

THE COMMODITY TERMS OF TRADE, UNIT ROOTS, AND NONLINEAR ALTERNATIVES: A SMOOTH TRANSITION APPROACH

JOSEPH V. BALAGTAS AND MATTHEW T. HOLT

This article extends the recent literature on the Prebisch–Singer hypothesis of a long-run decline in the relative prices of primary commodities. Our main innovation is testing for and estimating nonlinear alternatives to a secular deterioration. Specifically, we use bootstrap procedures to test the linear unit root model against models belonging to the family of smooth transition autoregressions (STARs) for twenty-four commodities, 1900–2003. In nineteen cases we reject the linear null at usual significance levels. In sixteen cases we are able to successfully fit STAR-type models. Simulation results show there is little support for the Prebisch–Singer hypothesis.

Key words: nonlinear model, primary commodities, smooth transition autoregression, time-varying autoregression, unit root tests.

Introduction

Have prices for primary commodities been in a long-term secular decline relative to those of manufactured goods? This basic question, posed independently by Raul Prebisch (1950) and Hans Singer (1950) over 50 years ago, and since labeled as the Prebisch–Singer hypothesis (PSH), remains at the core of a continuing debate among trade and development economists. Both authors concluded that relative commodity prices have declined and, moreover, will likely continue to do so. From an empirical perspective the basic question posed by the PSH seems innocuous enough, and is one that, at least in principle, should be relatively easy to answer. And yet the PSH continues to receive considerable attention. See, for example, Grilli and Yang (1988), Cuddington and Urzua (1989), Powell (1991), Ardeni and Wright (1992), Cuddington, Ludema, and Jayasuriya (2006), and Kellard and Wohar (2006), among others.

Joseph Balagtas is assistant professor and Matthew Holt is professor and Wickersham Chair of Excellence in Agricultural Research in the Department of Agricultural Economics, Purdue University.

This research has been supported by the Agricultural Experiment Station at Purdue University. Senior authorship is not assigned. We are grateful to Timo Teräsvirta and Aaron Smith for many useful comments and suggestions during the initial phases of this research and to the editor, Chris Barrett, and two anonymous referees for numerous helpful suggestions on prior drafts. Part of this research was completed while the second author was visiting the Department of Economic Statistics, Stockholm School of Economics, whose kind hospitality is gratefully acknowledged.

There are at least two reasons for continued interest in the PSH. First, the policy implications associated with the long-run behavior of commodity prices are fundamentally important. Deaton (1999, pp. 28–29), writing in the context of economic development in Africa, made this point concisely: “Sensible development and macroeconomic policy rules for commodity-exporting countries must be grounded in an understanding of the behavior of commodity prices. The urgency and attractiveness of export diversification depend greatly on whether real prices can be expected to trend up or down in the future.”

The other reason for sustained interest in this topic is that the fundamental question implied by the PSH is naturally an empirical one; it is simply impossible to answer at some point without turning to data. Perhaps not surprisingly, it seems the answer to the empirical question posed by the PSH hinges on appropriate specification of the time series behavior of relative commodity prices. Until rather recently empirical work on this topic assumed that the natural logarithm of (relative) commodity prices were stationary, typically around a linear trend. See, for example, Spraos (1980), Sapsford (1985), Thirlwall and Bergevin (1985), and Grilli and Yang (1988). Tests of the PSH typically boiled down to questions pertaining to the sign and significance of the estimated trend term. Of course many researchers recognized that structural breaks have seemingly occurred and, as well, that

autocorrelation corrections might be appropriate. This early work tended to support the PSH of a negative secular trend in relative commodity prices.

More recently researchers recognized that primary commodity prices might be difference stationary, and therefore might evolve according to a stochastic trend. Among the first to examine this possibility were Cuddington and Urzua (1989) and Cuddington (1992). Others following in this vein include Bleaney and Greenaway (1993), Newbold and Vougas (1996), and Kim et al. (2003). Here interest focuses on the sign and significance of the drift term (i.e., intercept term), inasmuch as this would be indicative of a stochastic trend in a unit root specification. A general conclusion is that when difference stationarity is imposed (i.e., when the data are simply first differenced) there is less support for the notion of long-term attenuation in the commodity terms of trade. Naturally the results may be sensitive to the period used as well as to whether or not perceived structural breaks are accounted for.

Indeed, a key issue in testing the PSH in recent years is whether the underlying price series has experienced structural change. Ocampo and Parra (2004) argued that deteriorations in the terms of trade have been discontinuous, with the 1920s and the 1980s being periods for which declines were particularly notable. To the extent that structural breaks have been observed in the data, standard unit root tests may provide misleading results (Perron, 1989). Therefore recent research examining the PSH has focused on employing unit root tests where the possibility of structural breaks is allowed. Relevant examples include Leon and Soto (1997), Zanas (2005), Cuddington et al. (2006), and Kellard and Wohar (2006). The empirical findings generally support the observation that, in contrast to the PSH, deteriorations in the terms of trade have been discontinuous and episodic.¹

While considerable progress has been made in examining the PSH, there is scope for additional work. To begin, research on this topic has yet to consider the question of whether commodity price dynamics are also changing

with time. Specifically, previous research has considered only the case where the intercept term and/or the linear trend term have experienced structural change. A more complete analysis would examine the possibility that structural change has also occurred in the model's autoregressive and/or moving average coefficients.

Alternatively, the data may also exhibit nonlinear features. Potential for nonlinearity in commodity prices is well documented. For example, by using a structural approach Deaton and Laroque (1995) show that the impossibility of negative storage gives rise to nonlinearity in prices of annual storable commodities. Similarly, Holt and Craig (2006) discuss why irreversibility in liquidation decisions pertaining to standing stocks of livestock and perennial crops can give rise to nonlinearities in the prices for these goods. Thirlwall and Bergevin (1985) also examined the possibility of cyclical asymmetries in the commodity terms of trade as a possible explanation for the PSH, but found little support for this thesis. In any event, only Persson and Teräsvirta (2003) have considered explicitly the possibility of a nonlinear relationship in the commodity terms of trade. They did so by using Grilli and Yang's (1988) aggregate price index to estimate a member of the family of smooth transition autoregressions (STARs). They did not, however, formally test for nonstationarity, a potentially important component of any formal assessment of the PSH.

In this article we examine the PSH within a modeling framework that allows formal testing and estimation of a broader range of time series behavior than has been considered in the extant literature. Our methodology builds upon previous research by Teräsvirta (1994) and examines the potential for STARs, time-varying autoregressions (TVARs) (Lin and Teräsvirta 1994), and models that contain both nonlinearity and time variation, or TV-STARs (Lundbergh, Teräsvirta, and van Dijk 2003) in modeling commodity prices. Each of these models employs a member of the family of univariate logistic or exponential functions to capture structural change and/or nonlinearity. As well, a key feature of the TVAR and the TV-STAR is that inclusion of higher-order trend terms in the logistic function allows for the possibility of nonmonotonic and/or instantaneous structural change. In short, these models permit multiple structural breaks of the sort allowed for by Zanas (2005) and Kellard and Wohar (2006).

¹ As well, Lipsey (1994) finds less evidence in favor of deterioration of the commodity prices relative to an aggregate price index once quality adjustments are accounted for. While we do not adjust for quality changes in the present study, this remains as a potentially fruitful area for future research.

