Unit roots, flexible trends and the Prebisch-Singer hypothesis*

Diego Winkelried

Universidad del Pacífico (Lima, Perú)
winkelried_dm@up.edu.pe

This version: November 6, 2015

Abstract

This paper studies the dynamic properties of relative commodity prices, especially the Prebisch-Singer hypothesis on the secular decline in these series, using a new family of unit root tests that is based on the Fourier approximation to the underlying trend of the data. The approximation controls for low-frequency variations such as structural breaks, or such as the long swings induced by hypothesized super cycles in the data. Relative to the extant literature, we find considerably more evidence in favor of trend stationarity in relative commodity prices, and relatively limited support for the Prebisch-Singer hypothesis.

JEL Classification : C22, O13.
Keywords : Primary commodity prices, unit roots, long swings, super cycles.

*I would like to thank Mariella Marmanillo for outstanding research assistance. I am also indebted to Pablo Lavado, Nelson Ramírez, Gabriel Rodríguez, Marco Vega and seminar participants at the XXXII Economists Meeting of the Central Reserve Bank of Peru for useful comments. I alone am responsible for any errors that may remain and for the views expressed in the paper.

1 Introduction

The long-run dynamics of primary commodity prices has been a subject of major interest in empirical development economics. In particular, a large body of literature has focused on studying whether the so-called *Prebisch-Singer hypothesis*, i.e. that the relative prices of primary commodities in terms of manufactures are driven by a secular downward trend (cf. Prebisch, 1950; Singer, 1950), holds in practice. Its importance stems from its key policy implications for commodity-exporting countries, as the adequacy of export diversification or industrialization efforts depends greatly on whether the terms of trade are expected to sustainably decline in the future (see Cuddington et al., 2008, for a comprehensive survey).

Since the relatively early contributions of Cuddington and Urzua (1989), Cuddington (1992) and Bleaney and Greenaway (1993), the bulk of the literature has examined the Prebisch-Singer hypothesis through the lens of unit root tests (see, *inter alia*, Ghoshray, 2011, for a recent review). This is so because ignoring a unit root when it is present in a given time series can have profound distortionary effects on the inferences made about its low frequency (i.e., long-run) behavior. Unit roots tests aid the researcher to decide whether a particular relative price should be modeled as a trend stationary or as difference stationary process, possibly subject to structural breaks. Then, evidence of a negative time slope (estimated from an equation in levels, the trend stationary case) or drift (estimated from an equation in differences, the unit root case) is taken as supportive of the Prebisch-Singer hypothesis.

Although usually consistent, unit root tests are known to have low power in finite samples, i.e the (false) null hypothesis of a unit root may not be rejected against a stationary but "persistent" alternative (see Choi, 2015, ch. 3, for a textbook account). In other words, such tests may fail to reject the unit root hypothesis not because of the merits of the null hypothesis, but because of the inadequacy of the alternative hypothesis. In particular, when the alternative model does not take into account certain components of the time series that affect its low frequency variation (where the "unit root" lies), a difference stationary model may result in a better characterization of the data. The leading example is that of unmodeled structural breaks: a covariance stationary series with a level shift or a broken trend may appear as a unit root process, if the structural change is ignored when testing. The same is truth for some forms of neglected nonlinearities.

There is little doubt that the long data span required to correctly assess a long-run phenomenon such as the Prebisch-Singer hypothesis, makes the presence of instabilities, structural breaks or regime shifts very likely. Even early studies, such as Grilli and Yang (1988) when released their famous dataset, acknowledge the possibility that the trends of relative commodity prices may be changing over time. Regarding unit root testing, even though results are somehow mixed, the latest studies that allow for multiple structural breaks under the alternative hypothesis (e.g. León and Soto, 1997; Zanias, 2005) tend to find more evidence *against* nonstationarity, relative to earlier contributions that allow for, at most, a single structural break (e.g. Cuddington and Urzua, 1989).[1] Analogously, the support for stationary models seems to be stronger after controlling for certain nonlinearities (e.g. Balagtas and Holt, 2009; Harvey et al., 2011; Enders and Holt, 2012).

What remains an open question is the nature of the underlying changes in primary commodity prices, i.e. whether they are smooth and manifest themselves gradually, or they are sharp showing their effects very fast. The latest contributions on testing unit roots in commodity prices (*inter alia*, Kellard and Wohar, 2006; Ghoshray, 2011; Arezki et al., 2014) use state-of-the-art techniques that allow for at most two structural (sharp) breaks of unknown dates. Yet, a revealing finding in Yamada and Yoon (2014) is that if the underlying trend of relative primary commodity prices is to be modeled as a broken linear trend, then the number of slope changes to be considered should be, at least, moderate and surely greater than two. Thus, despite the efforts of allowing for greater flexibility of the alternative model, the unit root tests considered in the latest studies may still have low power, leading to potentially erroneous conclusions.

[1] In a related analysis, Mariscal and Powell (2014) show that, after correcting the historical data for various level shifts, the logarithm of an aggregate commodity prices index cointegrates with the logarithm of a manufacturing unit value index. The cointegrating vector is $(1, -1)$ which indicates that the relative commodity price index is stationary.

Often economic theory is silent about the features, other that the absence of a unit root, that should be captured by the alternative model. As suggested, low power may be a simple manifestation of our ignorance about a more suitable characterization of the data. In the case of commodity prices, however, the so-called *Super Cycle hypothesis* provides a clear-cut conceptual basis for an alternative model worth exploring (see, for instance, Cuddington and Jerrett, 2008; Erten and Ocampo, 2012). Briefly, this hypothesis states that primary commodity prices display "long swings", a demand-driven phenomenon associated to the surge of large industrial economies, that may last for decades. In recorded modern history, prolonged expansive phases in commodity prices have occurred during the American industrialization by the turn of the 19th century, the post-war reconstruction of Europe and Japan by the mid 20th century, and the Chinese vigorous economic expansion of the late 20th century.

Cuddington and Jerrett (2008) and Erten and Ocampo (2012) suggest that super cycles in primary commodity prices are associated to an unobserved time series component lasting from 20 to 70 years (say, upswings of 10 to 35 years). Given the sample sizes on commodity prices available for empirical work, spanning few centuries, unit root tests may easily confound a long-lasting, but stationary, cycle with a unit root (see also Harvey et al., 2010). Furthermore, Prodan (2008) shows that modeling such cyclical behavior through the introduction of dummy variables may be inadequate, thereby reinforcing the suspicion that the unit root tests considered so far may lack power.[2]

The purpose of this paper is to reexamine the time series properties of relative commodity prices by considering the new family of unit root tests advanced in Enders and Lee (2012a,b). These tests place a "flexible trend" under the alternative model that controls for the effects of components affecting the long-run dynamics of the series. The idea, which stems from Becker et al. (2004), is to use a parsimonious Fourier approximation that allows for considerable flexibility in modeling the underlying trend, which may be subject to smooth shifts or even sharp structural breaks. In general, the Fourier approximation would capture various forms of instability or nonlinearity (as a function of time) in the underlying trend, and simulation studies in Becker et al. (2004), Becker et al. (2006) and Enders and Lee (2012a,b) show that these procedures are notoriously more powerful than alternative unit-root break-testing methodologies.

The application of the "flexible trend" unit root tests to relative commodity prices seems natural, given our conjecture that the perceived instabilities in these prices can be rationalized as the long swings induced by a super cycle.[3] Indeed, when these tests are applied to the relative price series from the celebrated Grilli and Yang (1988) dataset spanning the 20th century and the 2000s, it is found that out of 24 commodities, the unit root hypothesis cannot be categorically rejected in *at most* four cases. These findings provide, to the best of our knowledge, much stronger evidence against unit roots than in the extant literature. Moreover, these conclusions remain almost unaltered (and are, in fact, slightly more supportive of stationarity) after removing the volatile data of the 2000s.

The flexible trends are estimated as a by-product of the unit root tests, giving us the opportunity to reassess the Prebisch-Singer hypothesis (i.e., a negative long-run trend) from this new perspective. We find the flexible trends to follow a notorious (super)cyclical pattern in most of the cases. Thus, relatively little support to the Prebisch-Singer hypothesis is found when interpreted as a global prediction, i.e. that holds for the entire sample. The support is stronger when we treat this as a local prediction, since we do find several episodes of statistically significant and negative slopes in the flexible trend. As Yamada and Yoon (2014) put it, the Prebisch-Singer hypothesis holds *sometimes*.

The rest of the paper is organized as follows. Section 2 describes the flexible trend unit root tests. Section 3

[2] Prodan (2008) argues that the size and power of break-detecting procedures can be severely distorted in these situations. For instance, consider panels (c) and (e) in Figure 1 below. The discontinuous thick lines can be though of as the dummy-variable approximation to the smooth, cyclical lines. In both cases, the cycle period is about 70, and the cycle is approximated by either two or three offsetting level shifts, whose dates are rather difficult to locate in practice.

[3] Becker et al. (2006) successfully applied similar procedures to various real exchange rates, whose "long swings" are much more documented in the international economics literature. Similarly, Enders and Lee (2012b) study the behavior of the real interest rate in the US, also known to display a persistent cyclical behavior, whereas Jones and Enders (2012) examine the real price of crude oil. An important antecedent to our study is Enders and Holt (2012), who also study the dynamics of commodity prices.

presents the results of these tests applied to the Grilli and Yang dataset, and compares them to other findings in the literature, in particular in Kellard and Wohar (2006), Balagtas and Holt (2009) and Ghoshray (2011). Section 4 uses the models suggested by the unit root tests to compute the time-varying slopes of the trend functions, in order to assess the Prebisch-Singer hypothesis in our framework. Section 5 concludes and gives suggestions for future research.

2 Unit root tests with flexible trends

Consider the time series model

$$y_t = \tau(t) + u_t, \qquad \text{where} \qquad \Phi(L)u_t = \varepsilon_t\,, \tag{1}$$

with $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p - \phi_{p+1} L^{p+1}$ being a polynomial in the lag operator, $L^k y_t = y_{t-k}$. We refer to $\tau(t)$ as the *trend function* of y_t. The ultimate object of interest is the associated *slope*, defined as

$$\delta(t) = \tau(t) - \tau(t-1)\,. \tag{2}$$

Two possibilities to model $\delta(t)$ arise, depending on the properties of the roots of $\Phi(z) = 0$. First, if this polynomial contains no unit roots, i.e. the moduli of all roots are greater than one, then $u_t \sim I(0)$ and y_t is a trend stationary (TS) process. In this case, $\mathrm{E}(y_t) = \tau(t)$ and $\mathrm{E}(\Delta y_t) = \delta(t)$, and so the trend function can be estimated directly from the data in levels, and the slope easily inferred from it. Second, if $\Phi(1) = 0$ so that a unit root is present, then $u_t \sim I(1)$ and y_t is a difference stationary (DS) process. Here, $\mathrm{E}(y_t)$ is undetermined, but still $\mathrm{E}(\Delta y_t) = \delta(t)$, and so the slope should be modeled from the data in first differences. Whether to use the TS or DS specifications depends on the outcomes of unit root tests, which place the DS model as the null hypothesis, $H_0 : \Phi(1) = 0$.

