the Grilli and Yang (1988) index. A panel data framework not only allows us to examine the potential effect of cross-sectional dependence among commodity price indices, which may arise from common shocks or innovations (e.g. an increased demand for raw materials due to growth in developed countries), but also offers the advantage that, by combining information from the time-series and the cross-section dimensions, fewer time-series observations are required for statistical tests to have power. Second, we will apply statistical tests that take stationarity as the null hypothesis. Testing for stationarity, rather than for non-stationarity (i.e. the existence of a unit root), appears more suitable to assess the PS hypothesis because this hypothesis, as originally postulated, implies the existence of a long-run declining deterministic trend in relative commodity prices. For this purpose, we employ the residual-based Lagrange multiplier (LM) panel stationarity tests put forward by Hadri and Rao (2008), who extend the Hadri (2000) tests to accommodate one-time structural breaks and cross-section dependence.

It is interesting to notice that the use of panel stationarity tests in our empirical modelling approach is also attractive because it permits an alternative interpretation of the time path of relative commodity prices. Indeed, failure to reject the null hypothesis, that commodity prices relative to manufactured goods prices are jointly stationary, is equivalent to finding that they are cointegrated, and that the cointegrating parameter is equal to one (after allowing for the presence of one-time structural breaks and cross-section dependence). This implies that commodity prices and manufactured goods prices must be linked by a long-run equilibrium relationship; in the short run, however, prices may deviate from the long-run equilibrium relationship, although not by an ever growing amount since there will be economic forces that may be expected to act so as to restore equilibrium. This alternative interpretation turns out to be useful to test the 'strong' form of the definition of commodity price comovement proposed by Leybourne et al. (1994: 1751-2), according to which there exists comovement between a pair of series when they are cointegrated, and the cointegrating parameter is positive. In fact, notice that failure to reject the null hypothesis of joint stationarity gives rise to an even stronger form of the definition of commodity price comovement, because the cointegrating parameter is not only positive, but also equal to one for all of the individual commodity series in the panel.

The paper is organised as follows. Section 2 presents a brief review of the existing empirical literature on the PS hypothesis. Section 3 outlines the Hadri-based approaches to test for stationarity in panels of data, allowing for the presence of one-time structural breaks at unknown dates and cross-section dependence. Section 4 describes the data, presents the results of the empirical analysis and discusses some policy considerations. Section 5 offers some concluding remarks.

2 Brief literature review

Prebisch (1950) and Singer (1950) analyse average price indices of British imports and exports, which are used as proxy for the world prices of primary commodities and manufactured products, respectively, for the period 1876-80 to 1946-47. They find evidence that from the 1870s to the Second World War the trend of prices has moved against producers of primary commodities and in favour of producers of manufactures.[2] However, Spraos (1980) observes that early critics of Prebisch and Singer focused on the inappropriateness and quality of the price data used in their analyses. In particular, Spraos identifies four principal criticisms: (1) Great Britain cannot be considered as representative of the industrial countries as a whole; (2) primary products are also produced by developed countries; (2) exports are valued *f.o.b* while imports are valued *c.i.f*, so that changes in prices can be partly (or wholly) due to changes in transport costs; and (4) price indices do not adequately account for new products that are traded nor for improvements in the quality of existing ones. Kindleberger (1958), for instance,

[2] On a historical note, Singer (1950) does not present data but refers to his earlier unsigned United Nations (1949) study. Singer (1994: 43-8) indicates that '[i]t was Folke Hilgerdt, the Swedish economist and statistician ... who first mentioned this long-term data source to [him] and expressed puzzlement about its behavior. It was then that [Singer] wrote one of [his] first major research publications for the United Nations on "Relative Prices of Exports and Imports of Underdeveloped Countries," which was fortunate enough to catch [Prebisch's] eye. Meanwhile, independently [Singer] had worked on [his] own interpretation of the results of this study and presented it as a paper to the annual meeting of the American Economic Association in New York in December 1949, again practically to the day coinciding with Prebisch's own CEPAL studies on The Economic Development of Latin America and Its Principal Problems ... .' (p.48).

in a study that can be classified as belonging to category (i.) above, undertakes a more thorough statistical investigation that involves the construction of indices of industrial European export and import unit values, by commodity group and by trading area, from 1872 to 1952.[3] Kindleberger finds that although few generalizations can be made on the evolution of the terms of trade between world manufactures and world primary products, they tend to turn against developing and in favour of developed countries. Kindleberger also points out that caution must be exercised when interpreting these findings because they do not necessarily imply that the terms of trade of developing countries should be subjected to compensatory policies.

Since the 1980s there has been a revival of interest in the PS hypothesis. During this period authors have focused on the fact that formal statistical testing procedures were not considered in the original writings of PS. For instance, Spraos (1980) and Sapsford (1985) revisit the deterioration hypothesis by estimating semi-logarithmic regressions of (some measure of) developing countries' terms of trade against a constant and a time trend, and testing for the statistical significance of the estimated trend coefficient. Sapsford (1985) also tests for structural instability in the underlying trend coefficient using the Chow (1960) test for parameter constancy.

Grilli and Yang (1988) construct a US dollar commodity price index spanning from 1900 to 1986, consisting of 24 internationally traded non-fuel commodities. The Grilli and Yang (GY) dataset has become the most widely used data source in the literature related to the PS hypothesis; see, among others, Cuddington and Urzúa (1989), von Hagen (1989), Perron (1990), Powell (1991), Helg (1991), Ardeni and Wright (1992), Bleaney and Greenaway (1993), Newbold and Vougas (1996), León and Soto (1997), Kim et al. (2003), Zanias (2005), Kellard and Wohar (2006) and Ghoshray (2011). However, it is worth mentioning that the work of GY is not only important because they constructed a consistent dataset over a long period of time, but also because it is perhaps the first study that tests whether commodity prices can be viewed as trend-stationary (TS) or difference-stationary (DS) processes, based on the ADF unit root test of Dickey and Fuller (1979). Subsequently, the ADF test has also been employed by Bleaney and Greenaway (1993), Reinhart and Wickham (1994), and Kim et al. (2003), who analyse major commodity groupings, and by Cuddington (1992), who uses price data for individual commodities.[4]

Further extending the line of work based on unit root tests, Cuddington and Urzua (1989), Perron (1990), Helg (1991), Reinhart and Wickham (1994) and Newbold and Vougas (1996) apply the ADF test allowing for the presence of a known structural break.[5] León and Soto (1997) use the Zivot and Andrews (1992) unit root test in the presence of an endogenously determined structural break, while Zanias (2005) and Kellard and Wohar (2006) apply the Lumsdaine and Papell (1997) unit root test that allows for the presence of up to two endogenously determined structural breaks.

3 Industrial Europe is defined as Belgium, France, Germany, Italy, Luxemburg, Netherlands, Sweden, Switzerland and United Kingdom.

4 Reinhart and Wickham (1994) do not use the GY dataset, but quarterly data (1957q1 to 1993q2) for major commodity groupings: all non-oil commodities, beverages, food and metals.

5 In our literature review, Newbold and Vougas (1996) is the only paper that also tested the null hypothesis of stationarity, which is the testing strategy adopted in our paper. However, they do not study individual commodity price indices, nor account for structural breaks.

Ghoshray (2011) re-examines the PS hypothesis by employing the Lee and Strazicich (2003) unit root test with two endogenous breaks, which offers better size and power properties than both the Zivot and Andrews (1992) and Lumsdaine and Papell (1997) tests.