This study follows Cuddington (1992) and Kellard and Wohar (2006) in its use of data on individual commodity prices, thus avoiding smoothing problems associated with an aggregate index. The data are on the twenty-four commodities that comprise the Grilli and Yang (1998) index analyzed by much of the previous literature, and updated recently by Phaffenzeller, Newbold, and Rayner (2007) to span the 1900–2003 period. Building on Eklund (2003), we test a linear unit root model against STAR-type alternatives by using a nonparametric bootstrap. As well, Teräsvirta's (1994) testing framework is used to determine if the data are adequately characterized by a linear model, or whether a STAR- or TVAR-type specification might be required. When called for, a nonlinear model is econometrically estimated, then subjected to a battery of diagnostic tests. The resulting models are then used to recover any trends in the commodity terms of trade over time. Where nonlinear features are identified, we use stochastic simulations along forward paths to uncover the long-run price behavior, and to thus evaluate the PSH.

Notably, for nineteen of the twenty-four commodities examined, the linear unit-root model is rejected in favor of a STAR-type alternative. Moreover, we are able to successfully fit a STAR-type model to the data in sixteen instances. Thus parameter nonlinearity and parameter nonconstancy are often relevant features of the model. Moreover, only in the case of wool do we find evidence of a secular decline in the terms of trade; for all other commodities we find no evidence in support of the PSH.

The Modeling Framework

A central thrust of this article is to model appropriately identified commodity prices by using nonlinear time series techniques, specifically, by using one or more members of the family of STARS. The starting point in any STAR-type modeling exercise is the linear autoregressive (AR) model. Let y_t denote the (natural logarithm) of a commodity price. A corresponding AR model of order $p + 1$ is then specified as

$$(1) \quad \Delta y_t = \alpha + \beta y_{t-1} + \Phi' \mathbf{x}_t + \varepsilon_t$$

where Δ is a first difference operator such that $\Delta y_{t-k} = y_{t-k} - y_{t-k-1}$, $\mathbf{x}_t = (\Delta y_{t-1}, \dots, \Delta y_{t-p})'$, $\Phi = (\phi_1, \dots, \phi_p)'$ is a vector of autoregressive parameters to be estimated, and ε_t is an ad-

ditive error process such that $\varepsilon_t \sim iid(0, \sigma^2)$. As written in (1) the AR model does not impose a (global) unit root; such a specification may, however, be obtained by simply imposing the restriction $\beta = 0$. Aside from questions regarding unit roots, the modeling exercise must also determine the lag order p for the model, perhaps by using data-based procedures such as the AIC or BIC. See, for example, Hall (1994).

Variants of the AR model in (1) have been used in recent years to model commodity price data, either individually or in aggregate, and to otherwise examine the PSH. Examples include Newbold and Vogas (1996), Leon and Soto (1997), and Kim et al. (2000). In the context of the unit root version of (1), a statistically significant, negative estimate of the drift term α is taken as *prima facie* evidence in favor of the PSH.

A potential limitation of testing the unit root hypothesis in the context of (1) is that isolated structural breaks may bias the results in favor of finding that $\beta = 0$ (see, e.g., Perron 1989). For example, suppose that the model in (1) is associated with a single, discrete structural break and trend break. Assume these breaks occur at time t_b , such that $0 < t_b < T$. We might then specify the an alternative to (1) as

$$(2) \quad \Delta y_t = \alpha_0 + \beta_0 y_{t-1} + \Phi' \mathbf{x}_t + (\alpha_1 + \beta_1 y_{t-1}) D_{b,t} + \varepsilon_t$$

where $D_{b,t} = 1$ if $t > t_b$, and is 0 otherwise.² Tests of models similar to (2) against a linear unit root specification without trend breaks, that is, against equation (1) with $\beta = 0$, and where the breaks are determined as part of the testing/estimation framework, have been proposed by Banerjee, Lumsdaine, and Stoch (1992), Zivot and Andrews (1992), and Perron (1997), among others. Lumsdaine and Papell (1997) extend the framework to allow for two distinct shifts in intercept and trend terms. These methods have been applied recently by Zanas (2005), Kellard and Wohar (2006), and Wang and Tomek (2007) in testing for unit roots in, respectively, an aggregate commodity price index, a series of twenty-four individual commodity prices, and monthly and

² A more typical specification for (2) is $\Delta y_t = \alpha_0 + \beta_0 y_{t-1} + \Phi' \mathbf{x}_t + \tilde{\alpha}_1 D_{1,t} + \tilde{\beta}_1 D_{2,t} y_{t-1} + \varepsilon_t$, where $D_{1,t}$ is identical to $D_{b,t}$ above, and $D_{2,t} = (t - t_b) D_{1,t}$. These two specifications are essentially identical in that the model in (2) simply absorbs the term $-t_b D_{1,t}$ into the β_1 parameter, that is, $\beta_1 + t_b = \tilde{\beta}_1$.

weekly corn, soybean, hog, and milk prices. Although the results are somewhat mixed, there is considerable evidence that many commodity prices are stationary once structural breaks are considered. Moreover, once structural breaks are allowed for there is generally less evidence in support of the PSH (Kellard and Wohar 2006).

While allowing for one or possibly two breaks in intercept and trend terms admits richer alternatives in testing the unit root hypothesis—and, by extension, the PSH—the above does not exhaust the full range of alternatives. For example, none of these tests allow the model's higher-order dynamics, that is, the autocorrelation parameters in the vector Φ , to change. Moreover, prior research has not considered the possibility that structural change is a potentially smooth process over time. Finally, with the exception of Persson and Teräsvirta (2003), prior research has not investigated the possibility that nonlinearity could be a feature of historical commodity price data. To begin, consider the following generalization of (1)

(3)

$$\Delta y_t = \alpha_0 + \beta_0 y_{t-1} + \Phi_0' x_t + (\alpha_1 + \beta_1 y_{t-1} + \Phi_1' x_t) G(s_t; \gamma, c) + \varepsilon_t$$

where $s_t = t^* = t/T$. In (3) $G(t^*; \gamma, c)$ is the so-called transition function defined over the parameters γ and c , and which, in the spirit of the structural break model in (2), has a value bounded between 0 and 1. The main difference, however, is that values within the unit interval, are now admitted, and thus $D_{b,t}$ is no longer restricted to be a Heaviside indicator function. Moreover, unlike (2) the specification in (3) allows the autocorrelation coefficients, the ϕ_i 's, to vary over time. As written, the model in (3) is a member of the family of time-varying autoregressions, or TVARs, introduced initially by Lin and Teräsvirta (1994). Note in particular that if $\alpha_0 = \alpha_1$, $\beta_0 = \beta_1$, and $\Phi_0 = \Phi_1$, the model in (3) reduces to the linear model in (1).

Of course it is possible to generalize the model in (3) in one or more substantive ways. One important generalization is to note that regime change, that is, structural change, need not be triggered solely, or at all, by a trend variable. Specifically, if t^* is replaced in $G(\cdot)$ by a variable that is a continuous function of the lagged endogenous variable y_{t-d} , $d > 0$, say, $s_t = f(y_{t-d})$, then (3) becomes a member of the family of smooth transition autoregressions,

or STARs, as described initially by Teräsvirta (1994). Here s_t is referred to as the transition variable and d as the delay parameter. Henceforth we adopt the following notation. The variable s_t will be used to denote the transition variable, and unless otherwise indicated it will be used generically to signify either t^* or $f(y_{t-d})$.

A central question then is how exactly to specify the transition function in (3). Several alternatives exist, including the first-order logistic and the exponential functions. These are specified, respectively, as

(4)

$$G(s_t; \gamma, c) = (1 + \exp\{-\gamma(s_t - c)\})^{-1}, \quad \gamma > 0$$

and

(5)

$$G(s_t; \gamma, c) = 1 - \exp\{-\gamma(s_t - c)^2\}, \quad \gamma > 0.$$

If (4) is used in conjunction with (3), the resulting model is a member of Teräsvirta's (1994) logistic smooth transition autoregression (LSTAR) family. Alternatively, if (5) is used in conjunction with (3), the resulting model is a member of Teräsvirta's (1994) exponential smooth transition autoregression (ESTAR) family.