The trend function is often parameterized as

$$\tau(t) = \alpha_1 + \alpha_2 t + \text{Dummy variables for level shifts} + \text{Dummy variables for changes in slope}\,,$$

where the set of dummy variables is included to capture instabilities in $\tau(t)$.

Unit root tests, in general, are known to have low power (a high probability *not* to reject a false H_0), when $\tau(t)$ is misspecified. On one hand, $\tau(t)$ may include redundant terms, not present in the data generating process (1), which leads to overfitting and an unnecessary loss in degrees of freedom eventually resulting in a loss of power. A leading example occurs when the trend function, but not the data generating process, include the linear trend term t. On the other hand, $\tau(t)$ may exclude important terms, a form of misspecification that in practice is associated to considerably larger losses of power, as the excluded terms may be confound with a unit root in y_t. As mentioned, the leading example is an unmodeled structural break, and the power properties of unit root tests are sensitive to whether the presence of structural breaks is considered in the alternative model, and how many breaks are modeled.

The family of unit root tests we use in this paper, the "flexible trend approach" originally advanced in Becker et al. (2004) and Becker et al. (2006), is motivated by considering a trend function of the form

$$\tau(t) = \alpha_1 + \alpha_2 t + \sum_{k=1}^{n}\left[\beta_k^c \cos\left(\frac{2\pi k}{T}t\right) + \beta_k^s \sin\left(\frac{2\pi k}{T}t\right)\right], \tag{3}$$

where T is the sample size. The n cosine and sine terms form a *Fourier expansion* of an unknown and arbitrary function of t. The motivation for using the sine and cosine terms in the definition of the trend function is that a Fourier expansion is able to approximate absolutely integrable functions to a given degree of accuracy. When $n = T/2$, then the Fourier expansion is able to fit $\tau(t)$ perfectly. When $n < T/2$, then $\tau(t)$ provides a global approximation of the trend, associated to the frequencies $2\pi k/T$ for $k = 1, 2, \ldots, n$. Of course, the larger the number of terms at different frequencies n, the better the approximation provided by the Fourier expansion.

However, the purpose is not to fit $\tau(t)$ with no error, but to control in a simple parametric fashion for the effects that nonlinearities or the instabilities brought by structural breaks may have on the properties, especially the power, of the unit root tests of y_t. Since these phenomena are likely to affect the low frequency behavior of the time series, Becker et al. (2006) and Enders and Lee (2012a,b) propose to either select a small value for n (say, $n = 1, 2$), or even to pick up a single frequency and use

$$\tau(t) = \alpha_1 + \alpha_2 t + \beta^c \cos\left(\frac{2\pi k}{T} t\right) + \beta^s \sin\left(\frac{2\pi k}{T} t\right), \tag{4}$$

where k is a small number (say, $k = \{1, 2\}$), as the approximation to the trend function. In either way, the method filters out low frequency components that might interfere with an hypothesized unit root.

This formulation of the trend function in the context of unit root testing has several practical advantages. Firstly, the linear trend augmented by an extra sine-cosine pair is capable to approximate well many forms of instabilities that can be regarded as relevant for economic data. Several examples are shown in Figure 1, which is similar to figures shown in Becker et al. (2006) and Enders and Lee (2012b). A remarkable result from this figure is that even though (4) is particularly suitable to approximate “smooth breaks”, it does a decent work approximating “sharp breaks”. Furthermore, the figures illustrates that the approximation works well for small values of k, i.e. the suggestion of Enders and Lee (2012a,b) of using $k = 1$ or $k = 2$, when k is restricted to be an integer.

Secondly, the often difficult task of estimating unknown structural break dates is exchanged by the straightforward estimation of three parameters, namely k, β^c and β^s. To be more precise, the calibration of k and the *linear* estimation of β^c and β^s. The coefficients β^c and β^s affect the amplitude and displacement of the frequency component to the extent that even with a single frequency the trend can approximate multiple breaks, as shown in many examples in Figure 1. Such a parsimonious approach to model the effects of general forms of instabilities prevents overfitting $\tau(t)$, and to data-mine the break dates.

With these insights, Enders and Lee (2012a) propose an extension to the Dickey and Fuller (1979) test to allow for a flexible deterministic trend. In particular, for the single-frequency approximation, the unit root test consists in estimating the equation

$$y_t = C_1 + C_2 t + A\cos\left(\frac{2\pi k}{T} t\right) + B\sin\left(\frac{2\pi k}{T} t\right) + \rho y_{t-1} + \sum_{i=1}^{p} G_i \Delta y_{t-i} + \text{error}_t\,, \tag{5}$$

by ordinary least squares, and perform a standard one-tailed t test for $H_0 : \rho = 1$ against $H_1 : \rho < 1$. Under H_0, y_t is difference stationary, whereas under H_1 it is trend stationary around the moving mean $\tau(t)$. Note that, as it is customary, the testing equation (5) is augmented with p lags of Δy_t to prevent the residuals from exhibiting serial correlation. The choice of p can be made in standard fashion, for instance by minimizing an information criterion across many values of $p = 0, 1, \ldots, p_{\max}$.

As shown in Enders and Lee (2012a,b), the distribution of the associated t statistic under H_0 is nonstandard, and depends on whether the linear term trend is present in the testing equation (whether $C_2 = 0$ or not), and on the choice of k. The distribution, however, does not depend on the values of A and B, and so its critical values can be easily simulated.[4]

On the other hand, the Enders and Lee test boils down to the Augmented Dickey and Fuller test if $A = B = 0$. It also follows that at least one frequency component, typically with small k as suggested in Figure 1, must

[4] Other tests allowing for a flexible deterministic trend in the data generating process and/or the testing equations have been developed. Namely, Becker et al. (2006) extend the stationarity test of Kwiatkowski et al. (1992), Enders and Lee (2012b) proposes an LM unit root test similar to that of Schmidt and Phillips (1992), and Rodrigues and Taylor (2012) generalize the GLS detrending approach of Elliot et al. (1996). These tests are expected to be more powerful than the Dickey-Fuller test we adopt, for the very same reasons discussed in the papers proposing the original “linear” tests. Yet, in our application we find very small differences between the results of these tests and those reported below. No conclusion is altered in any significant way, and to save space we decided not to report these results, which are available upon request.

be significant if there are level shifts or structural breaks. Thus, a standard F test for $H_0 : A = B = 0$, which compares the fit of a model including the trigonometric for a given $k \neq 0$ to the fit of a model lacking such terms $k = 0$, i.e.

$$F(k) = \frac{\text{Sum of squared residuals } (k = 0) - \text{Sum of squared residuals } (k \neq 0)}{\text{Sum of squared residuals } (k \neq 0)} \left(\frac{T - q}{2} \right),$$

where q is the number of regressors in (5), can be interpreted as a test for a linear time trend in the data generating process, against a nonlinear ("unstable") trend. If the nonlinear term is absent from the data generating process, it is possible to increase the power of the unit root test by using the standard Augmented Dickey and Fuller test. Thus, in practice, it may be desirable to pretest for the absence of a nonlinear trend. Enders and Lee (2012a,b) suggest taking a conservative route and impose the unit root on the data generating process to derive the sampling distribution of this F statistic. This distribution is nonstandard, and depends on the presence of the linear term trend and, of course, of k. The resulting critical values are much larger (more than double) than those coming from an comparable F distribution.

A related point is the selection of k. When k is known, the unit root t test and the nonlinearity F test can be conducted directly by choosing the appropriate critical values. If k is unknown, on the other hand, Becker et al. (2006) and Enders and Lee (2012a,b) offer data-driven methods to estimate this frequency, and their corresponding critical values. Our approach, however, is admittedly more conservative. After obtaining a rejection in the $F(k)$ test, we inquiry whether the trigonometric terms are significant enough to remain significant in an equation featuring accumulated frequencies. This equations is

$$y_t = C_1 + C_2 t + \sum_{k=1}^{n} \left[A_k \cos\left(\frac{2\pi k}{T} t \right) + B_k \sin\left(\frac{2\pi k}{T} t \right) \right] + \rho y_{t-1} + \sum_{i=1}^{p} G_i \Delta y_{t-i} + \text{error}_t \,, \tag{6}$$

and a test on the significance of all trigonometric terms, up to $k = n$, is based on the statistic

$$F(n) = \frac{\text{Sum of squared residuals } (k = 0) - \text{Sum of squared residuals } (n \neq 0)}{\text{Sum of squared residuals } (n \neq 0)} \left(\frac{T - q}{2n} \right),$$

Thus, we only take a rejection in *both* $F(k)$ and $F(n)$ tests as evidence against linearity in the trend function.

3 Testing for unit roots in commodity prices

In our empirical analysis, we use the updated version of the celebrated Grilli and Yang (1988) dataset documented in Pfaffenzeller et al. (2007).[5] The data are annual over the period from 1900 to 2010, and include 24 primary commodity price series, 11 of which are foodstuffs, 7 are nonfood soft commodities and 6 are metals. The series of interest are 100 times the logarithm of the ratio of the prices of each commodity to a manufacturing unit value index. The main reason to choose this particular dataset is its wide popularity, as many influential empirical studies found in the literature have used either its individual commodity price series (León and Soto, 1997; Cuddington, 1992; Harvey et al., 2011; Yamada and Yoon, 2014) or its aggregate indices (Cuddington and Urzua, 1989; Bleaney and Greenaway, 1993; Zanias, 2005; Cuddington et al., 2008; Mariscal and Powell, 2014).

We are particularly interested in comparing our results to three recent comprehensive studies. Firstly, Kellard and Wohar (2006) who perform unit root tests allowing up to two endogenously determined structural breaks using the Grilli and Yang data up to 1998; secondly, Balagtas and Holt (2009) who apply unit root tests against smooth transition dynamics to the same series up to 2003; and, finally, Ghoshray (2011) who updates Kellard and Wohar's study to 2003, and extends the econometric analysis using also two-break tests. Hence, we perform our unit root testing for the full sample up to 2010 ($T = 107$ observations after adjusting for initial conditions) and for two subsamples: the first up to 2003 ($T = 100$), comparable to Balagtas and Holt and

[5] Publicly available at `www.stephan-pfaffenzeller.com`.

Ghoshray, and the second up to 1998 ($T = 95$), comparable to Kellard and Wohar. Besides the comparative study, we take the analysis of subsamples as a robustness check for our procedures. Recall that the excluded observations (roughly, the 2000s) are particularly volatile, with most commodity prices experiencing booms and seemingly unusual dynamics due to the influence of the Chinese booming economy.

3.1 Full sample results

For each relative commodity price, we run and report the outcomes of several unit root tests. Namely, the linear test ($k = 0$), the test augmented with the lowest frequency components ($k = 1$), the test augmented with the second frequency terms ($k = 2$) and the test that accumulates the first two frequencies ($n = 2$). Each test is furthermore run including and excluding the linear term t. The choice of the lag length p is the one that minimizes the Schwarz criterion in the linear test from a grid with a maximum of $p_{\max} = 4$ lags. The results are reported in Tables 1 to 4, including the estimated value of ρ, the t statistic for $H_0 : \rho = 1$ and, for the tests featuring Fourier frequencies, the F statistic on the significance of the trigonometric terms.[6]

We have classified the 24 relative commodity prices into 4 groups, according to the conclusions led by the unit root tests. The classification follows a testing strategy that begins with likely overparameterized testing equations, to progressively remove terms that may be lowering the power of the tests. Also, the classification is based on the full sample results using a significance level of 5%. We refer to the subsample results and those using looser significance levels below.