Among alternative approaches to unit root tests that have been implemented, von Hagen (1989) tests for cointegration between prices of commodities and manufactures using the two-step ordinary least squares (OLS) procedure of Engle and Granger (1987), while Powell (1991) tests for cointegration using the Engle-Granger procedure as well as the maximum likelihood estimator of cointegrated vector autoregressive (VAR) models of Johansen (1988), also allowing for the presence of known structural breaks.[6] Ardeni and Wright (1992) and Reinhart and Wickham (1994) employ the structural time-series approach advocated by Harvey (1989), according to which models are formulated directly in terms of three unobserved components of interest, namely trend, seasonal and irregular components.

In a recent contribution to the literature, Harvey et al. (2010) test the PS hypothesis using an entirely new and much longer dataset of 25 relative commodity price series (the specific commodity list, which includes 20 commodities already found in the GY dataset, is presented in Section 4). After consulting several historical sources, these authors manage to create an unbalanced panel of prices that goes back to 1650 for eight out of the 25 commodities under consideration. Then, they apply the Harvey et al. (2007, 2009) trend hypothesis tests, also allowing for one-time breaks, which involve the computation of a data-dependent weighted average of two trend statistics: one that is appropriate when the underlying series is stationary, and the other one when it is non-stationary. According to their results, there is evidence of a long-run negative trend in the relative price of eleven commodities, while no positive and significant trends were detected over all or part of the sample period for the remaining fourteen commodities. It is worth mentioning that, similar to the other studies existing in the literature, the results in Harvey et al. (2010) are based on a univariate analysis of the time-series properties of the commodity prices.

Perhaps it is no surprise that the results of the studies listed above provide mixed support for the PS hypothesis. Broadly speaking, these studies can be classified into three main groups. First, Spraos (1980), Sapsford (1985), Grilli and Yang (1988), Ardeni and Wright (1992), Bleaney and Greenaway (1993), Reinhart and Wickham (1994) and León and Soto (1997) confirm the negative sign (but not the magnitude) of the trend implicit in the works of PS. Second, Cuddington and Urzúa (1989), Perron (1990), Helg (1991), Powell (1991) and Zanias (2005) find that the relative price of primary commodities can be best characterised as a trendless process that exhibits a one-time negative shift. According to Cuddington and Urzúa (1989), this finding, strictly speaking, does not support the views of PS, because the latter refer to a secular terms of trade deterioration.[7] Third, von Hagen (1989), Cuddington (1992), Newbold and Vougas (1996), Kim et al. (2003), Kellard and Wohar (2006), Harvey et al. (2010) and Ghoshray (2011) do not find strong support for the PS hypothesis.

---

6 Powell (1991) refers to the effects of breaks on the critical values of the Engle-Granger test, but there is no mention of their effect for the Johansen test.

7 Singer (1999: 911), however, argues that '... it does not matter very much whether the data are interpreted as a persistent decline trend or as essentially stationary with intermittent downward breaks.' .

3 Testing for stationarity in panel data

In recent years, testing for unit roots in panel data has received a great deal of attention, as it is one possible way to achieve power gains over unit root tests applied to a single time series; see e.g. Breitung and Pesaran (2008: 279-322) for a literature review. Among the tests available in the literature, that of Im et al. (2003) (IPS) has proved to be one of the most commonly applied. The IPS test is based on averaging individual ADF statistics, and so it permits for all of the individual series in the panel to have a unit root under the null hypothesis. Within this framework, failure to reject the null hypothesis implies that all of the individual series can be characterised as DS (as opposed to TS) processes. However, given that the PS hypothesis implies that commodity prices exhibit a long-run declining deterministic trend, it appears that a more appropriate approach would be one that tests the null of stationarity around a level or around a (broken) trend.

Hadri (2000) develops a residual-based LM procedure to test the null hypothesis of stationarity for all the individual series in the panel, against the alternative that some (but not all) of the individual series have a unit root. As can be seen, this approach offers the key advantage that if the null hypothesis is not rejected, then one may conclude that all the commodity price indices in the panel are stationary. In particular, Hadri considers the following model specifications:

$$y_{it} = \alpha_i + r_{it} + \varepsilon_{it}, \tag{1}$$

$$y_{it} = \alpha_i + r_{it} + \beta_i t + \varepsilon_{it}. \tag{2}$$

where y_{it} denotes the observed series of commodity price index i at time t, $i = 1,..., N$, $t = 1,..., T$, r_{it} is a random walk, $r_{it} = r_{it-1} + u_{it}$, and ε_{it} and u_{it} are mutually independent normal distributions. In addition, ε_{it} and u_{it} are independent and identically distributed (i.i.d.) across i and over t, with $E[\varepsilon_{it}] = 0$, $E[\varepsilon_{it}^2] = \sigma_{\varepsilon,i}^2 > 0$, $E[u_{it}] = 0$ and $E[u_{it}^2] = \sigma_{u,i}^2 \geq 0$. Within this framework, the null hypothesis that all the individual series in the panel are stationary is $H_0 : \sigma_{u,i}^2 = 0$, where $i = 1,..., N$. The alternative hypothesis that some (but not all) of the individual series have a unit root is $H_1 : \sigma_{u,i}^2 > 0$, where $i = 1,..., N_1$; and $\sigma_{u,i}^2 = 0$, where $i = N_1 + 1,..., N$.

The models given in Equations (1) and (2) are used to test for level and trend stationarity, respectively. In a recent paper, Hadri and Rao (2008) extend the previous setup to allow for the presence of one-time structural breaks. More specifically, they postulate the following models of structural break under the null hypothesis:

$$y_{it} = \alpha_i + r_{it} + \delta_i D_{it} + \varepsilon_{it}, \tag{3}$$

$$y_{it} = \alpha_i + r_{it} + \delta_i D_{it} + \beta_i t + \varepsilon_{it}, \tag{4}$$

$$y_{it} = \alpha_i + r_{it} + \beta_i t + \gamma_i DT_{it} + \varepsilon_{it}, \tag{5}$$

$$y_{it} = \alpha_i + r_{it} + \delta_i D_{it} + \beta_i t + \gamma_i DT_{it} + \varepsilon_{it}, \tag{6}$$

where, in addition to the terms already defined, D_{it} and DT_{it} are dummy variables to specify the type of structural break, which are defined as:

$$D_{it} = \begin{cases} 1, & \text{if } t > T_{B,i} \\ 0 & \text{otherwise} \end{cases}, \tag{7}$$

and

$$DT_{it} = \begin{cases} t - T_{B,i}, & \text{if } t > T_{B,i} \\ 0, & \text{otherwise} \end{cases}, \tag{8}$$

where $T_{B,i}$ denotes the time of occurrence of the structural break for individual i. Also, $T_{B,i} = \omega_i T$, where $\omega_i \in (0,1)$ indicates the fraction of the break point relative to the whole sample period for individual i. The parameters δ_i and γ_i measure the extent (or magnitude) of the structural break, and allow for the possibility of different breaking dates across the individuals in the panel. The models in Equations (3) to (6) comprise the following characteristics. Equation (3) consists of an intercept term and allows for a shift in the level of the series. Equation (4) has intercept and linear trend terms, and admits a shift in the former (but not in the latter). Equation (5) includes intercept and linear trend terms, and permits a change in the latter (but not in the former). Lastly, Equation (6) also incorporates intercept and linear trend terms, and allows for a change in both the level and the slope of the series.[8]

Hadri and Rao (2008) use a systematic approach to find the appropriate model for each series y_{it}; it should be noticed that in implementing this approach, the models postulated in Equations (1) and (2) are also taken into account to allow for the possibility that there is no break in the underlying series y_{it}. Specifically, Hadri and Rao start off by determining the time of the break point endogenously, which involves estimating for each cross-section unit in the panel and for each model the break date, $\hat{T}_{B,i,k}$. This can be accomplished by minimising, with respect to $0 < \omega_i < 1$, the residual sum of squares (RSS) from the relevant model under the null hypothesis, where $i = 1,..., N$ denotes the commodity prices in the panel, and $k = 1, 2,..., 6$ refers to the models postulated in Equations (1) to (6). Then, given $\hat{T}_{B,i,k}$, for each individual in the panel, i, the preferred model, k, is chosen by minimising the Schwarz information criterion.[9]

---

8 Carrión-i-Silvestre et al. (2005) study the case of testing for panel stationarity with multiple structural breaks. However, they only consider the models formulated in Equations. (3) and (6).