For both transition functions γ is referred to as the speed of adjustment parameter. More specifically, the specification in (4) is such that as $\gamma \rightarrow \infty$ then $G(\cdot)$ becomes a Heaviside indicator function, that is, a function such that $G(\cdot) = 0$ for $s_t < c$ and $G(\cdot) = 1$ for $s_t > c$. In this manner the model in (2) is a special case of the LSTAR. Alternatively, for (5) as $\gamma \rightarrow \infty$ or $\gamma \rightarrow -\infty$ then $G(\cdot) \rightarrow 1$. In (4) c , the location parameter, indicates exactly the point where $G(c; \gamma, c) = 0.5$, whereas in (5) $G(c; \gamma, c) = 0$. Therefore, the structural break model in (2) is also a special case of the ESTAR. Finally, in the particular instance where $s_t = t^*$ and where the transition function is specified as in (4), that is, in the case where a TVAR is specified, and when $\gamma \rightarrow \infty$, the resulting model becomes one of discrete structural change with a single break point.

A generalization of (4) is also available, which, in fact, may be quite useful in empirical work when structural change over time is being modeled. Specifically, the general n th-order logistic function is defined as

$$(6) \quad G(s_t; \gamma, \mathbf{c}) = \left(1 + \exp \left\{ -\gamma \prod_{i=1}^n (s_t - c_i) \right\} \right)^{-1},$$

$$\gamma > 0, c_1 \leq c_2 \leq \dots \leq c_n$$

where \mathbf{c} is now a vector such that $\mathbf{c} = (c_1, c_2, \dots, c_n)$. If n is set to two in (6), the resulting model, called a quadratic STAR, or QSTAR, is similar in several respects to an ESTAR. For example, as $\gamma \rightarrow \infty$ and assuming $c_1 \neq c_2$, $G(s_t; \gamma, \mathbf{c})$ will equal unity for $s_t < c_1$ and for $s_t > c_2$ and zero in between. The central difference is that as $\gamma \rightarrow 0$ the underlying model becomes linear in parameters. Of course the ESTAR is slightly more parsimonious than the QSTAR. In the case where $n > 2$, considerable flexibility is afforded in the model's specification. In particular, and in the context of a TVAR, if $n \geq 3$ any structural change need no longer be monotonic function of t^* . As well, and in the spirit of the prior discussion on testing for unit roots against alternatives that are stationary but allow for one or more structural breaks, the transition function in (6) has enough flexibility to accommodate such a situation. See, for example, Lin and Teräsvirta (1994). To recapitulate, when (6) is combined with (3) and when $s_t = t^*$, the resulting model will readily nest the one- and two-regime change models considered, respectively, by Zanas (2005) and by Kellard and Wohar (2006).

A distinct possibility is that the model in (3) does not adequately capture all of the essential features of the data. Specifically, the model in (3) may require additional (additive) nonlinear components. For example, a three-regime model may be specified as

$$(7) \quad \Delta y_t = \alpha_0 + \beta_0 y_{t-1} + \phi_0' \mathbf{x}_t$$

$$+ (\alpha_1 + \beta_1 y_{t-1} + \phi_1' \mathbf{x}_t) G_1(s_t; \gamma_1, c_1)$$

$$+ (\alpha_2 + \beta_2 y_{t-1} + \phi_2' \mathbf{x}_t) G_2(s_t; \gamma_2, c_2) + \varepsilon_t.$$

Additional details on this model specification are provided in Eitrheim and Teräsvirta (1996). Importantly, these authors also describe a framework in which Lagrange multiplier (LM) tests may be constructed to test the model in (3) for remaining additive nonlinearity of the sort implied in (7).³ Of interest

is that the specification in (7) may accommodate a situation in which both structural change (i.e., $s_t = t^*$) and nonlinearity (i.e., $s_t = f(y_{t-d})$) might be prevailing features of the data. That is, the model in (7) is a form of the TV-STAR models developed originally by Lundbergh, Teräsvirta, and van Dijk (2003). With the exception of Holt and Craig (2006), models of this sort have generally not been considered in prior research on commodity price behavior.

The Testing Framework

In view of the foregoing discussion, important questions remain. First, how is it possible to know if nonlinearity is truly a feature of the data? Second, and related to the first question, how might tests for (global) unit roots be performed when either nonlinearity or structural change is considered as an alternative? Both issues are now examined in greater detail.

Linearity Testing

Regarding the first question, it is desirable to have a method of testing the linear model in (1) against time-varying and/or nonlinear alternatives such as those in (3) or (7). A simple test of the statistical significance of the estimated γ parameter in the relevant transition function is not appropriate because there are unidentified nuisance parameters under the null hypothesis of linearity, notably the autoregressive coefficients implied in ϕ_1 and the constant term a_1 . In the statistics literature this is generally referred to as the Davies (1977, 1987) problem. The implication is that the sampling distribution for the estimator of γ no longer has the usual asymptotic properties, and therefore standard asymptotic tests (i.e., standard t -tests, etc.) no longer apply.

In the case where smooth transition models are considered as an alternative, Luukkonen, Saikkonen, and Teräsvirta (1988) have proposed one workable solution to the Davies-type problem. Specifically, they recommend replacing $G(s_t; \gamma, c)$ in (3) with a suitable Taylor series approximation. For example, if a third-order Taylor series in s_t is used to approximate $G(\cdot)$, we may rewrite (3) as follows

$$(8) \quad \Delta y_t = \delta_1' \tilde{\mathbf{x}}_t + \delta_2' \tilde{\mathbf{x}}_t s_t + \delta_3' \tilde{\mathbf{x}}_t s_t^2 + \delta_4' \tilde{\mathbf{x}}_t s_t^3 + e_t$$

³ More elaborate versions of (7), referred to as multiple regime STARs, or MRSTARs, have been developed by van Dijk and

Franses (1999). These models are not considered here, however, because substantially more data than are presently available for empirical tests of the PSH are required for their implementation.

where $\tilde{x}_t = (1, y_{t-1}, x'_t)'$ and where the parameters in (8) are functions of the original parameters in (3) including the speed-of-adjustment parameter γ and the location parameter c . Also, the error term e_t in (8) is a function of the original error term ε_t as well as a remainder term $R(s_t; \gamma, c)$. Under the null hypothesis of linearity $R(s_t; \gamma, c) = 0$ and $e_t = \varepsilon_t$; standard testing strategies may therefore be applied.

As noted by van Dijk, Teräsvirta, and Franses (2002), it is straightforward to deduce that the parameter $\delta_{1,1}$ and the parameter vectors $\delta_i, i = 2, 3, 4$, are functions of the original parameters $a_0, a_1, \beta_0, \beta_1, \phi_0$, and δ_1 in (3) such that testing $H'_0 : \delta_2 = \delta_3 = \delta_4 = 0$ now constitutes a direct test of the linearity hypothesis $H'_0 : \gamma = 0$. This test, denoted as F_L , may be performed as a standard extra-sum-of-squares test, which under the null hypothesis will have a test statistic that is distributed asymptotically as an F statistic with $3(p + 2)$ and $T - (p + 2)$ degrees of freedom.⁴ As outlined by Teräsvirta (1994), an additional testing sequence may be pursued in an attempt to identify whether an LSTAR or ESTAR model is more appropriate. Specifically, (8) may be used to test the following sequence of hypotheses:

$$(9) \quad H_{04} : \delta_{4j} = 0, j = 1, \dots, p + 2$$

$$(10) \quad H_{03} : \delta_{3j} = 0 \mid \delta_{4j} = 0, j = 1, \dots, p + 2$$

$$(11) \quad H_{02} : \delta_{2j} = 0 \mid \delta_{3j} = \delta_{4j} = 0, \\ j = 1, \dots, p + 2.$$

F versions of the LM tests associated with testing the hypotheses in (9), (10), and (11) are referred to, respectively, as F_4, F_3 , and F_2 . Assuming that linearity is rejected, that is, that H'_0 is rejected, then an LSTAR is chosen if either F_4 or F_2 has the minimal p -value in the sequence. Otherwise, an ESTAR is chosen. See Teräsvirta (1994) for additional details.