Table 1 presents group A that includes 8 commodities (palmoil, maize, rice, sugar, wheat, hides, jute and zinc). For these commodities the most parameterized versions of the unit root tests, including the linear t term, strongly reject the unit root hypothesis, regardless on whether the testing equation is augmented with trigonometric terms or not. According to the F linearity tests, in most cases we fail to reject the linear trend hypothesis (the exceptions are sugar, jute, and zinc with a 10% significance level) leading us to conclude that all but the linear testing equations overfit the data. As such, we expect these unit root tests to have low power, making rejection by chance rather improbable. The evidence against nonstationarity is quite strong for the members of this group.

Table 2 presents group B which consists of 9 commodities. Here, strong evidence against unit roots is found in the tests pointed out by the F pretest. For the first 3 commodities in this group (lamb, timber and aluminum), the linear trend hypothesis cannot be rejected, suggesting overfitted, low-powered tests whenever trigonometric terms are included. The unit root is then rejected in the adequate linear tests. For the following 4 commodities (tea, cotton, tobacco and wool) the F test strongly rejects linearity when using $k = 1$, a conclusion that is also supported when using accumulated frequencies, $n = 2$. These results suggest that the linear and $k = 2$ tests may be misleading for these commodity prices, due to potentially serious misspecifications. The unit root is then rejected in the appropriate $k = 1$ and $n = 2$ tests. For the remaining 2 commodities (rubber and copper), the F test rejects linearity for $k = 2$ and $n = 2$, and the unit root is subsequently rejected in the corresponding tests.

The relative prices of banana, coffee and cocoa form a third group, C, whose results are shown in Table 3. For these cases the evidence against a unit root or a linear trend was inconclusive when the testing equations included the linear t term. However, a visual inspection of the data (see Figure 2 below) reveals these to be unambiguously driftless series and so the inclusion of the linear trend t may be hindering the power of the unit root tests. Indeed, upon removing the t term, the evidence against the unit root becomes much stronger and categorical. For banana, the F test points out to a nonlinear mean, whereas the linear test appears to be appropriate for coffee and cocoa.

Finally, Table 4 presents group D that contains the 4 series (beef, lead, tin and silver) for whom the unit root

[6] For reference, the tables tabulate critical values for $T = 100$, but inferences are made using the exact number of observations used in the estimation, i.e $T = \{95, 100, 107\}$. The critical values where computed by simulation using 20,000 replications. Since the asymptotic distribution of the t and F statistics are invariant to the magnitude of the coefficients of $\tau(t)$, we follow Enders and Lee (2012a) and set all of them to zero in the data generating process.

hypothesis could not be conclusively rejected, at least using the 5% level of significance. The remarkable finding so far is that upon controlling for potential instabilities in the mean or trend, or for hypothesized super cycle swings, nonstationarity cannot be rejected for only 4 out of 24 commodities.

3.2 Subsample results

In the vast majority of cases the conclusions reached with the full sample are also supported when the tests are run using subsamples. The results are in general quite robust to the inclusion or exclusion of the booming episodes of the 2000s, i.e. to the use of samples with different end points.

Specifically, our full sample classification remains exactly the same for 20 commodities, even with very similar point estimates and test statistics. The four cases where some reclassification is required are indicated in the Tables (in the column labeled "Δ?"). In the case of jute (Table 1) unit root rejections are not obtained for all possible tests, but for the relevant cases indicated by the F test. Thus, jute passes from group A to group B, with evidence of a nonlinear trend captured by the first pair of trigonometric terms. On the other hand, rubber and copper (Table 2) still belong to group B with slight modifications in the results of the pretesting strategy. Finally, the unit root is rejected in the case of lead (Table 4) in the subsamples, favoring a stationary alternative around a nonlinear trend. Thus, only in 3 out of 24 cases the unit root is not rejected with data up to 2003.

3.3 Comparison to previous studies

As mentioned, the Fourier approximation approach to unit root testing provides stronger evidence against nonstationarity in relative commodity prices than that documented in previous empirical studies. Next, we compare the main findings of three important studies to ours. Table 5 presents a qualitative appraisal of the (non)stationarity of each relative price across studies. Recall that for the subsamples ending in 1998 and in 2003, we could not reject the unit root hypothesis only in the cases of beef, tin and silver.

Consider Kellard and Wohar (2006) and Ghoshray (2011), which allow for up to two structural breaks under the alternative hypothesis. With the exception of beef in Kellard and Wohar and beef and silver in Ghoshray, all relative prices regarded as difference stationary in these studies appear to be trend stationary here. A plausible explanation is low power of the two-break tests, which may fail to adequately control for smooth underlying changes in the series using a limited number of dummy variables.

On the other hand, the case of tin is interesting since it is the only instance in which both Kellard and Wohar (at a 10% significance level) and Ghoshray (at a 5% significance level) reject the unit root hypothesis but we fail to do so. The series, shown in the second block of Figure 2 below, seems to display the U-shaped instabilities that Prodan (2008) have concluded to be rather difficult to detect and correct with dummy variables. In fact, Kellard and Wohar find the dates of the structural breaks of this series to be 1918 and 1975, whereas the dates found in Ghoshray are 1957 and 1984 (and 1941 and 1985 using the same approach as in Kellard and Wohar). Such sensitivity of the alternative model puts in doubts the suitability of the two-break tests for this series. Consistently with this unstable behavior, the linear trend is rejected by the F test in Table 4, and so the test that could not reject the unit root captures the U-shaped instabilities by including Fourier terms. In this respect, we feel more confident with our conclusions about this series. A similar argument applies to the case of silver (shown in the first block of Figure 2), found to be trend stationary only by Kellard and Wohar.

Consider now Balagtas and Holt (2009), which perform tests that place a linear unit root model as the null hypothesis against a stationary alternative governed by various forms of smooth transition dynamics. We find again that all relative prices regarded as difference stationary in this study (banana, maize, rice, wheat and jute) appear to be trend stationary here. Becker et al. (2006) and Enders and Lee (2012a,b) stress that the Fourier approximation approach is able to mimic the workings of smooth transition models in particular (when time is the state variable), and other forms of trend nonlinearities in general. Thus, again, the unit root nonrejections in Balagtas and Holt may be the result of low power stemming for a relatively restrictive alternative model. On the other hand, all of the relative prices that we found to be difference stationary (beef, tin and silver) are regarded as nonlinear stationary in Balagtas and Holt. The case of tin is again problematic and these authors find it to display explosive dynamics quite often when fitted to a nonlinear model.

4 Reassessing the Prebisch-Singer hypothesis

Following a standard practice in the literature, we now use the testing equations combined with the classification rendered by the unit root tests to estimate the time-varying slopes of the trend functions. We could then evaluate empirically whether these are negative, as implied by the Prebisch-Singer hypothesis.

It is important to stress that conditional on the knowledge that a particular series is trend stationary or not, the inferences about the significance of the trigonometric terms in the trend functions (or their slopes) can be made using standard distributions. For instance, in the case of trend stationary models, the associated F statistics would be identical to those reported in Tables 1 to 3, but with critical values (from F or χ^2 distributions) about half the critical values used when pretesting in a unit root context. Thus, a computed F statistic above, say, 3 gives an indication of trend nonlinearity. A glimpse at the figures reported in these tables, therefore, point out to the presence of these nonlinearities in a large number of series under analysis. Nonetheless, in order to give a homogenous treatment to all relative commodity prices, and to avoid pretesting biases in the estimation of the slope of the trend functions, we base our inferences on the most parameterized model, namely the one that includes a constant, a linear t term and 2 accumulated Fourier frequencies, i.e. equation (3) with $n = 2$.

Before continuing, it is interesting to enquire about the role of the t term in the trend function. Since the frequencies of the Fourier approximation are such that k is a integer, the trigonometric terms are equal at $t = 0$ and $t = T$. Thus, $\tau(0) = \alpha_1 + \beta_1^c + \beta_2^c$ and $\tau(T) = \tau(0) + \alpha_2 T$. Thus, if $\alpha_2 = 0$, then $\tau(t)$ is restricted to have the same endpoints. This is, of course, not true if y_t exhibits a drift, and so a better approximation is obtained by allowing $\alpha_2 \neq 0$. But the same happens in the more subtle case of an underlying cycle in y_t with a non-integer value of k. For concreteness, suppose that $y_t = \beta_1^c \cos(ft) + \beta_2^s \sin(ft) + u_t$ with $2\pi < fT < 4\pi$ (or $1 < k < 2$). In this case, $\tau(0) \neq \tau(T)$ and allowing $\alpha_2 \neq 0$ will let α_2 absorb the differences of the endpoints due to the "non-integer frequency", *regardless* on whether there is a drift in the data or not. Put it differently, the terms $\{\cos(ft), \sin(ft)\}$ with $2\pi < fT < 4\pi$ can be better approximated with the trigonometric terms featuring integer values of k *and* a linear trend.

This discussion is relevant under the super cycle conjecture. Recall that the frequency $2\pi k/T$ is associated to a period of T/k, i.e. the number of years it takes for a complete cycle to develop. Take the sample size to be $T = 100$ years, as in our data, so the trigonometric terms with frequency $f_1 = 2\pi/T$ ($k = 1$, a period of 100 years) cycle once over the sample period, whereas the trigonometric terms with frequency $f_2 = 4\pi/T$ ($k = 2$, a period of 50 years) cycle twice over the sample period. If we follow Cuddington and Jerrett (2008) and Erten and Ocampo (2012) and consider that a super cycle lasts about 70 years, then its corresponding frequency would be $2\pi k^*/T$ with $k^* \simeq 1.4$.[7]

We proceed as follows, in the case of trend stationary series. Recall that $y_t = \tau(t) + u_t$ and $\Phi(L)u_t = \varepsilon_t$, where $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_{p+1} L^{p+1}$ contains no unit roots, such that $\mathrm{E}(y_t) = \tau(t)$ has the form in (3). The slope is then

$$\delta(t) = \tau(t) - \tau(t-1) = \alpha_2 + \sum_{k=1}^{n} \left[\beta_k^c \Delta \cos(f_k t) + \beta_k^s \Delta \sin(f_k t)\right], \tag{7}$$

where we have defined $f_k = 2\pi k/T$ for brevity. Inferences drawn on $\delta(t)$ can be made after the estimation of

[7] It is important to emphasize that in the unit root tests the value of k is fixed, and so the frequency $f = 2\pi k/T \to 0$ as $T \to \infty$. This is an important force driving the asymptotic results in Enders and Lee (2012a,b). When we think of a super cycle what is fixed is the frequency (or period) rather than k^*, and the asymptotic properties of the t statistic for a unit root are very likely to be severely distorted in this setup. Thus, instead of extending Enders and Lee's testing framework, which is a task beyond the scope of this paper, we choose to use their formulation to approximate the behavior of a hypothesized super cycle in our sample.

the parameters in (7), namely $\boldsymbol{\theta} = (\alpha_2, \beta_1^c, \ldots, \beta_n^c, \beta_1^s, \ldots, \beta_n^s)'$. Upon multiplying y_t by $\Phi(L)$, we obtain the estimable equation

$$y_t = C_1 + C_2 t + \sum_{k=1}^{n} [A_k \cos(f_k t) + B_k \sin(f_k t)] + \sum_{i=1}^{p+1} \phi_i y_{t-i} + \varepsilon_t \,. \tag{8}$$