9 Notice that in practice the models in Equations (1) and (2) are estimated only once, since they do not include the dummy variables D_{it} and DT_{it}.

Let $\hat{\varepsilon}_{it}$ be the residuals that result from estimating the chosen model (with or without a break). The individual univariate Kwiatkowski et al. (1992) (KPSS) stationarity test is given by:

$$\eta_{i,T,k}(\hat{\omega}_i) = \frac{\sum_{t=1}^{T} S_{it}^2}{T^2 \hat{\sigma}_{\varepsilon_i}^2}, \tag{9}$$

where S_{it} denotes the partial sum process of the residuals, that is:

$$S_{it} = \sum_{j=1}^{t} \hat{\varepsilon}_{ij}, \tag{10}$$

and $\hat{\sigma}_{\varepsilon_i}^2$ is a consistent estimator of the long-run variance of $\hat{\varepsilon}_{it}$ from the appropriate regression. KPSS use in their paper a nonparametric estimator of the long-run variance, $\hat{\sigma}_{\varepsilon_i}^2$, which is based on a Bartlett window with a truncation lag parameter of $l_q = \text{integer}\left[q(T/100)^{1/4}\right]$, where $q = 4,12$ (the value of the statistic turns out to be sensitive to the choice of q). However, Caner and Kilian (2001) indicate that stationarity tests, like the KPSS test, exhibit very low power after correcting for size distortions. Thus, in this study we follow recent work by Sul et al. (2005), who advocate the use of a new boundary condition rule to obtain a consistent estimate of $\hat{\sigma}_{\varepsilon_i}^2$. This rule is implemented in the following stages. First, an autoregressive (AR) model for the residuals is estimated, that is:

$$\hat{\varepsilon}_{it} = \rho_{i,1}\hat{\varepsilon}_{i,t-1} + \ldots + \rho_{i,p_i}\hat{\varepsilon}_{i,t-p_i} + \upsilon_{it}, \tag{11}$$

where the lag length of the autoregression, p_i, can be determined for example using the general to specific (GTS) algorithm proposed by Hall (1994) and Campbell and Perron (1991). The idea in the GTS algorithm is to start with some upper bound on p_i, denoted $p_i^{\max}$, then estimate Equation (11) with $p_i = p_i^{\max}$, and test the statistical significance of $\rho_{i,p_i^{\max}}$. If this coefficient is statistically significant, using for instance a significance level of 10 per cent, one selects $p_i = p_i^{\max}$. Otherwise, the order of the estimated autoregression in (11) is reduced by one until the coefficient on the last included lag is found to be statistically significant. Second, $\hat{\sigma}_{\varepsilon_i}^2$ is obtained after applying the boundary condition rule:

$$\hat{\sigma}_{\varepsilon_i}^2 = \min\left\{T\hat{\sigma}_{\upsilon_i}^2, \frac{\hat{\sigma}_{\upsilon_i}^2}{(1-\hat{\rho}_i(1))^2}\right\}, \tag{12}$$

where $\hat{\rho}_i(1) = \hat{\rho}_{i,1}(1) + \ldots + \hat{\rho}_{i,p_i}(1)$ denotes the autoregressive polynomial evaluated at $L=1$. Third, the long-run variance estimate of the residuals in Equation (11), $\hat{\sigma}_{\upsilon_i}^2$, is obtained using a quadratic spectral window Heteroskedastic and Autocorrelation Consistent (HAC) estimator. Sul et al. (2005) report Monte Carlo simulation results that reveal that the new boundary condition rule to estimate $\hat{\sigma}^2_{\varepsilon_i}$ improves the size and power properties of the KPSS tests; see also Carrión-i-Silvestre and Sansó (2006).

Having consistently estimated $\hat{\sigma}^2_{\varepsilon_i}$, the panel stationarity test is calculated as the simple average of the individual univariate KPSS stationarity tests:

$$\widehat{LM}_{T,N,k}(\widehat{\omega_i}) = \frac{1}{N}\sum_{i=1}^{N}\eta_{i,T,k}(\widehat{\omega}_i), \quad (13)$$

which after a suitable standardization, using appropriate moments of the statistics associated to the models postulated in Equations (1) to (6), follows a standard normal limiting distribution:

$$Z_k(\widehat{\omega}_i) = \frac{\sqrt{N}\left(\widehat{LM}_{T,N,k}(\widehat{\omega}_i) - \bar{\xi}_k\right)}{\bar{\zeta}_k} \Rightarrow N(0,1) \quad (14)$$

where $\bar{\xi}_k = \frac{1}{N}\sum_{i=1}^{N}\xi_{i,k}$ and $\bar{\zeta}^2_k = \frac{1}{N}\sum_{i=1}^{N}\zeta^2_{i,k}$ denote the mean and variance required for standardization, respectively. The proof of the previous result can be found in Hadri (2000). Furthermore, Hadri and Rao (2008) show that in the presence of breaks, that is for the models in Equations (3) to (6), the individual means, $\xi_{i,k}$, and variances, $\zeta^2_{i,k}$, depend upon the relative position of the break in the sample or, in other words, $\xi_{i,k}$ and $\zeta^2_{i,k}$ are functions of $\hat{\omega}_i$; see Hadri and Rao (2008), Theorem 3.

The Hadri and Rao (2008) test critically relies on the assumption that the individual time series in the panel are independent from each other.[10] To allow for cross-section dependence, Hadri and Rao recommend employing an AR-based bootstrap method, the steps of which are as follows: First, to account for serial correlation Equation (11) is estimated, and the resulting residuals (centred around zero) are denoted $\hat{\upsilon}_{it}$. Second, following Maddala and Wu (1999), the residuals $\hat{\upsilon}_{it}$ are re-sampled with replacement with the cross-section index fixed, so that their cross-correlation structure is preserved; the resulting bootstrap innovations are denoted $\hat{\upsilon}^*_{it}$. Third, $\hat{\varepsilon}^*_{it}$ is generated using the following mechanism:

$$\hat{\varepsilon}^*_{it} = \hat{\rho}_{i,1}\hat{\varepsilon}^*_{i,t-1} + \ldots + \hat{\rho}_{i,p_i}\hat{\varepsilon}^*_{i,t-p_i} + \upsilon^*_t, \quad (15)$$

[10] Monte Carlo simulation results by Giulietti et al. (2009) indicate that, even for relatively large *T* and *N*, the Hadri tests (with no serial correlation and no structural breaks) suffer from severe size distortions in the presence of cross section dependence. Using the bootstrap method results in statistics that are approximately correctly sized.

where $\hat{\rho}_{i,1},...,\hat{\rho}_{i,p_i}$ are the corresponding OLS coefficient estimates from the fitted AR model in (11). To ensure that the bootstrap samples, $\hat{\varepsilon}_{it}^{*}$, generated by (15) are stationary processes, we generate a larger number of $\hat{\varepsilon}_{it}^{*}$, let us say $T+Q$ values, and then discard the first Q values. This strategy also offers the advantage that the method used to obtain the initial values of $\hat{\varepsilon}_{it}^{*}$ becomes unimportant, and so one might as well use zeros for the initial values; see Chang (2004: footnote 6). For our purposes, we choose $Q=40$. Fourth, the bootstrap samples of y_{it}, denoted y_{it}^{*}, are calculated by adding $\hat{\varepsilon}_{it}^{*}$ to the deterministic component of the corresponding chosen model, and the Hadri and Rao LM test statistic is calculated for each y_{it}^{*}. The four steps described earlier are repeated several times to derive the empirical distribution of the LM statistic, and then bootstrap p-values (or alternatively bootstrap critical values) may be obtained.