Combined Unit Root and Linearity Testing

Newbold and Vougas (1996) demonstrated that the evidence for the Prebisch–Singer hy-

pothesis is weaker if the commodity price data are generated by a unit-root process. Thus, testing for the presence of unit roots is a crucial part of evaluating the PSH.

As noted previously, in the context of a linear autoregressive model standard unit root tests may be conducted by testing whether or not $\beta = 0$ in (1), assuming of course that structural change and/or nonlinearly are not a feature of the data. See Dickey and Fuller (1979, 1981). And as also observed before, numerous studies have investigated tests of linear unit root models against stationary alternatives that allow for one and possibly two structural breaks. However, testing the unit-root version of (1) against nonlinear alternatives such as those depicted in (3) is still an emerging field of inquiry. For example, Enders and Granger (1998) and Caner and Hansen (2001) test the unit root hypothesis against stationary threshold autoregressive (TAR) alternatives. Likewise, Eklund (2003) has recently proposed a test of the linear unit root model against a specific LSTAR alternative. Among other things, Eklund (2003) finds that bootstrapping the relevant test statistics enhances the size properties of the relevant tests in certain limiting cases. In what follows we adopt the overall strategy proposed by Eklund (2003), although with some modifications.

What is desired here is a test of the unit root version of the linear autoregression in (1) against the more general STAR-type alternative in (3). That is, we seek a test of

$$(12) \quad \Delta y_t = \alpha + \phi'x_t + \varepsilon_t$$

against the alternative

$$(13) \quad \Delta y_t = \alpha_0 + \beta_0 y_{t-1} + \phi'_0 x_t \\ + (\alpha_1 + \beta_1 y_{t-1} + \phi'_1 x_t)G(s_t; \gamma, c) + \varepsilon_t$$

where x_t and s_t are as previously defined and $G(\cdot)$ is given by either (4) or (5). As discussed above, it is not possible to directly test (12) against (13) because of unidentified nuisance parameters under the null, however, it may be possible to circumvent this problem by replacing $G(s_t; \gamma, c)$ with a suitable Taylor series approximation. For example, if, as before, $G(s_t; \gamma, c)$ is replaced by a third-order approximation, we obtain

⁴ Alternatively, and in the usual spirit of LM-type tests, a χ^2 version of the linearity test with $3(p + 2)$ degrees of freedom could also be constructed. In small samples of the sort used here, however, the χ^2 test may be seriously oversized while the F version of the test generally has better size properties. For this reason we rely throughout on the corresponding F version of all LM-type tests.

$$(14) \quad \Delta y_t = \delta_0 + \lambda_0 y_{t-1} + \mathfrak{D}'_0 \mathbf{x}_t + \sum_{i=1}^3 \delta_i s_t^i + \sum_{i=1}^3 \lambda_i y_{t-1} s_t^i + \sum_{i=1}^3 \mathfrak{D}'_i \mathbf{x}_t s_t^i + \xi_t$$

where, as before, ξ_t is a function of a remainder term $R(s_t; \gamma, c)$ as well as ε_t in (13). It is now possible to use the auxiliary regression in (14) to directly test for *both* linearity and a unit root in the underlying y_t series. Specifically, testing the hypothesis $H_0^{lur} : \lambda_0 = \delta_1 = \delta_2 = \delta_3 = \lambda_1 = \lambda_2 = \lambda_3 = \mathfrak{D}_{1,1} = \dots = \mathfrak{D}_{3,p} = 0$ constitutes such a test inasmuch as the linear autoregression with a single unit root in (12) is obtained. Note also that $\varepsilon_t = \xi_t$ under the null hypothesis since the remainder term R is identically zero. It is, moreover, possible to construct the usual F test, denoted here by F_{lur} , associated with imposing H_0^{lur} . In general F_{lur} will be associated with $(7 + 3p)$ and $T - (8 + 4p)$ degrees of freedom.

The problem in the present case is that for conventional reasons the standard F statistic is no longer associated with the usual limiting distribution under the null of linearity and a unit root. Eklund (2003) did, in fact, obtain asymptotic results for a special case of (14). Specifically, he worked with the alternative model

$$(15) \quad \Delta y_t = \delta_0 + \lambda_0 y_{t-1} + \mathfrak{D}_{0,1} \Delta y_{t-1} + \mathfrak{D}_{0,1} \Delta y_{t-1} s_t + \xi_t$$

where, moreover, $s_t = y_{t-1}$.

While testing (12) against (15) allows for the development of asymptotic results, the tradeoff is that the alternative in (15) lacks generality and flexibility. Alternatively, it is always possible, as in Li and Maddala (1996) and Enders and Granger (1998), and as suggested by MacKinnon (2002), to simply obtain the empirical distributions of the relevant F_{lur} statistic by using simulation methods. Indeed, Leybourne, Newbold, and Vougas (1998) and Sollis (2005) employed a similar approach in testing for unit roots in the residuals of TVAR models. That is, the appropriate test statistics were generated by using Monte Carlo simulations under the linearity assumption implied by the null hypothesis. Of course the downside to this approach is that (1) the nonlinear model must first be estimated; (2) there is no obvious way to combine linearity testing with testing for a unit root; and (3) the approach is restricted to the class of TVAR models. In

any event, the bootstrap approach we employ does allow for greater flexibility in modeling and testing than would otherwise be the case. Moreover, given that we have a specific null hypothesis in mind (a linear unit root model), the bootstrap can substantially improve the finite sample properties of the desired test statistic.

In implementing our test of the nonstationary AR model in (12) against the (possibly locally) stationary nonlinear model in (15), we estimate both models using the observed sample data and obtain the sample estimate for F_{lur} . We then use a dynamic bootstrap of the null model's estimated residuals to construct a reasonably large number, B , of pseudo samples (Li and Maddala 1996). On each pseudo data set both the null and the auxiliary regression models are re-estimated and simulated values for the F_{lur} statistic are obtained. To construct an empirical p -value for the F_{lur} test statistic we simply observe the fraction of times that the sample value exceeds the corresponding simulated values. In the empirical implementation we base our results on $B = 999$ bootstrap replications.

Data

In the empirical analysis we use annual data on prices for twenty-four primary commodities spanning 1900–2003. Nominal prices are deflated by the United Nations Manufactures Unit Value index. The original data set through 1986 was used by Grilli and Yang (1988) to develop their commodity price index. An extension of the data through 1998 by Cashin and McDermott (2002) was used recently by Kellard and Wohar (2006) and Kim et al. (2003) to evaluate the PSH. We use the more recent extension of the data provided by Pfaffenzeller, Newbold, and Rayner (2007), who updated the data set through 2003.⁵ We conduct our analysis on the natural logarithms of the various price series over the time period 1900–1998, withholding the last

⁵ In addition to offering an extension, the data developed by Pfaffenzeller et al. (2007) differs in a few instances from the Cashin and McDermott (2002) data. We use the Pfaffenzeller et al. (2007) data set here because it is more recent and because the authors clearly document their data sources and weighting methods, making this the likely starting point for future extensions. For the sake of comparison, we also conducted the full array of specification tests and estimations on the Cashin and McDermott (2002) data. Results are very similar, and our main conclusion remains unchanged.

five observations to evaluate post-sample forecasting performance.⁶

Augmented Dickey–Fuller (ADF) tests were conducted to evaluate the unit root hypothesis against an alternative that is stationary in the levels. A detailed description of the test results are reported in a technical appendix (Balagtas and Holt 2008). The overall picture emerging from these tests is one of general support for the unit root hypothesis in commodity prices. This conclusion is consistent with results reported by Kim et al. (2003), and Kellard and Wohar (2006), among others.