Each parameter in the trend function (3) is a unique nonlinear function of the coefficients in (8), namely $\boldsymbol{q} = (C_1, C_2, A_1, \ldots, A_n, B_1, \ldots, B_n, \phi_1, \ldots, \phi_{p+1})'$. We now describe the map $\boldsymbol{\theta} = \mathcal{F}(\boldsymbol{q})$ that relates the vector of estimable coefficients $\boldsymbol{q}$ to the parameters of interest $\boldsymbol{\theta}$ through a method of undetermined coefficients. First, note that $C_1 + C_2 t = \Phi(L)(\alpha_1 + \alpha_2 t)$. Since,

$$\Phi(L)t = \left(1 - \sum_{i=1}^{p+1} \phi_i\right) t + \sum_{i=1}^{p+1} i\phi_i \,,$$

it follows that

$$C_1 = \alpha_1 \left(1 - \sum_{i=1}^{p+1} \phi_i\right) + \alpha_2 \sum_{i=1}^{p+1} i\phi_i \qquad \text{and} \qquad C_2 = \alpha_2 \left(1 - \sum_{i=1}^{p+1} \phi_i\right),$$

which is a linear system in α_1 and α_2 that can be easily solved. Similarly, for a given k, we have that $A_k \cos(f_k t) + B_k \sin(f_k t) = \Phi(L)(\beta_k^c \cos(f_k t) + \beta_k^s \sin(f_k t))$. Using basic trigonometric formulae, it can be verified that

$$\Phi(L) \cos(f_k t) = \cos(f_k t) \left(1 - \sum_{i=1}^{p+1} \phi_i \cos(f_k i)\right) - \sin(f_k t) \sum_{i=1}^{p+1} \phi_i \sin(f_k i) \,,$$

$$\Phi(L) \sin(f_k t) = \sin(f_k t) \left(1 - \sum_{i=1}^{p+1} \phi_i \cos(f_k i)\right) + \cos(f_k t) \sum_{i=1}^{p+1} \phi_i \sin(f_k i) \,.$$

After grouping the $\cos(f_k t)$ and $\sin(f_k t)$ terms,

$$A_k = \beta_k^c \left(1 - \sum_{i=1}^{p+1} \phi_i \cos(f_k i)\right) + \beta_k^s \sum_{i=1}^{p+1} \phi_i \sin(f_k i) \quad \text{and} \quad B_k = \beta_k^s \left(1 - \sum_{i=1}^{p+1} \phi_i \cos(f_k i)\right) - \beta_k^c \sum_{i=1}^{p+1} \phi_i \sin(f_k i) \,,$$

which gives a linear system that can be used to recover β_k^c and β_k^s.

To obtain confidence intervals of the parameters in $\boldsymbol{\theta}$ and, more importantly, on the slope function $\delta(t) \equiv \delta(t, \boldsymbol{\theta})$, we perform a parametric bootstrap, based on suggestions in MacKinnon (2002). The procedure is straightforward:

1. Estimate (8) by ordinary least squares, compute the point estimates $\hat{\boldsymbol{\theta}} = \mathcal{F}(\hat{\boldsymbol{q}})$, a point estimate of the trend function $\hat{\tau}(t) = \tau(t, \hat{\boldsymbol{\theta}})$ and store the residuals $\hat{\varepsilon}_t$. Compute the standardized residuals

$$e_t = \sqrt{\frac{T}{T-1}} \left(\frac{\hat{\varepsilon}_t}{\sqrt{1-h_t}} - \frac{1}{T} \sum_{s=1}^{T} \frac{\hat{\varepsilon}_s}{\sqrt{1-h_s}} \right),$$

   where h_t is the (t,t)-th element of the so-called "hat" matrix. By construction, the standardized residuals are homocedastic and have a sample average of zero.

2. Generate a bootstrap sample e_t^* of size $T + T_0$ by resampling e_t with replacement. Then, generate pseudo data for u_t following the recursion $\hat{\Phi}(L) u_t^* = e_t^*$. The recursion is initialized at $u_{-s}^* = 0$ for $s = 1, \ldots, p$ and the first $T_0 = 50$ observations are discarded to mitigate the effects of these initial conditions.

3. Generate pseudo data $y_t^* = \hat{\tau}(t) + u_t^*$ for $t = 1, 2, \ldots, T$. Estimate (8) by ordinary least squares using the pseudo data and store the results. The estimates are stored in vector $\hat{\boldsymbol{q}}_{(b)}$ which can be used to obtain the parameters of interest, $\hat{\boldsymbol{\theta}}_{(b)} = \mathcal{F}(\hat{\boldsymbol{q}}_{(b)})$ and $\hat{\delta}(t)_{(b)} = \tau(t, \hat{\boldsymbol{\theta}}_{(b)}) - \tau(t-1, \hat{\boldsymbol{\theta}}_{(b)})$.

4. Repeat steps 2 and 3 a large number of times (say, $B = 20{,}000$). The confidence limits for $\delta(t)$ are given by the 5th and 95th percentiles of the empirical distribution $\{\hat{\delta}(t)_{(1)}, \hat{\delta}(t)_{(2)}, \ldots, \hat{\delta}(t)_{(B)}\}$.

For difference stationary series the procedure is similar, with minor modifications due to the imposition of the unit root. In this case, $\Phi(L)$ can be factorized as $\Phi(L) = (1-L)\Xi(L)$, where $\Xi(L) = 1 - \xi_1 L - \xi_2 L^2 - \cdots - \xi_p L^p$ is a polynomial of degree p with no unit roots. Now $\Delta y_t = \delta(t) + \Delta u_t$, where $\Xi(L)\Delta u_t = \varepsilon_t$. Upon multiplying Δy_t by $\Xi(L)$, we obtain the estimable equation

$$\Delta y_t = C_2 + \sum_{k=1}^{n} \left[A_k \Delta \cos(f_k t) + B_k \Delta \sin(f_k t)\right] + \sum_{i=1}^{p} \xi_i \Delta y_{t-i} + \varepsilon_t \,. \tag{9}$$

The slope is directly assessed from a model in first differences. The map $\boldsymbol{\theta} = \mathcal{F}(\boldsymbol{q})$ is essentially the same as the one described above, with $\boldsymbol{\theta}$ and $\boldsymbol{q}$ suitably redefined. In particular, each parameter in the (7) can be written as a unique function of the coefficients in (9), $\boldsymbol{q} = (C_2, A_1, \ldots, A_n, B_1, \ldots, B_n, \xi_1, \ldots, \xi_p)'$.

Following Kellard and Wohar (2006) we construct measures for the prevalence of a negatively sloped trend. As a reference, define

$$\psi = \frac{\text{Number of periods such that } \delta(t) \text{ is negative}}{T} \,,$$

which gives the proportion of years in our sample that the point estimate of the slope is negative. Define also,

$$\Psi = \frac{\text{Number of periods such that the bootstrap upper confidence limit of } \delta(t) \text{ is negative}}{T} \,,$$

which gives the proportion of years in our sample that the point estimate of the slope takes *statistically significant* negative values. Since Ψ correctly accounts for the sampling variability in the slope function, this is our preferred measure and the subsequent analysis builds on it.

Figure 2 presents the results for the 24 relative commodity prices, sorted according to the values of Ψ and ψ. Each panel presents the point estimate of $\delta(t)$ as well as its 95% bootstrap confidence interval. The dates for which $\hat{\delta}(t) < 0$ are marked with a hollow circle, whereas the periods where the slope is statistically negative are marked with filled circles. The proportions ψ and Ψ are also reported. Finally, a rescaled version of the data *in levels* is included in each graph, for reference (the data in first differences turned out to be too volatile to be visually informative).

It is interesting to note how despite of having the same basis (a constant, a time trend and the trigonometric terms), the underlying slopes show a wide variety of patterns, not necessarily regular, across different commodities. This is, in fact, a manifestation of the ability of the Fourier terms to approximate arbitrary functions. The slopes generally display long swings (i.e., it tends to take on the same sign for long stretches of time) and rarely a global behavior (i.e., always positive). This observation reinforces the conjecture that a long-lasting cyclical component drives the medium-to-long-run dynamics of commodities prices. We must stress that this is not an artifact derived from the use of sinusoidal functions. The Fourier approximation can easily take a non-cyclical almost-global behavior like in the case of timber, tin or aluminum. The fact that it shows cycle-like patterns in many instances follows from the mere presence of such patterns in the data.

With regards to the prevalence of negative slopes, the Ψ criterion takes the value of zero for 7 commodities (timber, lamb, tin, beef, silver, copper and zinc). This means that for almost a third of the commodities, the Prebisch-Singer hypothesis does not hold *at all* in the sense that, although the point estimate of the slope may be negative, we fail to find any period of a statistically significant downward trend. On the other extreme, the

Ψ indicator is greater that 0.5 for 6 commodities (rubber, maize, hides, wool, cotton and aluminum), which are the cases we found to be more supportive of the Prebisch-Singer hypothesis, as most of the time the underlying slope of these series is negative and statistically significant.

The remaining commodities represent intermediate cases. Whether these are supportive of the Prebisch-Singer hypothesis or not depends on what values of Ψ the researcher considers to be large. The Prebisch-Singer hypothesis refers to a strong prediction about the *global* behavior of commodity prices that is not unambiguously supported by the data, even though we may be able to find individual prices which seems governed by such prediction. As stated by Yamada and Yoon (2014), the only save conclusion is that the Prebisch-Singer hypothesis holds *sometimes*. Note that the same is true for the prevalence of positively sloped trends: for example, no case is found where the lower confidence limit is always positive. The ambiguity stems from the fact that a global hypothesis is unlikely to hold in a group of series that display rather local dynamics, in the form of long swings that are most likely driven by super cycles.

5 Conclusions

We have confirmed a recent trend in the results of unit root testing of relative commodity prices: the unit root hypothesis is often rejected if the alternative model is flexible enough to characterize the behavior of a stationary but persistent series. The literature has considered dummy-variable structural breaks or smooth transition dynamics, whereas in this paper we use a flexible Fourier approximation to the trend of the data. Of the 24 relative commodity prices analyzed, depending on the sample used, the unit root cannot be rejected in only 3 or 4 instances.

Besides its flexibility to account for the effects of structural breaks or related phenomena, the use of the Fourier approximation to an arbitrary trend as a competing alternative model of commodity prices is motivated by conceptual advances pointing out to a long-lived cycle, i.e. a super cycle, in such prices. A persistent cycle can be easily confound with a unit root in a small sample, so by explicitly accounting for an alternative that encompasses such cyclical behavior, a difference stationary model ceased to be a good characterization of the data. A visual inspection to our estimates of the slope of the trend, i.e. how the trend evolves through time, reinforces the notion of an underlying cycle.

This conclusion has important implications for the Prebisch-Singer hypothesis. In particular, the dynamics of commodity prices are difficult to characterize as a global phenomenon, so a sustained decline of these prices may be better interpreted as a long-lasting downswing rather than a secular trend. In this sense, we subscribe the conclusion reached by Kellard and Wohar (2006) that it might be more useful to concentrate research efforts on examining the extent and causes of local, rather than global, trends in commodity prices.