4 Empirical results

4.1 Data

We employ the commodity price index dataset constructed by Grilli and Yang (1988) for the period 1900-86, and extended to 2010 by Stephan Pfaffenzeller, for a total of 111 time observations; the sample period is thus seven years longer than that recently analysed by Ghoshray (2011).[11] The dataset consists of price information on 24 commodities that account for (approximately) 54 per cent of all non-fuel commodities traded in the world in the period 1977-79; see Grilli and Yang (1988: footnote 2). These commodities are: aluminium, banana, beef, cocoa, coffee, copper, cotton, hides, jute, lamb, lead, maize, palm oil, rice, rubber, silver, sugar, tea, timber, tin, tobacco, wheat, wool and zinc. We use these commodities to conform four balanced panels of data: (i) Food commodities, which comprises banana, beef, cocoa, coffee, lamb, maize, palm oil, rice, sugar, tea and wheat; (ii) non-food commodities, which consists of cotton, hides, jute, rubber, timber, tobacco and wool; (iii) metals, which includes aluminium, copper, lead, silver, tin and zinc; and (iv) all commodities, which contains all 24 commodity price indices. Following the tradition of studies that have used the GY dataset, the price indices of the 24 commodities are deflated using a trade-weighted unit value index of the exports of manufactured commodities of five major industrial countries (France, Germany, Japan, United Kingdom and USA) to developing countries; see Pfaffenzeller et al. (2007). The resulting deflated series of commodity price indices are considered in logarithms.

As indicated in Section 2, in a recent contribution to the literature, Harvey et al. (2010) construct an entirely new dataset with commodity price information pre-dating 1900 up until 2005. The dataset is unbalanced because the commodity price series do not start in the same year. More specifically, the dataset consists of time-series observations for

[11] See Pfaffenzeller et al. (2007) for practical advice on how to update the GY commodity price indices, as well as for a full description of the data series and their sources.

beef, coal, gold, lamb, lead, sugar, wheat and wool that start in 1650; cotton in 1670; tea in 1673; rice and silver in 1687; coffee in 1709; tobacco in 1741; pig iron in 1782; cocoa, copper and hide in 1800; tin in 1808; nickel in 1840; zinc in 1853; oil in 1859; aluminium in 1872; and banana and jute in 1900. Thus, there are 20 commodities that are already included in the GY dataset. An important distinction between the GY dataset and the one collected by Harvey et al. (2010) relates to the construction of the price of manufactures. Indeed, while the former use manufacturing export unit value indexes for selected industrial countries, the latter employ value-added price deflators for manufacturing products; see Harvey et al. (2010) and the references therein for a discussion of the advantages and disadvantages of both measures.

It is interesting to observe that if we attempt to balance the Harvey et al. (2010) dataset, a requirement that is needed to apply the panel stationarity tests, then the relative commodity price series in the resulting balanced panel would begin in 1900, which is the same year when the GY dataset starts (since Banana and Jute prices are only available starting in that year).

4.2 Testing for cross-section dependence

We begin our empirical investigation with an analysis of cross-sectional dependence among commodity price indices. To do this, we calculate the Pesaran (2004) general diagnostic test for cross-section dependence in panels, denoted CD statistic. The procedure to implement this test works as follows. First, we fit an ADF(p) type regression model for each cross-section unit i separately, and take the resulting residuals as individual series $\hat{e}_{it}$. The importance of this initial stage is that it allows us to get rid of any serial correlation pattern in the individual time series i. Second, we compute the cross-correlation coefficient between the residuals of cross-section units i and j as:

$$\hat{\rho}_{ij} = \frac{\sum_{t=1}^{T} \hat{e}_{it}\hat{e}_{jt}}{\left(\sum_{t=1}^{T} \hat{e}_{it}^2\right)^{1/2}\left(\sum_{t=1}^{T} \hat{e}_{jt}^2\right)^{1/2}}. \tag{16}$$

Finally, we calculate the CD statistic as:

$$\mathrm{CD} = \sqrt{\frac{2T}{N(N-1)}}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\rho}_{ij}\right) \sim N(0,1). \tag{17}$$

Table 1 summarises the results of applying the CD statistic to the four panels defined earlier, using $p = 1, 2$ and 3 lags in the ADF(p) regressions. As can be seen from the table, the null hypothesis that commodity price innovations are cross-sectionally independent is strongly rejected for all commodity groupings and for all augmentation orders. The highest degree of cross-section dependence is found to be across all 24 commodities, followed by that observed in the group of metals. Thus, these results provide a justification for analysing commodity prices jointly, within a panel data framework, rather than as individual time series. In addition, they highlight the importance of allowing for the presence of cross-section dependence when performing inference to determine the time-series properties of commodity prices.

4.3 Testing for panel stationarity with structural breaks

Next, we turn our attention to testing for panel stationarity. The analysis starts off by identifying the presence of structural breaks (if any) in the prices of the 24 commodities included in the GY study. This issue has been examined by León and Soto (1997), Kellard and Wohar (2006), and Ghoshray (2011) using the sample periods 1900-92, 1900-98 and 1900-2003, respectively. Thus, in determining the position of the structural breaks, we carry out our estimations over four sample periods, namely the three that we already mentioned, as well as 1900-2010 (that is, the longest sample period currently available). This approach allows us to examine the effect of extending the sample period on the position of the break date, and also compare our results with those obtained in these three papers.

The results of determining the position of breaks over different sample periods are summarised in Table 2. There are two main aspects worth noticing in this table. First, the results reveal evidence of one structural break in all individual commodity prices. Indeed, notice that the model specifications that do not account for the presence of a structural break, i.e. the models in Equations (1) and (2), are never selected. This finding is in sharp contrast with the earlier work by León and Soto (1997), Kellard and Wohar (2006) and Ghoshray (2011), who find that there is no evidence of structural breaks in some commodities. Second, the results suggest that there are ten commodities (namely coffee, cocoa, beef, lamb, banana, palm oil, cotton, rubber, timber and aluminium) for which the position of the break does not change as the sample period is extended. More importantly, for the remaining fourteen commodities extending the sample period appears to have an effect on the estimated position of the break date; notice, in particular, the cases of tea, sugar, wool, tobacco, copper, lead and zinc for which extending the sample period over the years 2004-10 changes the position of the break date.