Results

Unit Roots vs. Nonlinearity

Of course the ADF test does not take account of any structural change or nonlinearities, features that may be part of the data generating process. To examine this issue in greater detail, we perform tests of the linear unit root model against stationary nonlinear and time-varying alternatives by using the general testing framework outlined in the previous section. Specifically, for each commodity a model consistent with (12) is tested against a model consistent with (15), the point being to construct the F_{lur} statistic. Approximate p -values for this statistic are then constructed by using bootstrap procedures. Regarding tests for nonlinearity, we follow Persson and Teräsvirta (2003) in using $s_t = \Delta y_{t-d} = y_{t-d} - y_{t-d-1}$ as the candidate transition variable, where $d = 1, \dots, 6$. As with the ADF tests, lag lengths are determined by using the AIC. Results are recorded in table 1. With several exceptions, noted below, we report results associated with only the minimal p -value for the F_{lur} test statistic across candidate transition variables for all twenty-four commodities.

As reported in table 1, the null model is rejected at the 5% level for seventeen commodities, aluminum, beef, cocoa, copper, cotton, hides, lamb, lead, palmoil, rubber, silver, sugar, tea, timber, tobacco, wool, and zinc. Two others, coffee and tin, are associated with a rejection of the null at the 10% level. Only in the case of bananas, jute, maize, rice, and wheat does a linear unit root model seem to adequately characterize the data.⁷ For seven

commodities (aluminum, cotton, silver, sugar, timber, tin, and tobacco) the linear unit root model is rejected rather strongly for more than one candidate transition variable. We report these additional results because, as noted in additional detail below, some added degree of latitude might be afforded in specifying and estimating a STAR-type model for these commodities.

One surprising result revealed in table 1 is that every commodity for which linearity is rejected is associated with nonlinearity (i.e., an ESTAR, QSTAR, or LSTAR model) as opposed to structural change (i.e., a TVAR model). For each series associated with rejection of linearity, we perform the testing sequence described by Teräsvirta (1994) in an attempt to identify whether an ESTAR, QSTAR, or LSTAR specification is more appropriate. These results are also reported in table 1. There are nineteen occasions when an LSTAR model is identified and nine instances where an ESTAR or QSTAR is called for.⁸ Of course, the presence of nonlinearity and/or structural change says nothing of the PSH. To evaluate the PSH, we now turn to the task of fitting appropriate nonlinear models for the eighteen commodity price series for which the linear unit root model is rejected.

Estimated STAR-Type Models

We attempt to fit the STAR or TVAR models called for in table 1 by using a nonlinear algorithm to estimate each model's parameters, including those that characterize the relevant transition function (van Dijk, Teräsvirta, and Franses 2002). Of course model estimation is only a preliminary part of the modeling cycle used to fit and assess the performance of the fitted models. We employ the diagnostic methods—in the form of LM tests—described by Eitrheim and Teräsvirta (1996) to evaluate the estimated models for: (1) remaining additive nonlinearity, that is, for specifications similar to that described in (7); and (2) remaining autocorrelation.⁹ In order to conduct tests for remaining nonlinearity we

and wheat). Only in the case of wheat, however, does a one-sided test that the drift term is negative have significance at any usual level (p -value of 0.054). We therefore conclude there is generally little support for the Prebisch–Singer hypothesis for these commodities.

⁸ The total is greater than nineteen because, as already noted, for several commodities we consider more than one candidate delay variable.

⁹ Because overall sample sizes are relatively small, we only consider alternatives to first-order STAR-type model specifications that include a second additive term.

⁶ Plots of the data used, 1900–2003, are reported in the technical appendix (Balagtas and Holt 2008).

⁷ Of the five commodities for which the linear unit root model is not rejected, two are associated with positive drift terms (bananas and jute) and three with negative drift terms (maize, rice,

Table 1. Results of Testing a Linear Unit Root Model Against STAR or TVAR Alternatives and of Applying Teräsvirta's (1994) Model Selection Sequence

Commodity	p	$\min[p_{F_{lur}}]$	d	p_{F_4}	p_{F_3}	p_{F_2}	STAR/TVAR
Aluminum	2	0.017**	1	0.006	0.100	0.539	LSTAR
	2	0.028**	2	0.235	0.532	0.013	LSTAR
Banana	2	0.140					
Beef	1	0.023**	3	0.006	0.044	0.577	LSTAR
Cocoa	2	0.025**	3	0.044	0.184	0.122	LSTAR
Coffee	2	0.094*	3	0.040	0.541	0.257	LSTAR
Copper	5	0.021**	6	0.005	0.674	0.347	LSTAR
Cotton	3	0.003***	3	0.001	0.058	0.619	LSTAR
	3	0.050**	2	0.689	0.056	0.107	ESTAR/QSTAR
Hides	4	0.033**	2	0.177	0.167	0.066	LSTAR
Jute	4	0.135					
Lamb	5	0.046**	1	0.210	0.017	0.481	ESTAR/QSTAR
Lead	4	0.018**	6	0.725	0.028	0.004	LSTAR
Maize	4	0.322					
Palmoil	5	0.014**	2	0.036	0.412	0.080	LSTAR
Rice	4	0.125					
Rubber	1	0.015**	4	0.003	0.140	0.333	LSTAR
Silver	2	0.001***	2	0.001	0.166	0.215	LSTAR
	2	0.001***	6	0.020	0.459	0.001	LSTAR
	2	0.019**	4	0.020	0.930	0.044	LSTAR
Sugar	5	0.025**	1	0.156	0.494	0.066	LSTAR
	5	0.086*	2	0.156	0.984	0.048	LSTAR
Tea	2	0.050**	6	0.083	0.056	0.367	ESTAR/QSTAR
Timber	2	0.040**	2	0.313	0.001	0.069	ESTAR/QSTAR
	2	0.053*	1	0.327	0.029	0.128	ESTAR/QSTAR
Tin	1	0.096*	4	0.287	0.065	0.195	ESTAR/QSTAR
	1	0.100*	3	0.859	0.147	0.048	LSTAR
Tobacco	4	0.019**	2	0.955	0.009	0.079	ESTAR/QSTAR
	4	0.028**	4	0.411	0.002	0.185	ESTAR/QSTAR
Wheat	6	0.299					
Wool	4	0.035**	1	0.317	0.073	0.387	ESTAR/QSTAR
Zinc	6	0.050**	3	0.103	0.488	0.127	LSTAR

Note: The column headed p denotes the optimal number of lags in the linear AR model. The column headed $\min[p_{F_{lur}}]$ denotes the minimum p -value of the linear unit root test over delays $d = 1, \dots, 6$. The column headed d denotes the delay corresponding to $\min[p_{F_{lur}}]$, and columns headed, respectively, p_{F_4} , p_{F_3} , and p_{F_2} correspond to p -values of tests in the model selection sequence. All p -values are obtained by performing $B = 999$ recursive bootstraps of the model's residuals under the respective null hypothesis. The final column indicates whether an LSTAR, ESTAR/QSTAR, or TVAR model is chosen. A single asterisk (*) indicates significance at the 10% level, double asterisks (**) indicate significance at the 5% level, and triple asterisks (***) indicate significance at the 1% level.

reserve the first six observations and use $s_t = \Delta y_{t-1}, \dots, \Delta y_{t-6}$ as candidate transition variables. After setting aside the last five observations to evaluate post-sample forecasts, ninety-two observations remain available for model estimation and diagnostic testing. Diagnostic test results for the fitted models are reported in the technical appendix (Balagtas and Holt 2008), as are plots of transition functions with respect to both the identified transition variable and to time.