The claim made in this paper is that once we allow a flexible *deterministic* component to filter away the frequencies associated to a super cycle, there is little evidence of nonstationarity. Put it differently, the unit root tests we have considered prevents the effects of a possible long cycle in the data to be attributed to a unit root. But in order to fully understand the properties of these cyclical behavior in commodity prices, to explore its commonalities across commodities, to forecast its future developments, and to derive policy recommendations, the super cycles ought to be modeled as *stochastic* processes. For instance, it would be of interest to assess whether the regime switching model of Engel and Hamilton (1990) that has been successfully applied to exchange rates to unveil their long swings, is also useful to model relative commodity prices. It would also be interesting to model these prices in an unobservable component framework, as done in Ardeni and Wright (1992) though the literature did not favor this route, armed with the prior information that the observed persistence is driven by a stationary cycle. In our opinion, these are open questions for a fruitful research agenda.

References

Ardeni, P. G. abd B. Wright (1992), "The Prebisch-Singer hypothesis: a reappraisal independent of stationarity hypotheses", *The Economic Journal*, 102(413), 803–812.

Arezki, R., K. Hadri, P. Loungani and Y. Rao (2014), "Testing the Prebisch–Singer hypothesis since 1650: Evidence from panel techniques that allow for multiple breaks", *Journal of International Money and Finance*, 42(C), 208-223.

Enders, W. and M. T. Holt (2012), "Sharp breaks or smooth shifts? An investigation of the evolution of primary commodity prices", *American Journal of Agricultural Economics*, 94(3), 659-673.

Balagtas, J. V. and M. T. Holt (2009), "The commodity terms of trade, unit roots, and nonlinear alternatives: A smooth transition approach", *American Journal of Agricultural Economics*, 91(1), 87-105.

Becker, R., W. Enders and S. Hurn (2004), "A general test for time dependence in parameters", *Journal of Applied Econometrics*, 19(7), 899-906.

Becker, R., W. Enders and J. Lee (2006), "A stationarity test in the presence of an unkwon number of smooth breaks", *Journal of Time Series Analysis*, 27(3), 381-409.

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*, 4(3), 349-363.

Choi, I. (2015), *Almost All About Unit Roots. Foundations, Developments, and Applications*, Themes in Modern Econometrics, Cambridge: Cambridge 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(2), 207-227.

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

Cuddington, J. T., R. Ludema and S. A. Jayasuriya (2007), "Prebisch-Singer redux", in Lederman, D. and W. F. Maloney (eds.), *Natural Resources: Neither Curse nor Destiny*, Standford Univesity Press, ch. 5, 103-140.

Cuddington, J. T. and C. M. Urzua (1989), "Trends and cycles in the net barter terms of trade: A new approach", *Economic Journal*, 99(396), 426-42.

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(366), 427-431.

Elliot, G., T. J. Rothenberg and J. H. Stock (1996), "Efficient tests for an autorregresive unit root", *Econometrica*, 64(4), 813–836.

Enders, W. and J. Lee (2012a), "The flexible Fourier form and Dickey–Fuller type unit root tests", *Economics Letters*, 117(1), 196–199.

Enders, W. and J. Lee (2012b), "A unit root test using a Fourier series to approximate smooth breaks", *Oxford Bulletin of Economics and Statistics*, 74(4), 574-599.

Engel, C. and J. D. Hamilton (1990), "Long swings in the dollar: Are they in the data and do markets know it?", *American Economic Review*, 80(4), 689-713.

Erten, B. and J. A. Ocampo (2013), "Super cycles of commodity prices since the mid-nineteenth century", *World Development*, 44(C), 14-30.

Ghoshray, A. (2011), "A reexamination of trends in primary commodity prices", *Journal of Development Economics*, 95(2), 242-251.

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

Harvey, D. I., N. M. Kellard, J. B. Madsen and M. E. Wohar (2010), "The Prebisch-Singer hypothesis: Four centuries of evidence", *Review of Economics and Statistics*, 92(2), 367-377.

Harvey, D. I., S. J. Leybourne and A. M. R. Taylor (2011), "Testing for unit roots and the impact of quadratic trends, with an application to relative primary commodity prices", *Econometric Reviews*, 30(5), 514-547.

Jones, P. M. and W. Enders (2014), "On the use of the flexible Fourier form in unit root tests, endogenous breaks, and parameter instability", in Ma, J. and M. Wohar (eds.), *Recent Advances in Estimating Nonlinear Models: With Applications in Economics and Finance*, Springer, 59-83.

Kellard, N. M. and M. E. Wohar (2006), "On the prevalence of trends in primary commodity prices", *Journal of Development Economics*, 79(1), 146-167.

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 (1–3), 159–178.

León, J. and R. Soto (1997), "Structural breaks and long-run trends in commodity prices", *Journal of International Development*, 9(3), 347-366.

MacKinnon, J. (2002), "Bootstrap inference in econometrics", *Canadian Journal of Economics*, 35(4), 615-645.

Mariscal, R. and A. Powell (2014), "Commodity price booms and breaks: Detection, magnitude and implications for developing countries", IDB Working Paper 444.

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

Prebisch, R. (1950), *The Economic Development of Latin America and its Principal Problems*, New York: United Nations Publications.

Prodan, R. (2008), "Potential pitfalls in determining multiple structural changes with an application to purchasing power parity", *Journal of Business and Economic Statistics*, 26(1), 50-65.

Rodrigues, P. M. M. and A. M. R. Taylor (2008), "The flexible Fourier form and local generalised least squares detrended unit root tests", *Oxford Bulletin of Economics and Statistics*, 74(5), 736-759.

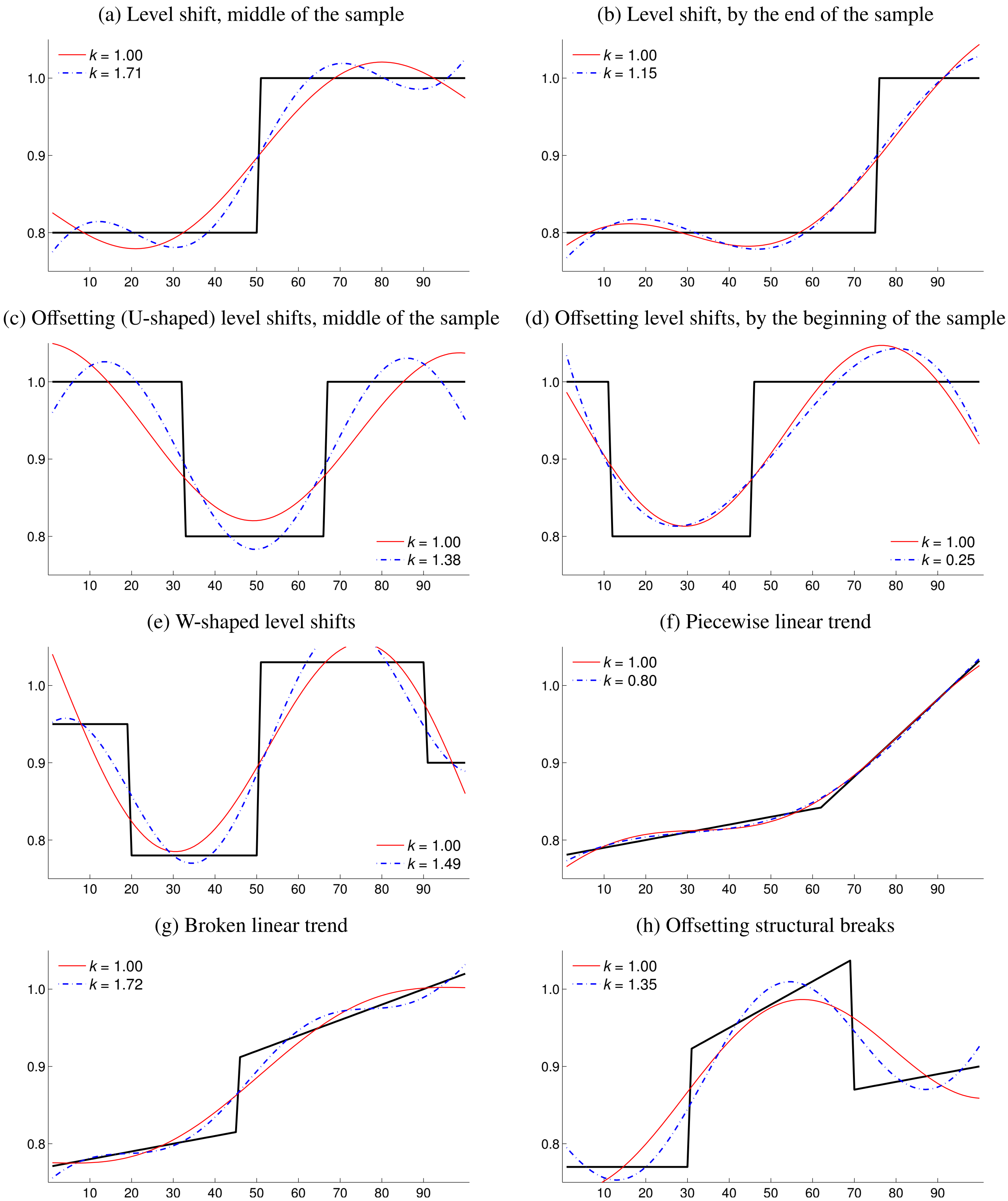
Schmidt, P. and P. C. B. Phillips (1992), "LM tests for a unit root in the presence of deterministic trends", *Oxford Bulletin of Economics and Statistics*, 54(3), 257-287.

Singer, H. W. (1950), "The distribution of gains between investing and borrowing countries", *American Economic Review*, 40(2), 473-485.

Yamada, H. and G. Yoon (2014), "When Grilli and Yang meet Prebisch and Singer: Piecewise linear trends in primary commodity prices", *Journal of International Money and Finance*, 42(C), 193-207.

Zanias, G. P. (2005), "Testing for trends in the terms of trade between primary commodities and manufactured goods", *Journal of Development Economics*, 78(1), 49-59.

Figure 1. *Single-frequency Fourier approximations of structural breaks*



Notes: Fitted values of the function $\tau(t) = \alpha_1 + \alpha_2 t + \beta_1 \cos(ft) + \beta_2 \sin(ft)$, where $f = 2\pi k/T$, using $T = 100$ observations. The red (continuous) line shows the approximation for a given $k = 1$, whereas the blue (dotted) line gives the "best" single-frequency approximation by estimating k by least squares.