In what follows, we use Table 2 to compare our results on the position of break dates with the dates reported by León and Soto (1997), Kellard and Wohar (2006), and Ghoshray (2011). We focus on the commodities for which there is evidence of one structural break, and regard discrepancies of up to two years in the break date (in either direction) as negligible. Thus, in comparison to León and Soto (1997), who based their analysis on the period 1900-92, we find similar break dates for cocoa, beef, banana, palm oil, wool, tobacco, rubber, copper and aluminium. With respect to Kellard and Wohar (2006), who extend the sample period to include the years 1993-98, similar break dates are found for rice, palm oil and aluminium. Finally, after further extending the sample period to cover the 1999-2003 period, as in Ghoshray (2011), we find similar break dates for tea, hides and zinc. Overall, it appears that the results on break date determination are dependent on the econometric strategy used to identify the breaks, as well as on how the breaks are characterised, that is on whether we allow for a change in level, a change in slope, or both. Without a doubt, extending the sample period implies that new important events and/or changes in commodity markets are included in the analysis, and these in turn make previously chosen functional forms no longer appropriate.

Figure 1 plots the commodity price indices over the 1900-2010 period, along with the chosen broken-trend model that is fitted to each commodity.[12] The estimated trend components tell very diverse stories indeed. Commodities such as cocoa, rice, sugar, cotton, rubber, copper, aluminium and silver offer support for the PS hypothesis by exhibiting a negative trend both before and after the break. Partial support for the PS hypothesis is provided by some commodities where a negative trend is only observed either before the break (wheat, maize, and palm oil) or after the break (coffee, tea, banana, jute, wool, and hides). At the other end of the spectrum, lamb, tobacco, timber, tin and lead offer no support for the PS hypothesis, as they display a positive trend both before and after the break. Lastly, beef and zinc can be viewed as trendless series that show evidence of a one-time positive level shift after the break, which does not corroborate the PS hypothesis either. At this point, it could well be argued that while the Hadri and Rao (2008) procedure accounts for unknown structural breaks, it is limited insofar as only a single break is permitted for each individual in the panel. However, informal visual inspection of the residuals from the chosen models reveals no evidence of further structural breaks, apart from the presence of potential outlier observations.[13]

The residuals of the broken-trend models are then used to construct the univariate KPSS stationarity tests (with breaks) based on Equation (9). To correct for serial correlation we include $p = 10$ lags in Equation (11), and then determine the optimal number of lags using the GTS algorithm outlined earlier.

Table 3 reports the Hadri and Rao panel stationary test statistics for the panels described above. At this point it is worth recalling that given that inference cannot be based on the standard normal density, due to the presence of cross-section dependence among the commodity price indices, we implement the AR-based bootstrap procedure outlined in the previous section (based on 2,000 bootstrap replications).[14] The results reveal that when analysing commodity prices within a panel context, the null hypothesis that they are jointly stationary around a (broken) trend cannot be rejected at traditional significance levels. This finding supports the view that commodity prices maintain a long-run cointegration relationship with the price of manufactured goods or, to put it in the terminology of Leybourne et al. (1994), there exists comovement between commodity prices and the price of manufactures. The idea of comovement requires that the response of both commodity prices and manufactured goods prices to common macroeconomic determinants (such as aggregate demand, inflation, exchange rates, and interest rates) be qualitatively the same, i.e. the partial derivatives must have the same sign.

---

12 León and Soto (1997), Kellard and Wohar (2006) and Ghoshray (2011) report plots of the commodity price series, but not of the (broken) trend components that are estimated.

13 Bai and Perron (1998) provide a framework for estimating and testing linear regression models with multiple structural breaks that occur at unknown dates. However, the Bai-Perron methodology is not implemented here because it does not permit the use of trending regressors, which is of particular relevance when assessing the validity of the PS hypothesis. Bai and Perron (2003) and Camarero et al. (2010) illustrate the use of the Bai-Perron methodology for modelling interest rates in the presence of multiple abrupt changes in the mean of the series in a univariate and panel data contexts, respectively.

14 Indeed, notice that in the case of food commodities, incorrectly performing inference based on the upper tail of the standard normal distribution leads to rejection of the null of panel stationarity at the five per cent significance level.

4.4 Policy considerations

From a statistical point of view, the finding that commodity prices exhibit TS (as opposed to DS) behaviour is important because the effect of a shock will be transitory (as opposed to permanent). From an economic point of view, as discussed by, among others, Deaton and Miller (1996), the key issue is the choice of the appropriate policy response to commodity price shocks: stabilization when the shock is transitory, or adjustment when it is permanent. Empirical evidence on the response of developing countries to trade shocks is quite interesting indeed. For instance, Collier and Gunning (1999: 1-63) compare 23 case studies from a sample of countries in Africa, Asia and Latin America that experienced different sorts of commodity shocks, and conclude that, generally speaking, although the countries under investigation did not fail to save a large share of the windfalls (as the permanent income theory of consumption predicts), they did tend to invest these savings badly. However, in practice the implementation of stabilization policies can be made much more difficult because of uncertainties surrounding the magnitude and (perhaps more importantly) the duration of the price booms or busts. Indeed, as indicated by Deaton and Miller (1996), the high degree of persistence of commodity prices might complicate macroeconomic management. On the one hand, when times are good countries may need to accumulate reserves over prolonged periods of time, and this strategy could turn out to be expensive and possibly not even feasible (or sustainable) from a political point of view. On the other hand, when times are bad countries may face limitations in their ability to borrow to finance their consumption levels.

Taking the above policy issues into consideration, it is of some interest to measure the persistence of commodity price shocks after structural breaks are accommodated in the analysis. For this, we use half-life estimates based on the Pesaran and Shin (1998) generalised impulse response (GIR) functions that result from estimating VAR models.[15] GIR functions, unlike standard impulse response functions based on a Cholesky decomposition, offer the advantage of being invariant to the way shocks in the underlying VAR model are orthogonalised. The empirical analysis starts off by estimating VAR models for each primary commodity group under consideration, namely food, non-food and metals, where the VAR models themselves consist of the residuals that result from estimating the chosen break-type model. It should be recalled that after accounting for structural breaks (and cross-section dependence), the resulting commodity price series turn out to be jointly stationary and therefore suitable for modelling in a VAR framework.

An important initial stage in the analysis is the selection of the optimal order of the VAR models, which involves selecting an order high enough such that one can be reasonably confident that the optimal order will not exceed it. Bearing in mind that the sample size ($T = 110$ observations) might become small relative to the number of lagged level variables included in each VAR model, we set four lags as the maximum order of the models, and use the Schwarz information criterion to select the optimal order. This criterion selects the optimal number of lags to be equal to one.[16] Then, the

[15] Seong et al. (2006) recommend using impulse response functions to estimate the half-life of a shock. The traditional formula to estimate the half-life of a shock $-(\ln(2) \div \ln(\delta))$, where δ refers to the value of the autoregressive parameter, is only applicable in the case of simple AR(1) models.

[16] The Akaike information criterion also selects the same optimal lag order.

underlying VAR(1) models can be used to compute the associated GIR functions, which plot the time profile of the effect of an own unit shock in a commodity price (measured by one standard deviation). Lastly, the resulting statistically significant lag weights are normalised so that they add up to one, and the half-life is calculated as the number of years required for 50 per cent (or the first half) of the adjustment to take place.