Initial results revealed that STAR-type models were inappropriate in three instances. Specifically, for cocoa, rubber, and zinc the estimated nonlinear models fail to improve on the fit of their linear counterparts as in-

dicated by the AIC.¹⁰ In each instance this failure to improve upon the fit of the linear model seems to be a result of the identified nonlinearity stemming from a relatively small number of outliers. We therefore restrict our attention to the remaining sixteen commodities. For each fitted model the estimated speed-of-adjustment parameters, the $\hat{\gamma}_i$'s, and the

¹⁰ For these commodities cocoa is associated with a positive drift term while rubber and zinc are associated with negative drift terms. In each case, however, the drift terms are statistically insignificant at all usual levels. Likewise, results for the alternative models (i.e., the models that imply that the data are trend stationary) revealed that in each of these cases the trend term is not statistically significant. There is therefore apparently little support for the PSH for these commodities as well.

Table 2. Estimated Speed of Adjustment Parameters, γ , and Location Parameters, c , in the Transition Functions of the Estimated Models

Commodity	First Transition Function					Second Transition Function			
	Type	d	$\hat{\gamma}_1$	\hat{c}_1	\hat{c}_2	Type	d	$\hat{\gamma}_2$	\hat{c}_3
Aluminum	LSTAR	2	500	0.144 (0.003)	–	LTVAR	t	249.196 (537.020)	0.824 (0.036)
Beef	LSTAR	3	500	0.158 (0.003)	–				
Coffee	LSTAR	3	500	–0.195 (0.007)	–				
Copper	LSTAR	6	324.56	–0.095 (0.117)	–				
Cotton	ESTAR	2	4.224 (2.896)	–0.019 (0.017)	–	LTVAR	t^3	27.161 (11.454)	0.744 (0.019)
Hides	LSTAR	2	396.050	–0.073 (0.225)	–				
Lamb	ESTAR	1	0.388 (0.167)	–0.204 (0.054)	–	LSTAR	2	4.137 (2.438)	0.109 (0.047)
Lead	LSTAR	6	500	0.074 (0.001)	–				
Palmoil	LSTAR	2	24.142 (16.616)	0.175 (0.011)	–				
Silver	LSTAR	2	23.827 (21.255)	0.026 (0.011)	–	LSTAR	2	5.361 (4.152)	0.151 (0.043)
Sugar	LSTAR	1	35.612 (21.207)	–0.006 (0.012)	–	LSTAR	2	36.287 (45.144)	0.084 (0.034)
Tea	ESTAR	4	0.489 (0.707)	0.031 (0.032)	–				
Timber	ESTAR	1	0.874 (0.068)	0.068 (0.024)	–	LTVAR	t	1.823 (1.162)	0.674 (0.281)
Tin	LTVAR	3	4.143 (5.827)	0.207 (0.066)	–				
Tobacco	ESTAR	4	0.307 (0.328)	0.040 (0.021)	–	ETVAR	t	3.526 (1.806)	0.353 (0.029)
Wool	QSTAR	4	407.546	–0.070 (0.010)	0.197 (0.006)				

Note: Values in parentheses are asymptotic standard errors. LTVAR denotes a logistic transition function with time as an argument. Alternatively, ETVAR denotes an exponential transition function with time as an argument. Under the column headed d the entry t^3 denotes a restricted third-order LTVAR transition function.

estimated location parameters, the \hat{c}_i 's, along with asymptotic standard errors, are reported in table 2.

As indicated in table 2, seven of the remaining sixteen commodities are associated with a TV-STAR or three-regime STAR model of the form in (7). Specifically, aluminum, cotton, lamb, silver, sugar, timber, and tobacco require additional additive components to adequately characterize the data. Of these, four are TV-STAR models (aluminum, cotton, timber, and tobacco), characterized by both nonlinearity and structural change, while the remaining three are three-regime STAR models (lamb, silver, and sugar). Based on the estimated values of γ reported in table 2, seven commodities are associated with TAR-type specifi-

cations (aluminum, beef, coffee, copper, hides, lead, and wool). In the case of wool, the TAR-type behavior is best characterized as a QSTAR model. In no case does a TVAR model alone, that is, a model characterized simply by structural change, appear to be an entirely adequate specification. Moreover, in every instance in which parameter nonconstancy is a feature (i.e., aluminum, cotton, timber, and tobacco), parameter change is smooth.

Model Simulations to Evaluate Prebisch-Singer

While the foregoing results provide ample evidence of nonlinearity and, in some cases, parameter nonconstancy for a relatively large

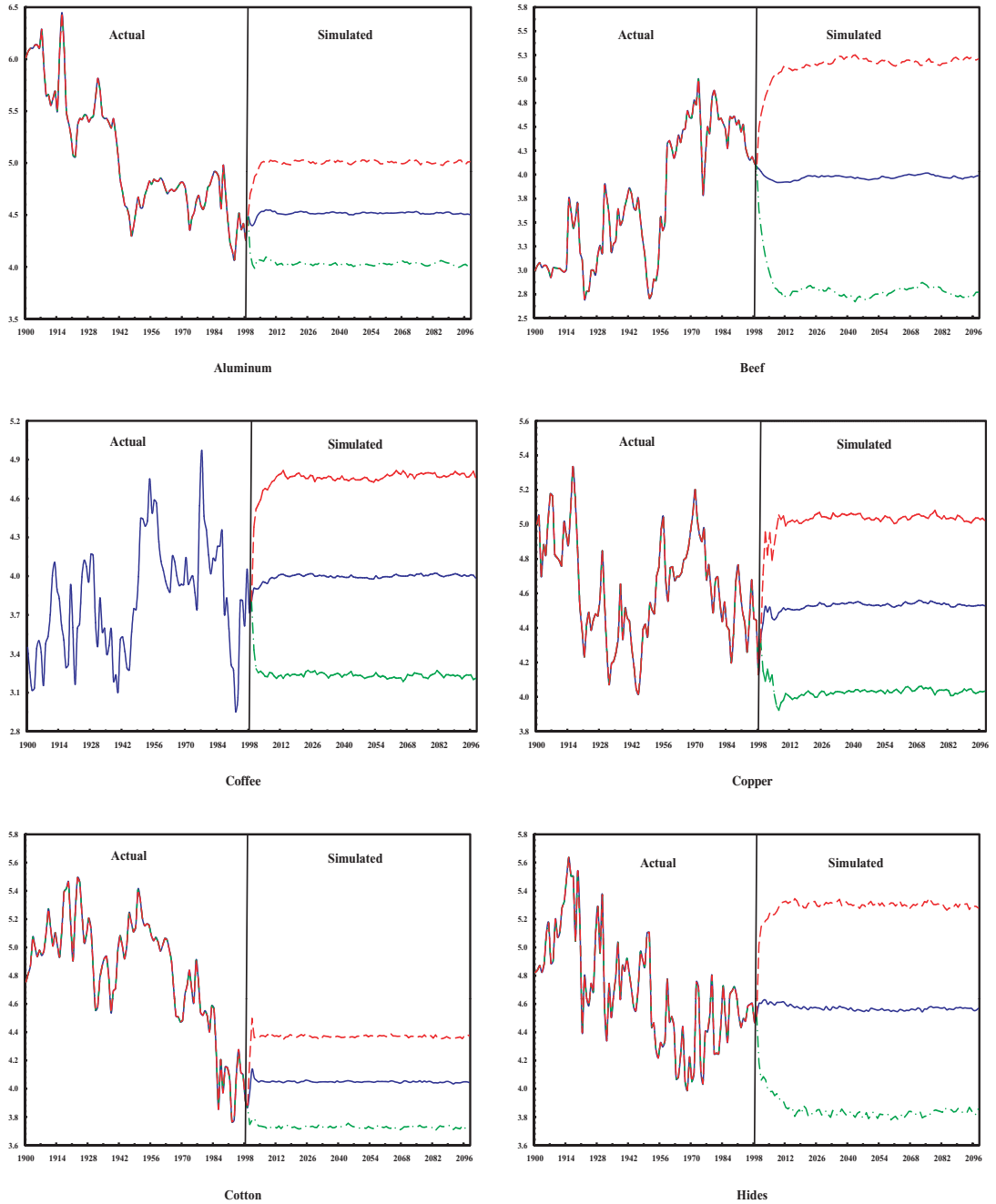


Figure 1. Actual (1900–1998) and simulated (1999–2097) commodity price data for sixteen commodities