Table 1. *Unit root and linear trend tests: Group A (rejection of the unit root regardless the behavior of the trend)*

				$k = 0$		$k = 1$			$k = 2$			$n = 2$			
	Last	LT?	p	ρ	t stat	F stat	ρ	t stat	F stat	ρ	t stat	F stat	ρ	t stat	Δ?
Palmoil	2010	Y	1	0.68	−4.48**	4.08	0.61	−4.88**	4.49	0.64	−4.90**	14.30**	0.49	−5.88**	
	2003	Y	1	0.68	−4.45**	9.95**	0.53	−5.60**	2.06	0.65	−4.62**	12.70*	0.49	−5.81**	
	1998	Y	1	0.68	−4.31**	6.20	0.56	−5.05**	2.59	0.64	−4.53**	8.89	0.53	−5.23**	
Maize	2010	Y	0	0.68	−4.52**	5.08	0.58	−5.06**	1.19	0.67	−4.57**	9.49	0.53	−5.46**	
	2003	Y	0	0.67	−4.45**	8.34*	0.53	−5.42**	0.47	0.66	−4.45**	9.13	0.51	−5.46**	
	1998	Y	0	0.65	−4.42**	7.12*	0.53	−5.26**	0.53	0.64	−4.44**	7.73	0.52	−5.29**	
Rice	2010	Y	1	0.74	−4.39**	3.60	0.67	−4.76**	5.37*	0.68	−4.89**	11.79	0.59	−5.54**	
	2003	Y	1	0.75	−4.12**	16.02**	0.57	−5.95**	2.67	0.70	−4.42**	17.88**	0.53	−6.07**	
	1998	Y	1	0.71	−4.43**	10.76**	0.57	−5.68**	3.31	0.66	−4.77**	12.87*	0.54	−5.82**	
Sugar	2010	Y	1	0.60	−4.83**	1.17	0.57	−4.90**	6.05**	0.52	−5.44**	5.96	0.52	−5.36**	
	2003	Y	1	0.60	−4.71**	4.76	0.53	−5.24**	3.02	0.55	−5.04**	5.41	0.52	−5.26**	
	1998	Y	1	0.60	−4.65**	3.71	0.54	−5.06**	2.46	0.55	−4.91**	4.27	0.51	−5.22**	
Wheat	2010	Y	1	0.65	−4.88**	0.44	0.64	−4.85**	4.52*	0.59	−5.31**	8.48	0.55	−5.68**	
	2003	Y	1	0.62	−5.10**	2.82	0.58	−5.39**	2.93	0.57	−5.38**	4.99	0.54	−5.57**	
	1998	Y	1	0.61	−4.98**	2.34	0.58	−5.20**	2.65	0.56	−5.24**	3.91	0.55	−5.33**	
Hides	2010	Y	0	0.59	−5.15**	6.61	0.51	−5.82**	1.87	0.56	−5.32**	11.55	0.45	−6.28**	
	2003	Y	0	0.60	−4.88**	9.06**	0.47	−5.84**	2.29	0.57	−5.11**	12.02*	0.44	−6.06**	
	1998	Y	0	0.59	−4.84**	7.94*	0.47	−5.70**	3.05	0.55	−5.16**	11.30	0.43	−5.96**	
Jute	2010	Y	1	0.76	−3.70**	6.11	0.59	−4.55**	5.91**	0.70	−4.24**	24.43**	0.39	−6.07**	A
	2003	Y	1	0.77	−3.43*	16.87**	0.48	−5.79**	2.22	0.72	−3.63	24.24**	0.38	−6.09**	B_1
	1998	Y	1	0.75	−3.42*	15.39**	0.49	−5.62**	3.32	0.67	−3.80*	22.99**	0.37	−5.98**	B_1
Zinc	2010	Y	1	0.55	−5.38**	1.22	0.53	−5.46**	4.98*	0.49	−5.87**	5.99	0.47	−5.92**	
	2003	Y	1	0.56	−4.91**	8.23*	0.44	−5.78**	0.72	0.54	−4.94**	9.58	0.43	−5.85**	
	1998	Y	1	0.54	−4.97**	6.38	0.44	−5.66**	1.24	0.52	−5.06**	7.95	0.42	−5.75**	
Critical values	$T = 100$	Y	5%		−3.45	8.50		−4.35	5.84		−4.02	13.80		−5.09	
		Y	10%		−3.15	7.07		−4.05	4.51		−3.69	11.94		−4.78	

Notes: The column "Last" shows the last observation of the sample, which begins in 1904 after adjusting for lags. The column "LT?" indicates if a linear trend term is included (Y) or excluded (N) in the testing equation. The column "p" shows the lag length in the testing equation. The $k = 0$ case is the Augmented Dickey Fuller test, the $k = 1$ and $k = 2$ cases use the single frequency testing equation (5), and the $n = 2$ case uses the accumulated frequencies testing equation (6). The column "Δ?" shows if a group reclassification occurs in the subsample analysis; its entries should be read as (Group letter)$_k$. * [**] denotes rejection of the unit root hypothesis (t stat) or the linear trend/mean hypothesis (F stat) at a 10% [5%] significance level.

Table 2. *Unit root and linear trend tests: Group B (rejection of the unit root for the relevant trend model)*

	Last	LT?	p	$k=0$ ρ	$k=0$ t stat	$k=1$ F stat	$k=1$ ρ	$k=1$ t stat	$k=2$ F stat	$k=2$ ρ	$k=2$ t stat	$n=2$ F stat	$n=2$ ρ	$n=2$ t stat	Δ?
Lamb	2010	Y	4	0.78	-3.69^{**}	0.23	0.78	-3.68	3.78	0.72	-4.04^{**}	4.80	0.71	-4.10	
	2003	Y	4	0.78	-3.58^{**}	0.36	0.77	-3.59	4.88^{*}	0.69	-4.03^{**}	5.25	0.70	-3.96	
	1998	Y	4	0.78	-3.48^{**}	0.23	0.77	-3.46	6.81^{**}	0.66	-4.20^{**}	6.78	0.66	-4.15	
Timber	2010	Y	4	0.68	-4.09^{**}	1.99	0.64	-4.32^{*}	0.39	0.67	-4.08^{**}	2.39	0.62	-4.31	
	2003	Y	4	0.68	-3.94^{**}	1.39	0.65	-4.08^{*}	0.36	0.67	-3.93^{*}	1.63	0.64	-4.04	
	1998	Y	4	0.66	-3.92^{**}	0.62	0.65	-3.95	0.41	0.65	-3.90^{*}	1.01	0.64	-3.91	
Aluminum	2010	Y	1	0.81	-3.57^{**}	5.56	0.72	-4.33^{*}	1.39	0.80	-3.69	8.05	0.70	-4.54	
	2003	Y	1	0.79	-3.74^{**}	4.61	0.72	-4.34^{*}	1.74	0.77	-3.88^{*}	6.94	0.69	-4.54	
	1998	Y	1	0.79	-3.67^{**}	4.20	0.72	-4.21^{*}	1.70	0.77	-3.80^{*}	7.21	0.68	-4.51	
Tea	2010	Y	1	0.85	-2.98	15.53^{**}	0.62	-4.79^{**}	0.98	0.84	-3.07	22.48^{**}	0.54	-5.47^{**}	
	2003	Y	1	0.85	-2.73	20.09^{**}	0.56	-5.28^{**}	0.09	0.85	-2.68	22.20^{**}	0.52	-5.48^{**}	
	1998	Y	1	0.84	-2.85	18.57^{**}	0.53	-5.15^{**}	0.39	0.83	-2.87	20.44^{**}	0.52	-5.28^{**}	
Cotton	2010	Y	1	0.81	-3.46^{**}	9.85^{**}	0.63	-4.63^{**}	1.77	0.80	-3.57	16.18^{**}	0.57	-5.21^{**}	
	2003	Y	1	0.81	-3.39^{*}	12.35^{**}	0.62	-5.00^{**}	1.01	0.79	-3.44	15.50^{**}	0.57	-5.26^{**}	
	1998	Y	1	0.80	-3.35^{*}	11.06^{**}	0.63	-4.83^{**}	1.35	0.77	-3.46	14.60^{**}	0.57	-5.14^{**}	
Tobacco	2010	Y	1	0.87	-3.05	11.01^{**}	0.71	-4.53^{**}	0.60	0.87	-3.02	23.83^{**}	0.56	-5.73^{**}	
	2003	Y	1	0.88	-2.67	9.29^{**}	0.70	-4.40^{**}	1.26	0.86	-2.85	21.07^{**}	0.56	-5.41^{**}	
	1998	Y	1	0.86	-2.95	8.08^{*}	0.70	-4.45^{**}	1.52	0.83	-3.12	19.05^{**}	0.57	-5.32^{**}	
Wool	2010	Y	1	0.78	-3.57^{**}	17.68^{**}	0.48	-5.51^{**}	1.39	0.77	-3.65	31.53^{**}	0.33	-6.69^{**}	
	2003	Y	1	0.77	-3.56^{**}	24.27^{**}	0.34	-6.17^{**}	0.23	0.77	-3.46	29.22^{**}	0.28	-6.47^{**}	
	1998	Y	1	0.79	-3.10	25.38^{**}	0.35	-6.12^{**}	0.08	0.79	-3.07	29.05^{**}	0.28	-6.42^{**}	
Rubber	2010	Y	1	0.82	-3.05	1.84	0.78	-3.26	6.96^{**}	0.71	-4.11^{**}	12.63^{*}	0.59	-5.10^{**}	B_2
	2003	Y	1	0.78	-3.58^{**}	4.83	0.67	-4.22^{*}	4.91^{*}	0.70	-4.18^{**}	14.97^{**}	0.54	-5.33^{**}	B_0
	1998	Y	1	0.78	-3.51^{**}	6.34	0.66	-4.34^{*}	4.18	0.70	-4.00^{*}	15.13^{**}	0.54	-5.20^{**}	B_0
Copper	2010	Y	1	0.86	-2.64	0.58	0.85	-2.55	14.18^{**}	0.69	-4.33^{**}	16.37^{**}	0.62	-5.18^{**}	B_2
	2003	Y	1	0.82	-3.13	16.05^{**}	0.60	-5.15^{**}	6.85^{**}	0.71	-4.02^{**}	22.89^{**}	0.51	-5.70^{**}	B_1
	1998	Y	1	0.82	-3.03	17.04^{**}	0.59	-5.11^{**}	5.75^{*}	0.71	-3.84^{*}	21.71^{**}	0.51	-5.46^{**}	B_1
Critical values	$T=100$	Y	5%		-3.45	8.50		-4.35	5.84		-4.02	13.80		-5.09	
		Y	10%		-3.15	7.07		-4.05	4.51		-3.69	11.94		-4.78	

Notes: The column "Last" shows the last observation of the sample, which begins in 1904 after adjusting for lags. The column "LT?" indicates if a linear trend term is included (Y) or excluded (N) in the testing equation. The column "p" shows the lag length in the testing equation. The $k = 0$ case is the Augmented Dickey Fuller test, the $k = 1$ and $k = 2$ cases use the single frequency testing equation (5), and the $n = 2$ case uses the accumulated frequencies testing equation (6). The column "Δ?" shows if a group reclassification occurs in the subsample analysis; its entries should be read as (Group letter)$_k$. * [**] denotes rejection of the unit root hypothesis (t stat) or the linear trend/mean hypothesis (F stat) at a 10% [5%] significance level.