The results of the persistence analysis, reported in Table 4, reveal that the estimated half-life to own-price shocks is lower than the typical half-lives estimated in previous studies using other methodologies. For example, the half-life for food and non-food commodities is two years (except for tea and hides where it is one year). More persistence is found in the group of metals where the average half-life is just over three years, varying between two years (for aluminium and zinc) and five years (for silver).[17] It should be noticed that the relatively short-lived persistence of shocks is achieved after removing broken-time deterministic trends from the underlying commodity price series, and that this result is consistent with the findings in the econometrics literature that relate unaccounted structural breaks with spurious non-stationarity, and therefore low rates for mean reversion; see Perron (1989).

The finding that commodity price shocks exhibit low persistence rates, after accommodating structural breaks, suggests that in the short run there appears to be scope for the utilization of stabilization mechanisms in order to smooth the path of export revenues in developing countries. Needless to say, the implementation of stabilization mechanisms to a particular country should be based on a careful examination of the specific products and export markets on which the country is dependent. In the long run, the issue at the core of the discussion is that most developing countries rely heavily on export revenue from the production of few primary commodities. The evidence reported in this study indicates that all the 24 commodity prices under analysis have exhibited abrupt one-time changes of one form or another. Thus, it is in the interest of developing countries to develop strategies that help them achieve a diversified production structure, so that the impact of future commodity price shocks is cushioned. Related to this point, Singer (1999), among other authors, reiterates the importance for developing countries to diversify their exports by moving into the production of manufactures, in particular of those that are somewhat technologically complex.

5 Concluding remarks

In this study we have examined the validity of the Prebisch and Singer hypothesis of a long-run negative trend in the terms of trade between primary commodities and manufactures. For this, we use an up to date version of the widely used commodity price dataset assembled by Grilli and Yang (1988), and employ a panel stationarity testing procedure that addresses both structural breaks and cross-sectional dependence. This modelling approach differs from the one that has been used in the existing literature, which is based on univariate non-stationarity tests applied to individual commodity prices.

---

[17] Collier and Gunning (1999) indicate that in a sample of 19 positive shocks, in two out of three cases the duration is about 3-8 years.

The empirical analysis starts off by confirming the presence of cross-section dependence among commodity prices. This finding supports the view that when dealing with commodity prices it is not appropriate to assume that they are independent from each other, due to the existence of market linkages. Also, it provides a justification for treating commodity prices as a panel of data, which is advantageous since the power of statistical tests increases with the number of cross sections in the panel. The analysis proceeds by revealing that all 24 commodity prices exhibit a one-time structural break, which differs across commodities. In fourteen out of 24 cases the position of the break varies according to the time span of data that is used, while in the remaining ten cases the estimated position of the break does not change when the sample period is extended by including more recent observations.

The results of the panel stationarity tests suggest that commodity prices are jointly stationary after accommodating one-time structural breaks and cross-section dependence. This finding can be alternatively viewed as implying that commodity prices and manufactured goods prices must be linked by a long-run equilibrium relationship or, to put it in other words, that there is comovement between them. Underlying the existence of comovement, there are common macroeconomic variables that influence both prices in the same way.

Broadly speaking, support for the Prebisch and Singer hypothesis is mixed. The strongest evidence in favour is encountered for commodities such as cocoa, rice, sugar, cotton, rubber, copper, aluminium and silver, which display a negative trend both before and after the break. the remaining commodities provide either partial support (as some commodity prices exhibit a negative trend only before or after the break) or no support whatsoever for the hypothesis. The results also indicate that once the breaks are removed from the underlying series, the persistence of commodity price shocks (as measured by their half-life) is shorter than that obtained in other studies using alternative methodologies.

From an economic policy standpoint, our results support the adoption of prudent macroeconomic policies. On the one hand, finding that all 24 commodity prices exhibit abrupt structural breaks of one form or another, support the view that, in the long run, it is in the interest of developing countries to implement policy measures aimed at diversifying their production structure, so that their dependence on few commodities as a source of foreign exchange is reduced. On the other hand, the relatively low rates of persistence of commodity price shocks suggest that, in the short run, there is scope for developing countries to design and use stabilization mechanisms in response to trade shocks.

References

Agénor, P.-R. and P. J. Montiel (1996). *Development Macroeconomics*. Princeton, NJ: Princeton University Press.

Ardeni, P. G. and B. Wright (1992). ‘The Prebisch-Singer hypothesis: A reappraisal independent of stationarity hypotheses’. *Economic Journal*, 102: 803-812.

Bai, J. and P. Perron (1998). ‘Estimating and testing models with multiple structural changes’. *Econometrica*, 66: 47-78.

Bai, J. and P. Perron (2003). 'Computation and analysis of multiple structural change models'. *Journal of Applied Econometrics*, 18: 1-22.

Balassa, B. (1989). 'Outward orientation'. In H. Chenery and T. N. Srinivasan (eds), *Handbook of Development Economics, Volume II*. Amsterdam: Elsevier Science Publishers.

Bleaney, M. and D. Greenaway (1993). 'Long-run trends in the relative price of primary commodities and in the terms of trade of developing countries'. *Oxford Economic Papers*, 45: 349-363.

Bowley, A. L. (1903). 'The prices of imports and exports of the United Kingdom and Germany'. *Economic Journal*, 13: 628-632.

Breitung, J. and M. H. Pesaran (2008). 'Unit roots and cointegration in panels'. In L. Mátyás and P. Sevestre (eds), *The Econometrics of Panel Data*. Berlin: Springer-Verlag.

Camarero, M., Carrión-i-Silvestre, J., and C. Tamarit (2010). 'Does real interest rate parity hold for OECD countries? New evidence using panel stationarity tests with cross-section dependence and structural breaks'. *Scottish Journal of Political Economy*, 57: 568-590.

Campbell, J. Y. and P. Perron (1991). 'Pitfalls and opportunities: What macroeconomists should know about unit roots'. *NBER Macroeconomics Annual*, 6: 141-201.

Caner, M. and L. Kilian (2001). 'Size distortions of tests of the null hypothesis of stationarity: Evidence and implications for the PPP debate'. *Journal of International Money and Finance*, 20: 639-657.

Carrión-i-Silvestre, J., T. Del Barrio, and E. López-Bazo (2005). 'Breaking the panels: An application to the GDP per capita'. *The Econometrics Journal*, 8: 159-175.

Carrión-i-Silvestre, J. and A. Sansó (2006). 'A guide to the computation of stationarity tests'. *Empirical Economics*, 31: 433-448.

Chang, Y. (2004). 'Bootstrap unit root tests in panels with cross-sectional dependency'. *Journal of Econometrics*, 120: 263-293.

Chow, G. C. (1960). 'Tests of equality between sets of coefficients in two linear regressions'. *Econometrica*, 28: 591-605.

Collier, P. and J. W. Gunning (1999). 'Trade shocks: Theory and evidence'. In P. Collier, J. W. Gunning, and Associates (eds), *Trade Shocks in Developing Countries, Volume 1: Africa*. Oxford: Oxford University Press.

Cuddington, J. T. (1992). 'Long-run trends in 26 primary commodity prices. A disaggregated look at the Prebisch-Singer hypothesis'. *Journal of Development Economics*, 39: 207-227.

Cuddington, J. T. and C. M. Urzua (1989). 'Trends and cycles in the net barter terms of trade: A new approach'. *Economic Journal*, 99: 426-442.

Deaton, A. and R. Miller (1996). 'International commodity prices, macroeconomic performance and politics in Sub-Saharan Africa'. *Journal of African Economies*, 5 (Supplement): 99-191.

Dickey, D. A. and W. A. Fuller (1979). 'Distribution of the estimators for autoregressive time series with a unit root'. *Journal of the American Statistical Association*, 74: 427-431.