number of commodities in the sample data, the basic question still remains. Is there evidence that the PSH holds among the commodities for which STAR-type models are fitted? While there are several ways to investigate this issue, one approach is to examine forward iterations of each model, possibly where stochastic

shocks are introduced. That is, what is required are the k -step-ahead forecasts from the estimated models using the ending points of the sample data as initial values.

Forward simulations of the estimated models provide insight into the long-run behavior of the price series—and thus a way to evaluate

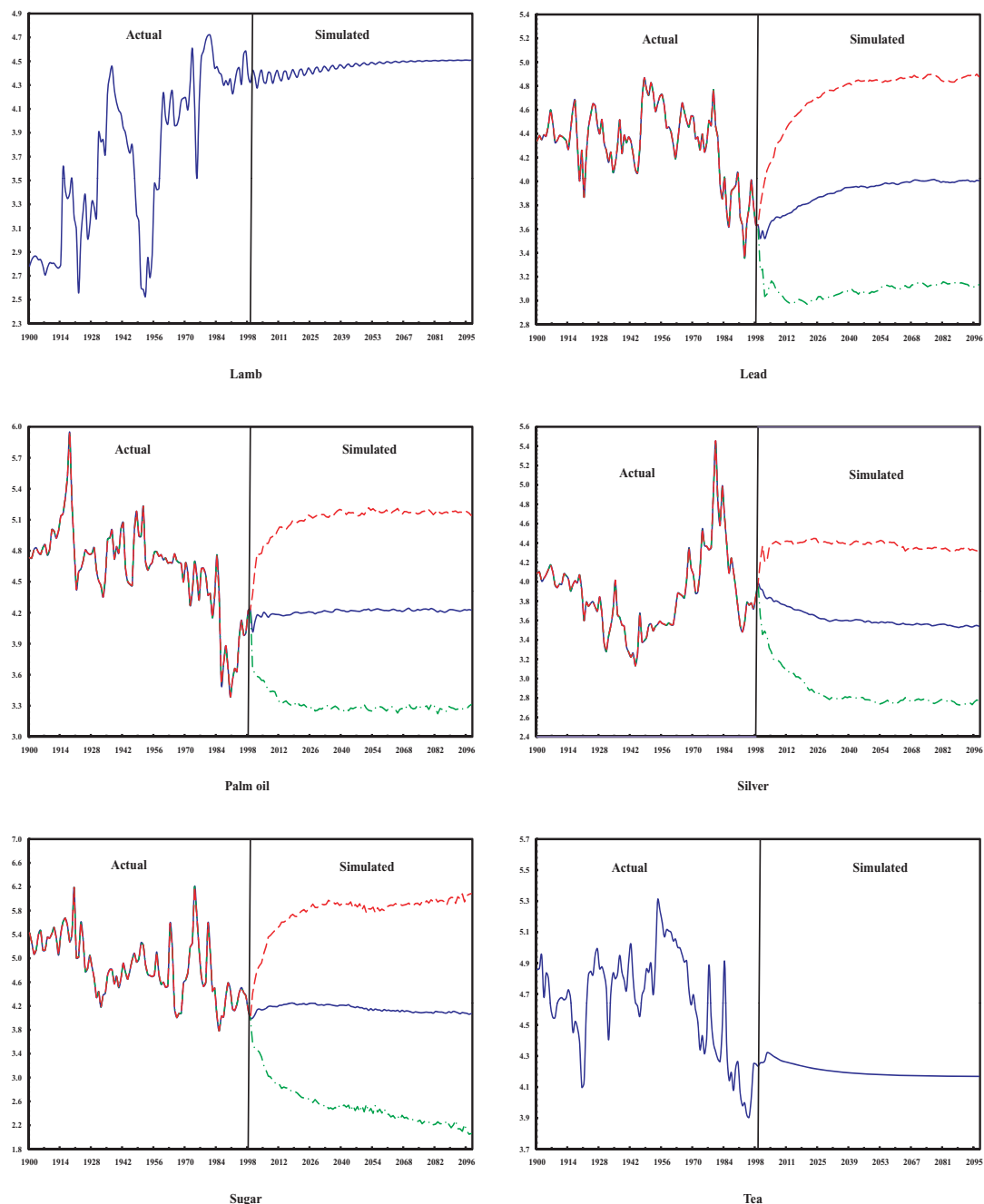


Figure 1. Continued

the PSH—among commodities for which STAR-type models have been estimated. Also, forward extrapolations of the model will reveal something about its dynamic properties. Indeed, a necessary condition for stability of an estimated STAR model is that forward iterations of its “skeleton,” that is, the forward iterations that do not include stochastic shocks

and therefore result in biased forecasts, either converge to a steady-state or a limit cycle path (Tong 1990). Alternatively, a necessary and sufficient condition for stability is that the forward iterations of the model obtained when shocks are included, that is, the unbiased forecasts, converge to a stable path (Tong 1990). In this manner useful information