Table 3. *Unit root and linear trend tests: Group C (rejection of the unit root for driftless data)*

				$k=0$		$k=1$			$k=2$			$n=2$			
	Last	LT?	p	ρ	t stat	F stat	ρ	t stat	F stat	ρ	t stat	F stat	ρ	t stat	Δ?
Banana	2010	N	1	0.86	−2.76*	11.59**	0.66	−4.26**	0.22	0.86	−2.77	12.51**	0.63	−4.53**	
	2003	N	1	0.88	−2.23	16.44**	0.61	−4.63**	0.50	0.88	−2.22	17.41**	0.58	−4.86**	
	1998	N	1	0.90	−2.15	11.68**	0.71	−3.93**	0.18	0.90	−2.14	12.00**	0.66	−4.42*	
Coffee	2010	N	1	0.81	−3.25**	3.89	0.74	−3.78*	2.14	0.79	−3.41**	7.66	0.69	−4.14	
	2003	N	1	0.82	−2.95**	5.00	0.73	−3.70*	1.11	0.80	−3.10*	7.36	0.70	−3.85	
	1998	N	1	0.80	−3.16**	4.00	0.72	−3.75*	0.94	0.79	−3.20*	5.60	0.69	−3.87	
Cocoa	2010	N	1	0.83	−3.34**	2.44	0.80	−3.65*	4.14*	0.77	−3.82**	8.69	0.71	−4.37*	
	2003	N	1	0.83	−3.20**	3.61	0.78	−3.63*	1.67	0.80	−3.41**	6.57	0.74	−4.00	
	1998	N	1	0.83	−3.05**	4.30	0.77	−3.57*	0.77	0.82	−3.14*	5.91	0.74	−3.77	
Critical values	$T = 100$	N	5%		−2.89	7.04		−3.80	4.27		−3.27	11.53		−4.50	
		N	10%		−2.58	5.68		−3.48	3.21		−2.90	9.77		−4.17	

Notes: The column "Last" shows the last observation of the sample, which begins in 1904 after adjusting for lags. The column "LT?" indicates if a linear trend term is included (Y) or excluded (N) in the testing equation. The column "p" shows the lag length in the testing equation. The $k = 0$ case is the Augmented Dickey Fuller test, the $k = 1$ and $k = 2$ cases use the single frequency testing equation (5), and the $n = 2$ case uses the accumulated frequencies testing equation (6). The column "Δ?" shows if a group reclassification occurs in the subsample analysis; its entries should be read as (Group letter)$_k$. * [**] denotes rejection of the unit root hypothesis (t stat) or the linear trend/mean hypothesis (F stat) at a 10% [5%] significance level.

Table 4. *Unit root and linear trend tests: Group D (non-rejection of the unit root)*

				$k = 0$		$k = 1$			$k = 2$			$n = 2$			
	Last	LT?	p	ρ	t stat	F stat	ρ	t stat	F stat	ρ	t stat	F stat	ρ	t stat	Δ?
Beef	2010	Y	4	0.80	−3.22*	3.70	0.72	−3.73	7.26**	0.68	−4.10**	9.29	0.64	−4.25	
		N	4	0.94	−1.66	5.99*	0.82	−2.93	0.47	0.95	−1.44	5.96	0.81	−2.71	
	2003	Y	4	0.79	−3.09	4.19	0.72	−3.65	5.71*	0.70	−3.81*	9.94	0.61	−4.28	
		N	4	0.94	−1.77	7.16**	0.80	−3.15	0.37	0.94	−1.56	9.47	0.74	−3.33	
	1998	Y	4	0.78	−3.08	3.42	0.72	−3.44	4.48	0.71	−3.62	9.54	0.60	−4.20	
		N	4	0.93	−1.79	6.60*	0.82	−3.04	0.64	0.93	−1.64	10.71*	0.72	−3.49	
Lead	2010	Y	2	0.83	−2.68	2.51	0.81	−2.61	4.99*	0.78	−3.21	13.11*	0.67	−3.96	D
		N	2	0.85	−2.69*	2.66	0.80	−3.01	4.85**	0.80	−3.15*	11.30*	0.67	−3.95	
	2003	Y	2	0.81	−2.94	20.99**	0.46	−5.56**	0.75	0.80	−2.99	22.50**	0.43	−5.67**	B_1
		N	2	0.88	−2.13	6.19*	0.79	−3.19	1.48	0.85	−2.40	12.42**	0.67	−4.04	
	1998	Y	2	0.79	−2.95	19.69**	0.46	−5.45**	0.88	0.77	−3.05	20.43**	0.44	−5.49**	B_1
		N	2	0.85	−2.35	5.31	0.76	−3.21	1.83	0.81	−2.68	10.57*	0.66	−3.94	
Tin	2010	Y	1	0.87	−2.72	0.73	0.85	−2.61	5.52*	0.83	−3.17	7.88	0.80	−3.13	
		N	1	0.88	−2.66*	1.14	0.85	−2.81	3.43*	0.86	−2.76	5.47	0.80	−3.06	
	2003	Y	1	0.89	−2.39	10.77**	0.76	−4.02	3.18	0.85	−2.86	18.55**	0.65	−4.85*	
		N	1	0.89	−2.51	4.39	0.83	−3.16	3.15	0.86	−2.88	11.65**	0.74	−4.03	
	1998	Y	1	0.87	−2.53	8.93**	0.77	−3.81	2.97	0.83	−2.98	18.86**	0.64	−4.96*	
		N	1	0.87	−2.66*	4.86	0.80	−3.38	2.45	0.85	−2.92*	10.54*	0.73	−4.02	
Silver	2010	Y	1	0.90	−2.25	1.02	0.86	−2.34	8.30**	0.82	−3.11	10.57	0.79	−3.14	
		N	1	0.92	−1.81	3.57	0.85	−2.60	2.94	0.91	−1.78	8.11	0.80	−2.86	
	2003	Y	1	0.88	−2.54	9.16**	0.74	−3.81	5.72*	0.82	−3.28	17.77**	0.60	−4.70	
		N	1	0.89	−2.47	5.04	0.80	−3.17	3.25*	0.85	−2.84	14.21**	0.68	−4.25*	
	1998	Y	1	0.87	−2.55	8.67**	0.74	−3.74	3.59	0.84	−2.94	18.54**	0.57	−4.83*	
		N	1	0.88	−2.42	6.67*	0.77	−3.39	2.14	0.87	−2.62	13.62**	0.67	−4.17*	
Critical values	$T = 100$	Y	5%		−3.45	8.50		−4.35	5.84		−4.02	13.80		−5.09	
		N	5%		−2.89	7.04		−3.80	4.27		−3.27	11.53		−4.50	
		Y	10%		−3.15	7.07		−4.05	4.51		−3.69	11.94		−4.78	
		N	10%		−2.58	5.68		−3.48	3.21		−2.90	9.77		−4.17	

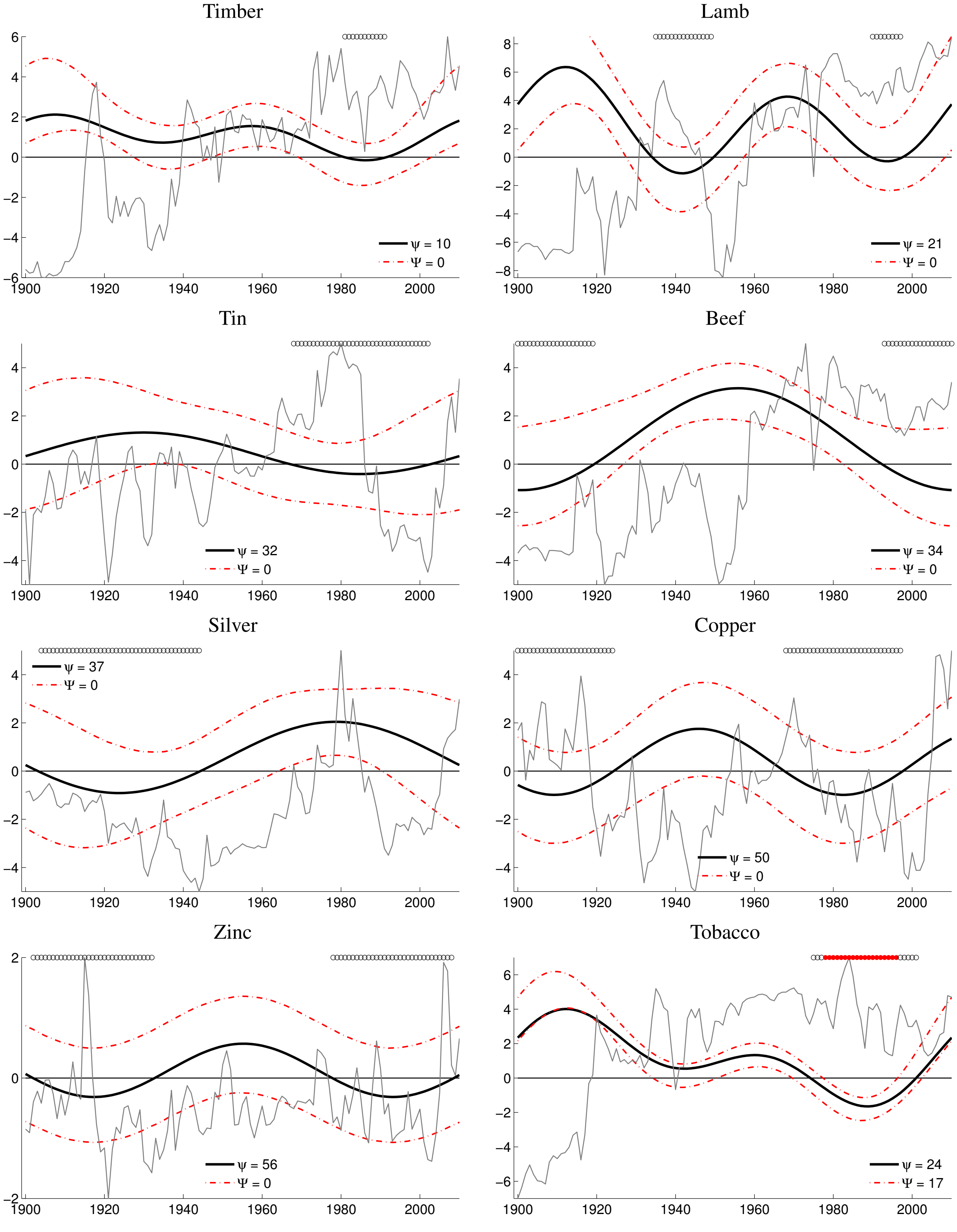
Notes: The column "Last" shows the last observation of the sample, which begins in 1904 after adjusting for lags. The column "LT?" indicates if a linear trend term is included (Y) or excluded (N) in the testing equation. The column "p" shows the lag length in the testing equation. The $k = 0$ case is the Augmented Dickey Fuller test, the $k = 1$ and $k = 2$ cases use the single frequency testing equation (5), and the $n = 2$ case uses the accumulated frequencies testing equation (6). The column "Δ?" shows if a group reclassification occurs in the subsample analysis; its entries should be read as (Group letter)$_k$. * [**] denotes rejection of the unit root hypothesis (t stat) or the linear trend/mean hypothesis (F stat) at a 10% [5%] significance level.