Engle, R. F. and Granger, C. W. J. (1987). 'Cointegration and error correction: representation, estimation and testing'. *Econometrica*, 55: 251-276.

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

Giulietti, M., J. Otero, and J. Smith (2009). 'Testing for stationarity in heterogeneous panel data in the presence of cross section dependence'. *Journal of Statistical Computational and Simulation*, 79: 195-203.

Grilli, E. R. and M. C. Yang (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.

Hadri, K. (2000). 'Testing for stationarity in heterogeneous panels'. *The Econometrics Journal*, 3: 148-161.

Hadri, K. and Y. Rao (2008). 'Panel stationarity test with structural breaks'. *Oxford Bulletin of Economics and Statistics*, 70: 245-269.

von Hagen, J. (1989). 'Relative commodity prices and cointegration'. *Journal of Business and Economic Statistics*, 7: 497-503.

Hall, A. (1994). 'Testing for a unit root in time series with pretest data-based model selection'. *Journal of Business and Economic Statistics*, 12: 461-470.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Harvey, D. I., S. J. Leybourne, and A. M. R. Taylor (2007). 'A simple, robust and powerful test of the trend hypothesis'. *Journal of Econometrics*, 141: 1302-1330.

Harvey, D. I., S. J. Leybourne, and A. M. R. Taylor (2009). 'Simple, robust and powerful tests of the breaking trend hypothesis'. *Econometric Theory*, 25: 995-1029.

Harvey, D. I., Kellard, N. M, Madsen, J. B, and M. E. Wohar (2010). 'The Prebisch-Singer hypothesis: Four centuries of evidence'. *The Review of Economics and Statistics*, 92: 367-377.

Helg, R. (1991). 'A note on the stationarity of the primary commodities relative price index'. *Economics Letters*, 36: 55-60.

Im, K., M. H. Pesaran, and Y. Shin (2003). 'Testing for unit roots in heterogeneous panels'. Journal of Econometrics, 115: 53-74.

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

Kellard, N. M and M. E. Wohar (2006). 'On the prevalence of trends in primary commodity prices'. *Journal of Development Economics*, 79: 146-167.

Keynes, J. M. (1912). Official papers. Tables showing for each of the years 1900-11 the estimated value of the imports and exports of the United Kingdom at the prices prevailing in 1900. *Economic Journal*, 22: 630-631.

Kim, T.-H., S. Pfaffenzeller, T. Rayner, and P. Newbold (2003). 'Testing for linear trend, with application to relative primary commodity prices'. *Journal of Time Series Analysis*, 24: 539-551.

Kindleberger, C. P. (1958). 'The terms of trade and economic development'. *Review of Economics and Statistics*, 40: 72-85.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (1992). 'Testing the null hypothesis of stationarity against the alternative of a unit root'. *Journal of Econometrics*, 54: 159-178.

Lee, J. and M. C. Strazicich (2003). 'Minimum Lagrange multiplier unit root test with two structural breaks'. *Review of Economics and Statistics*, 85: 1082-1089.

Leon, J. and R. Soto (1997). 'Structural change and long-run trend in commodity prices'. *Journal of International Development*, 9: 347-366.

Lewis, W. A. (1954). 'Economic development with unlimited supplies of labour'. The Manchester School, 22: 139-191.

Leybourne, S. J., Lloyd, T. A. and G. V. Reed (1994). 'The excess comovement of commodity prices revisited'. *World Development*, 22: 1747-1758.

Lumsdaine, R. L. and D. Papell (1997). 'Multiple trend breaks and the unit root hypothesis'. *Review of Economics and Statistics*, 79: 212-218.

Maddala, G. S. and S. Wu (1999). 'A comparative study of unit root tests with panel data and a new simple test'. *Oxford Bulletin of Economics and Statistics*, 61: 631-652.

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

Perron, P. (1989). 'The great crash, the oil price shock and the unit root hypothesis'. *Econometrica*, 57: 1361-1401.

Perron, P. (1990). 'Testing for a unit root in a time series with a changing mean'. *Journal of Business and Economic Statistics*, 8: 153-162.

Pesaran, M. H. (2004). 'General diagnostic tests for cross section dependence in panel'. *Cambridge Working Papers in Economics*, No. 435.

Pesaran, M. H. and Y. Shin (1998). 'Generalized impulse response analysis in linear multivariate models'. *Economics Letters*, 58: 17-29.

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

Powell, A. (1991). 'Commodity and developing countries terms of trade: What does the long run show?'. *Economic Journal*, 101: 1485-1496.

Prebisch, R. (1949). *El Desarrollo Económico de la América Latina y Sus Principales Problemas*. New York: United Nations Economic and Social Council.

Prebisch, R. (1950). *The Economic Development of Latin America and its Principal Problems*. Economic Comission for Latin America, United Nations, Department of Economic Affairs, New York: Lake Success.

Reinhart, C. M. and P. Wickham (1994). ‘Commodity prices: Cyclical weakness or secular decline?’. *IMF Staff Papers*, 41: 175-213.

Rostow, W. W. (1950). ‘The terms of trade in theory and practice’. *Economic History Review*, New Series, 3: 1-20.

Sapsford, D. (1985). ‘The statistical debate on the net barter terms of trade between primary commodities and manufactures: A comment and some additional evidence’. *Economic Journal*, 95: 781-788.

Sarkar, P. (1986). ‘The terms of trade experience of Britain in the Nineteenth Century’. *Journal of Development Studies*, 23: 20-39.

Seong, B., A. M. Morshed, and S. K. Ahn (2006). ‘Additional sources of bias in half-life estimation’. *Computational Statistics and Data Analysis*, 51: 2056-2064.

Singer, H. W. (1950). ‘The distribution of gains between investing and borrowing countries’. *American Economic Review*, 40: 473-485.

Singer, H. W. (1994). ‘Comments on Raúl Prebisch: The continuing quest’. In E. V. Iglesias (ed), *The Legacy of Raul Prebisch*. Washington DC: Inter-American Development Bank.

Singer, H. W. (1999). ‘Beyond terms of trade – convergence and divergence’. *Journal of International Development*, 11: 911-916.

Spraos, J. (1980). ‘The statistical debate on the net barter terms of trade between primary commodities and manufactures’. *Economic Journal*, 90: 107-128.

Sul, D., P. Phillips, and C. Choi (2005). ‘Prewhitening bias in HAC estimation’. *Oxford Bulletin of Economics and Statistics*, 67: 517-546.

United Nations (1949). ‘Relative Prices of Exports and Imports of Under-Developed Countries: A Study of Post-War Terms of Trade Between Under-Developed and Industrialized Countries’. United Nations, Department of Economic Affairs, New York: Lake Success.

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

Zivot, E. and D. W. K. Andrews (1992). ‘Further evidence on the great crash, the oil price shock, and the unit root hyothesis’. *Journal of Business and Economic Statistics*, 10: 251-270.

Table 1: CD statistic of cross-sectional independence

Panel	ADF(1)		ADF(2)		ADF(3)	
	Statistic	p -value	Statistic	p -value	Statistic	p -value
Food	11.980	[0.000]	11.693	[0.000]	11.700	[0.000]
Non-food	11.402	[0.000]	11.489	[0.000]	11.568	[0.000]
Metals	14.769	[0.000]	15.419	[0.000]	15.427	[0.000]
All	31.086	[0.000]	31.886	[0.000]	32.153	[0.000]

Notes: Food commodities include Banana, Beef, Cocoa, Coffee, Lamb, Maize, Palm Oil, Rice, Sugar, Tea and Wheat. Non-food commodities consist of Cotton, Hides, Jute, Rubber, Timber, Tobacco and Wool. Metals comprise Aluminium, Copper, Lead, Silver, Tin and Zinc. The Pesaran (2004) CD statistic is based on the cross-correlation of the residuals that result from estimating p th-order ADF type regressions (including constant and trend) for each of the individual commodities that conform a panel. The CD statistic follows a standard normal distribution under the null hypothesis of cross-sectional independence.

Source: Authors' computations.

Table 2: Estimated models and structural breaks

Commodity	1900-1992		1900-1998		1900-2003		1900-2010	
	Model	Date	Model	Date	Model	Date	Model	Date
Coffee	6	1950	6	1950	6	1950	6	1950
Cocoa	6	1947	6	1947	6	1947	6	1947
Tea	6	1954	6	1954	6	1954	6	1968
Rice	6	1973	4	1982	6	1982	4	1982
Wheat	6	1973	6	1921	4	1986	6	1987
Maize	5	1975	4	1986	4	1986	6	1986
Sugar	6	1972	6	1972	6	1972	4	1925
Beef	4	1959	3	1959	3	1959	3	1959
Lamb	4	1947	4	1947	4	1947	4	1947
Banana	6	1926	6	1926	6	1926	6	1926
Palm oil	4	1986	6	1986	4	1986	6	1986
Cotton	6	1946	6	1946	6	1946	6	1946
Jute	6	1947	5	1966	6	1947	6	1973
Wool	5	1952	5	1952	5	1952	5	1942
Hides	6	1921	6	1952	6	1952	6	1921
Tobacco	6	1918	6	1918	6	1918	6	1919
Rubber	4	1918	4	1918	4	1918	4	1918
Timber	6	1921	6	1921	6	1921	6	1921
Copper	4	1953	4	1953	4	1953	4	2006
Aluminium	6	1942	6	1942	6	1942	6	1942
Tin	6	1977	6	1977	6	1986	4	1986
Silver	4	1967	6	1974	6	1974	4	1967
Lead	6	1947	6	1947	6	1947	6	1982
Zinc	6	1918	6	1918	6	1918	3	2006

Note: The columns labelled 'Model' indicate the chosen model specifications, as postulated in Equations. (1) to (6).

Source: Authors' computations.

Table 3: Panel stationarity results for relative commodity prices with breaks

Panels	Statistic	p -value
Food	1.658	[0.151]
Non-food	0.867	[0.136]
Metals	-0.440	[0.769]
All commodities	0.765	[0.364]

Notes: The footnote in Table 1 lists the commodities included in each panel. In constructing the individual KPSS statistics, we set $p_i^{\max} = 10$ in Equation (11), and select the optimal number using the GTS algorithm with a significance level of 10 per cent. Bootstrap p -values are based on 2000 replications.

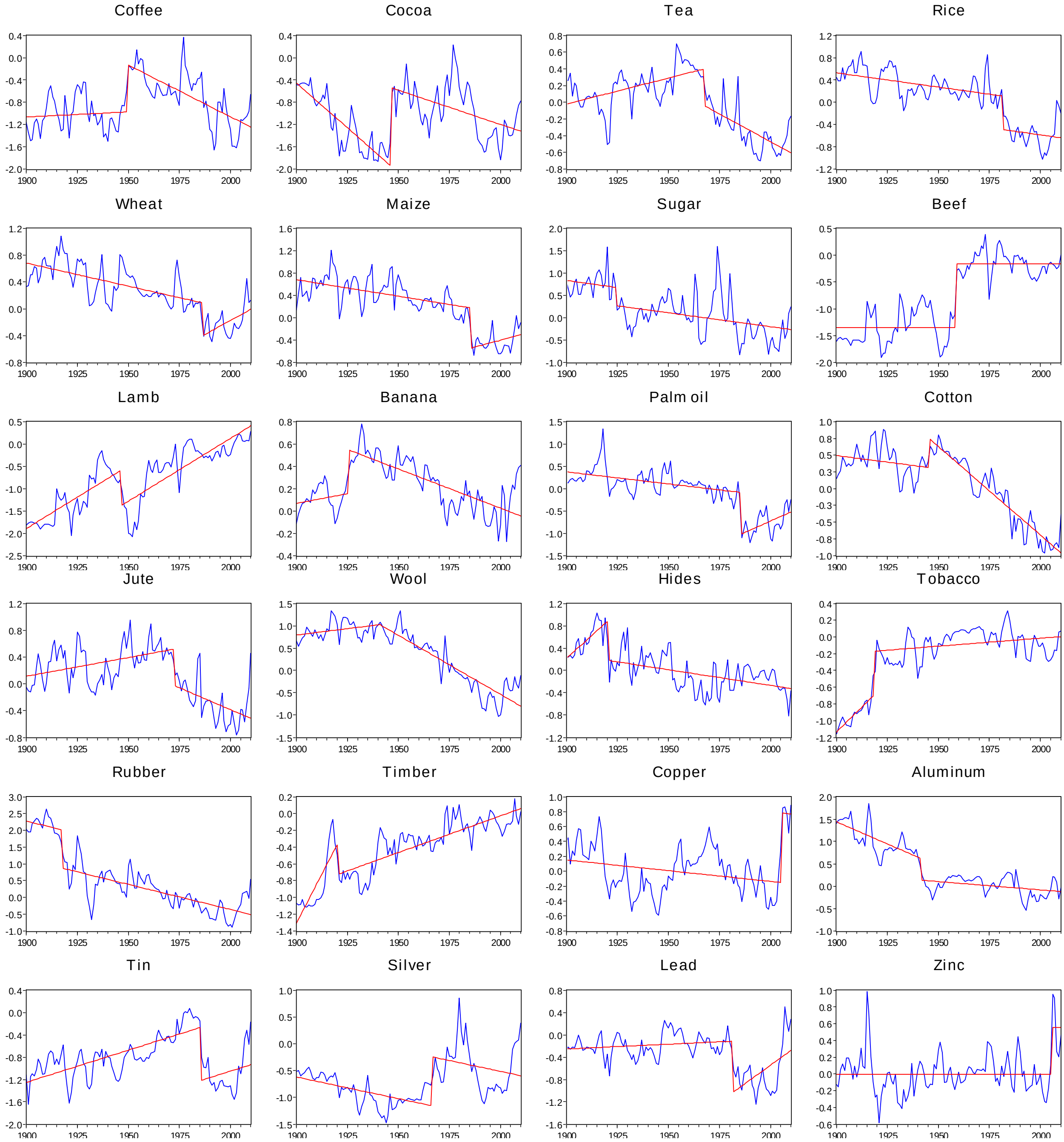
Source: Authors' computations.

Table 4: Half-life estimates (in years) from GIR functions

Food	Half-life	Non-food	Half-life	Metals	Half-life
Coffee	2	Cotton	2	Copper	4
Cocoa	2	Jute	2	Aluminium	2
Tea	1	Wool	2	Tin	3
Rice	2	Hides	1	Silver	5
Wheat	2	Tobacco	2	Lead	3
Maize	2	Rubber	2	Zinc	2
Sugar	2	Timber	2		
Beef	2				
Lamb	2				
Banana	2				
Palm oil	2				

Source: Authors' computations.

Figure 1: Plots of the commodity price series and fitted broken trend



Source: Authors' computations.