Table 5. *Comparisons to previous studies*

	From 1900 to 1998				From 1900 to 2003					
	Kellard and Wohar		Flexible		Balagtas and Holt		Ghoshray		Flexible	
	y	Breaks	y	Group	y	Nonlin	y	Breaks	y	Group
Banana	DS	-	TS	C_1	DS	L	TS	1	TS	C_1
Beef	DS	-	DS	D	TS	S	DS	-	DS	D
Cocoa	DS	-	TS	C_0	TS	L	DS	-	TS	C_0
Coffee	DS	-	TS	C_0	TS*	S	DS	-	TS	C_0
Lamb	DS	-	TS	B_0	TS	U	DS	-	TS	B_0
Maize	TS*	2	TS	A	DS	L	DS	-	TS	A
Palmoil	TS*	1	TS	A	TS	S	TS*	2	TS	A
Rice	TS	1	TS	A	DS	L	TS	2	TS	A
Sugar	DS	-	TS	A	TS	S	DS	-	TS	A
Tea	TS	2	TS	B_1	TS	U	DS	-	TS	B_1
Wheat	DS	-	TS	A	DS	L	TS	2	TS	A
Cotton	DS	-	TS	B_1	TS	U	TS	2	TS	B_1
Hides	TS*	2	TS	A	TS	S	DS	-	TS	A
Jute	TS*	1	TS	B_1	DS	L	DS	-	TS	B_1
Rubber	TS*	2	TS	B_0	TS	L	TS*	2	TS	B_0
Timber	TS*	2	TS	B_0	TS	U	TS	2	TS	B_0
Tobacco	DS	-	TS	B_1	TS	U	TS*	2	TS	B_1
Wool	TS*	2	TS	B_1	TS	U	TS*	2	TS	B_1
Aluminum	TS*	1	TS	B_0	TS	S	DS	-	TS	B_0
Copper	DS	-	TS	B_1	TS	S	TS*	2	TS	B_1
Lead	TS	2	TS	B_1	TS	S	TS	2	TS	B_1
Silver	TS	2	DS	D	TS	S	DS	-	DS	D
Tin	TS*	2	DS	D	TS*	S	TS	2	DS	D
Zinc	TS	2	TS	A	TS	L	TS*	1	TS	A
Number of TS	14		21		19		13		21	
Number of DS	10		3		5		11		3	

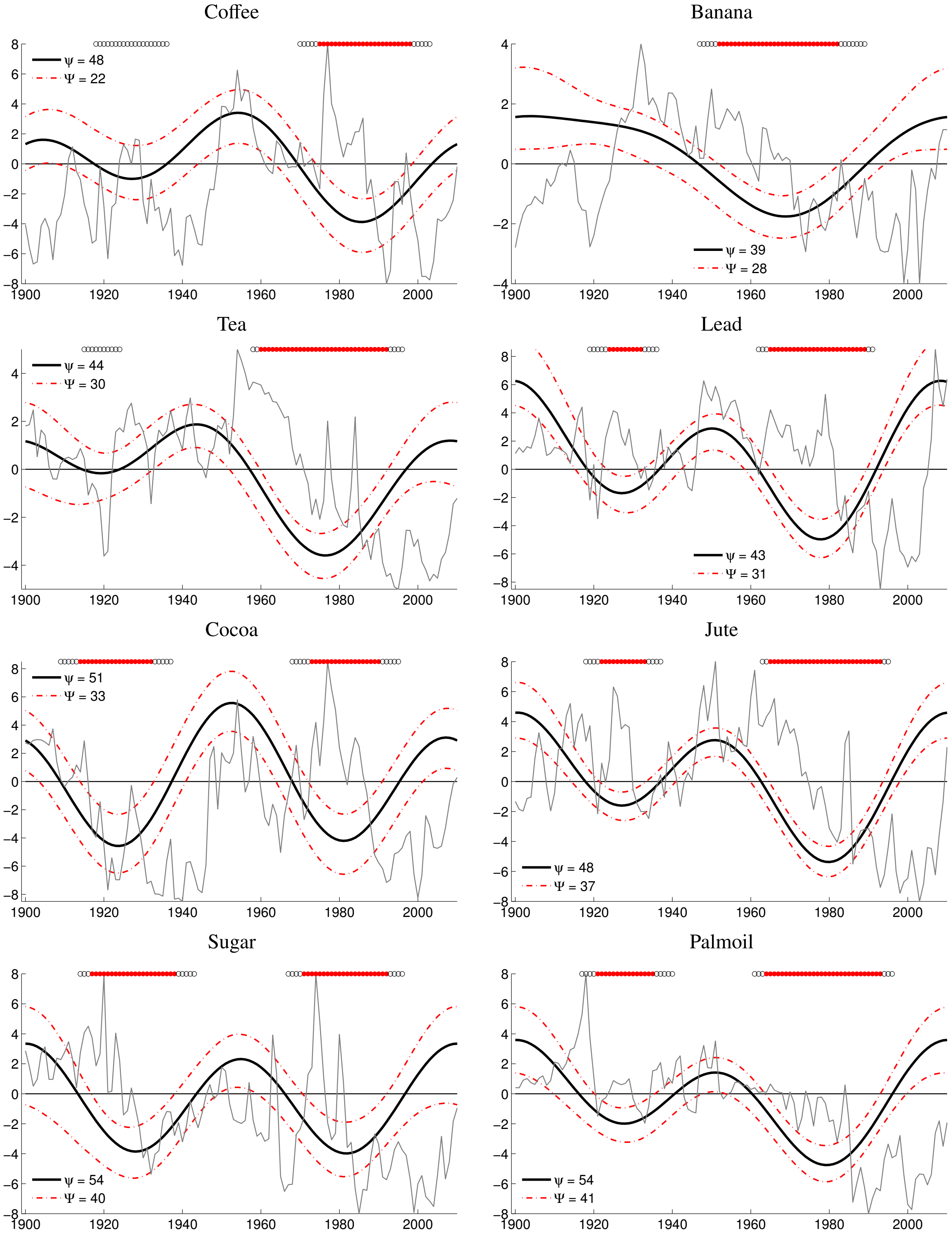
Notes: DS stands for a difference stationary process (the null hypothesis in all cases), TS for a trend stationary process at a 5% significance level, and TS* for a trend stationary process at a 10% significance level. The column "Breaks" shows the number of level shifts or structural breaks reported in Kellard and Wohar (2006) and Ghoshray (2011). The columns "Flexible" show the classifications obtained in Tables 1 to 4 for the corresponding subsamples. The "Nonlin" classification of Balagtas and Holt (2009) is as follows: L for a linear model, S for a smooth transition model with a logistic (S-shaped) transition function, and U for a smooth transition model with a quadratic (U-shaped) transition function.

Figure 2. *Relative commodity prices and estimated underlying slopes*



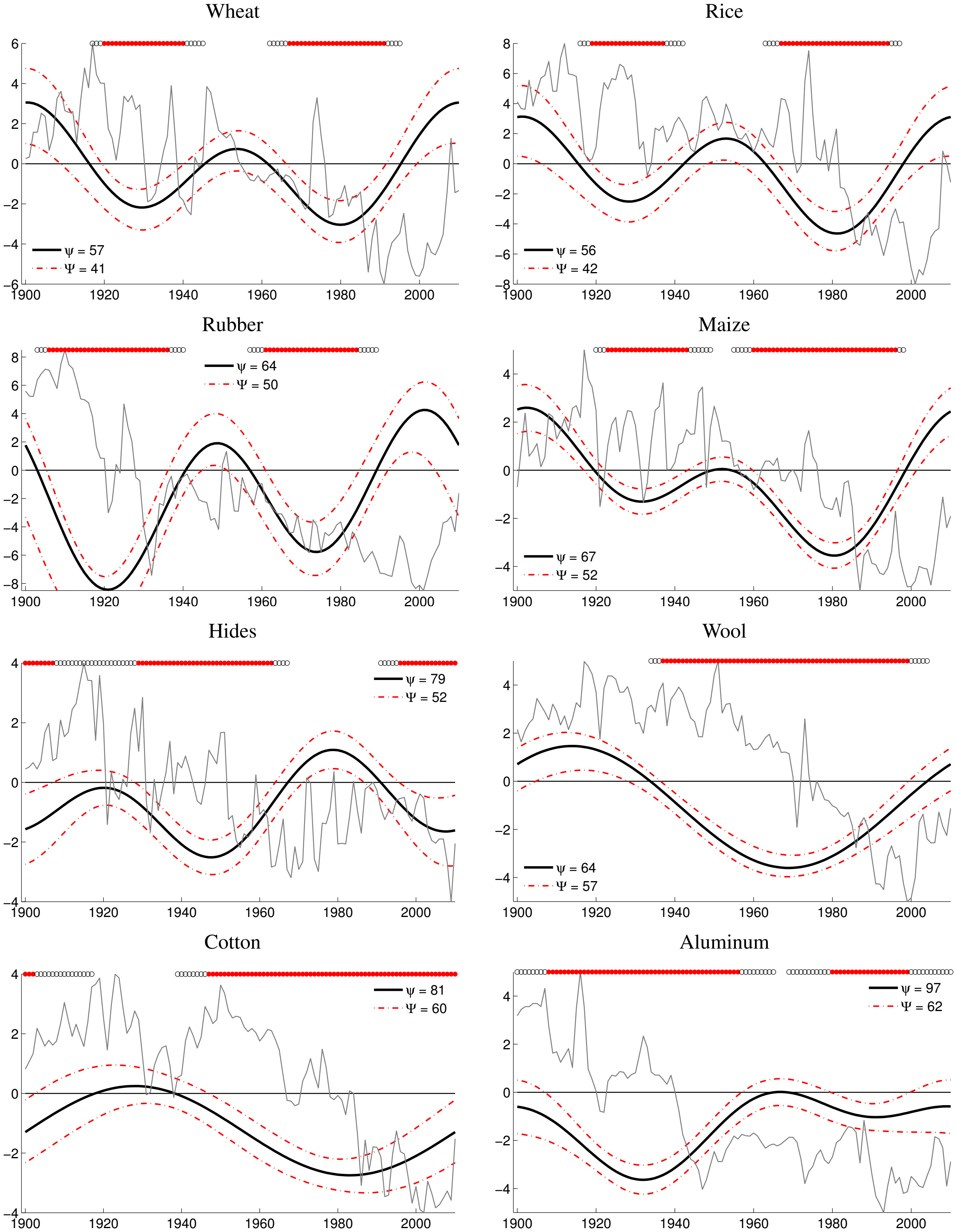
Notes: Thin (noisy, gray) line: Data *in levels* (rescaled); Thick (smooth, blue) line: point estimate of $\delta(t)$; Dotted (smooth, red) lines: bootstrap confidence limits of $\delta(t)$, using 20,000 replications. ψ is the frequency of $\delta(t) < 0$ (dates marked by a hollow circle), whereas Ψ is the frequency of a negative upper confidence limit (dates marked by a filled circle).

Figure 2. (cont') *Relative commodity prices and estimated underlying slopes*



Notes: Thin (noisy, gray) line: Data *in levels* (rescaled); Thick (smooth, blue) line: point estimate of $\delta(t)$; Dotted (smooth, red) lines: bootstrap confidence limits of $\delta(t)$, using 20,000 replications. ψ is the frequency of $\delta(t) < 0$ (dates marked by a hollow circle), whereas Ψ is the frequency of a negative upper confidence limit (dates marked by a filled circle).

Figure 2. (cont') *Relative commodity prices and estimated underlying slopes*



Notes: Thin (noisy, gray) line: Data *in levels* (rescaled); Thick (smooth, blue) line: point estimate of $\delta(t)$; Dotted (smooth, red) lines: bootstrap confidence limits of $\delta(t)$, using 20,000 replications. ψ is the frequency of $\delta(t) < 0$ (dates marked by a hollow circle), whereas Ψ is the frequency of a negative upper confidence limit (dates marked by a filled circle).