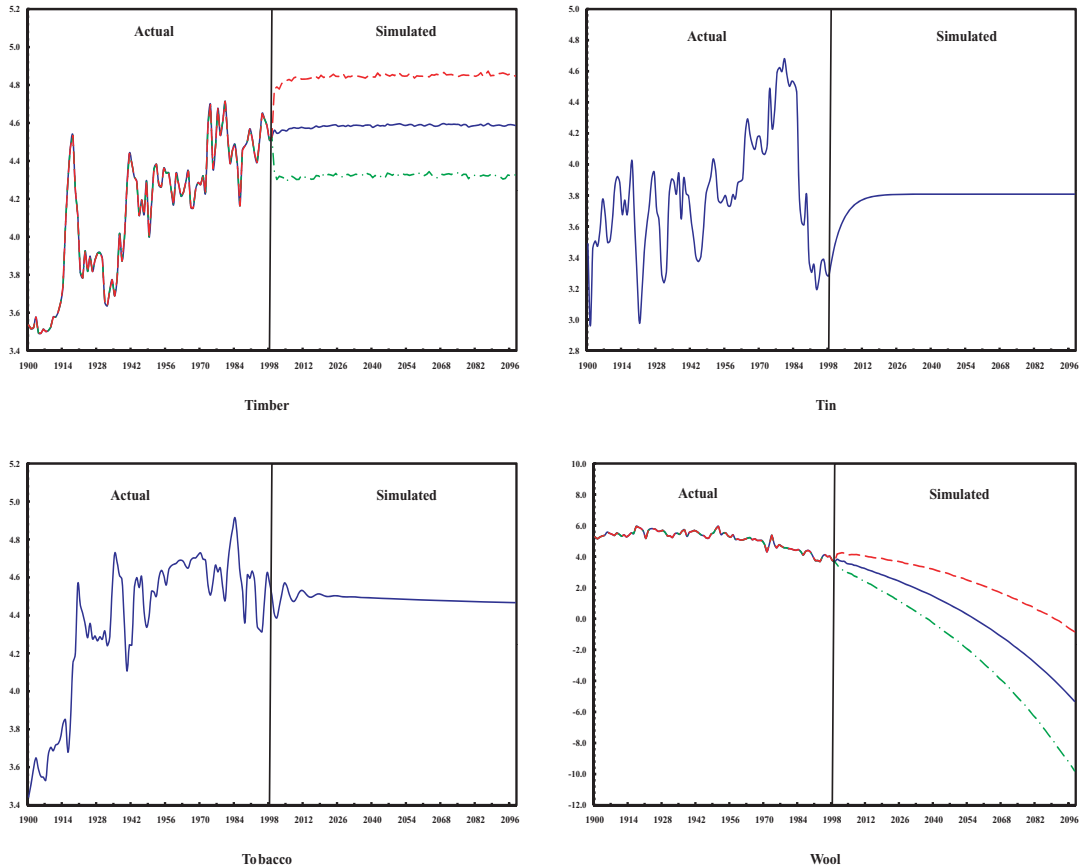


Figure 1. Continued

may be obtained about each model's dynamic properties.¹¹

The forward simulations obtained for each of the sixteen STAR-type models, along with the historical sample data, are presented in figure 1. In the case of lamb, tea, tin, and tobacco the necessary but not sufficient condition for dynamic stability is satisfied. The bootstrap simulations obtained for lamb display a slow tendency toward explosive behavior; the forecast confidence intervals move well beyond the observed range of the data rather quickly—after about thirty or so forward iterations. As illustrated in figure 1, the nonlinear dynamics associated with lamb prices are apparently dominated by a steep rise during WWII and a subsequent sharp decline in the early and mid 1950s. In any case, for this group of four commodities only the naive forecasts are plotted.

¹¹ Additional details on the bootstrap methods used to construct the forward simulations are available in the technical appendix (Balagtas and Holt 2008).

A general conclusion is that for lamb, tea, tin, and tobacco the nonlinear behavior is apparently rather extraordinary in that locally explosive behavior is a prominent feature. In the case of tin, for example, the characteristic polynomial associated with regime $G(\cdot) = 1$ has a dominant real root of 2.14. Even though this regime is very rarely observed in the data, it seems that the model is unable to break free of this dynamic once stochastic shocks are included. Similar results apply for lamb, tin, and tobacco. The overall implication is that the estimated STAR models for these commodities might be useful for short-run forecasting, but certainly not for obtaining longer-term predictions. See Hall, Skalin, and Teräsvirta (2001) for a similar example in the context of a nonlinear model of El Niño events.

For the remaining twelve commodities, that is, for aluminum, beef, coffee, copper, cotton, hides, lead, palmoil, silver, sugar, timber, and wool, plots of the bootstrap-based forecasts along with approximate $\pm 2\sigma$ confidence

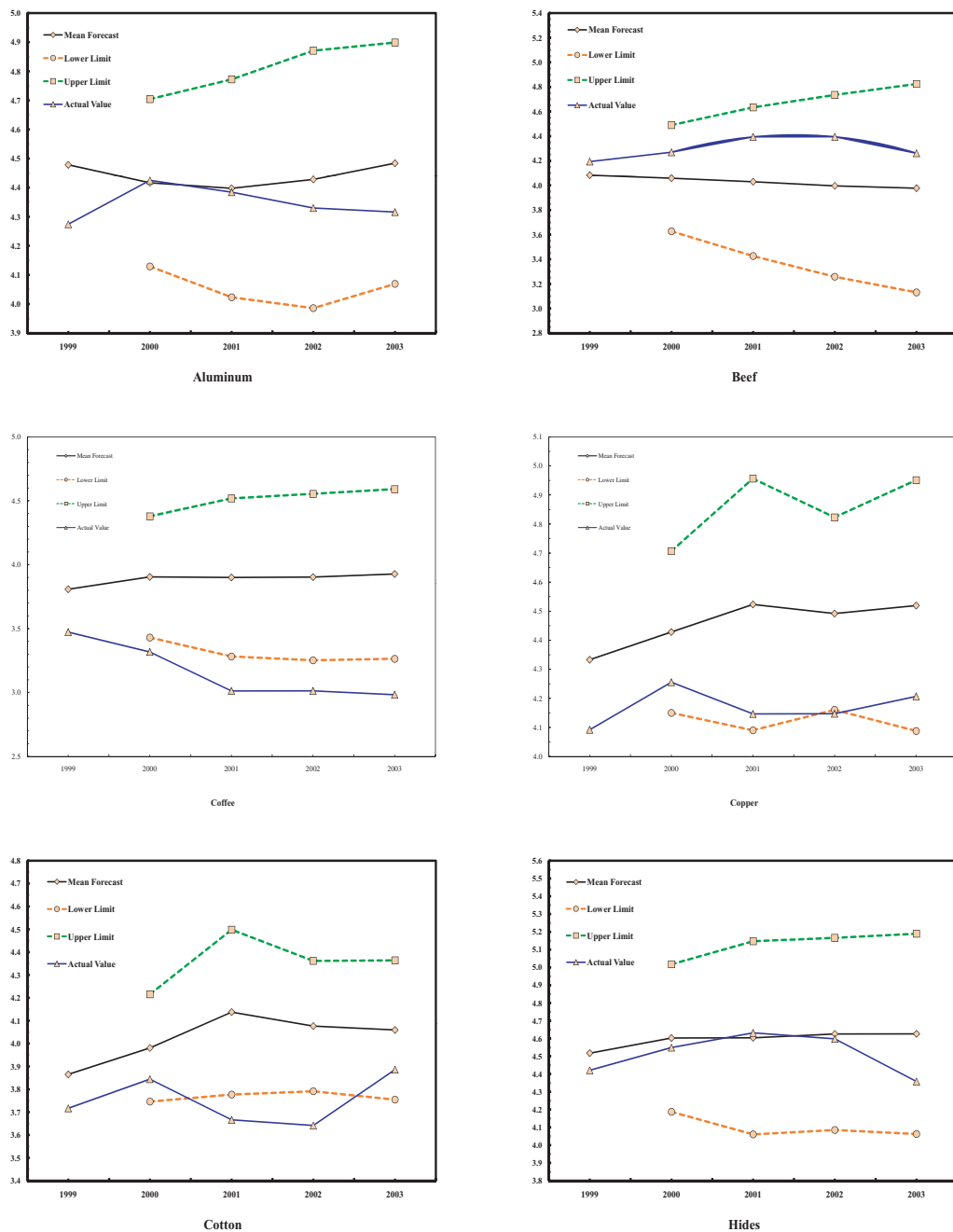


Figure 2. Actual and predicted commodity prices for sixteen commodities, 1999–2003. Dashed lines denote approximate 95% error bands. Actual values are denoted by lines with triangles and predicted values by lines with diamonds

bands are reported in figure 1.¹² The predictions (both biased and unbiased) plotted in

figure 1 reveal a lack of support for the PSH in every instance with the single exception of wool. Forward iterations of the QSTAR model estimated for wool does indeed indicate a secular decline in the terms of trade. In the case of wool both intercept (i.e., drift) terms are negative, and both the $G(\cdot) = 0$ and $G(\cdot) = 1$ regimes

¹² As noted by Teräsvirta, van Dijk, and Mederios (2005), a direct benefit of using the bootstrap-based approach to obtain forecasts is that information on the forecast density at each horizon $k \geq 2$ is readily available as a byproduct.

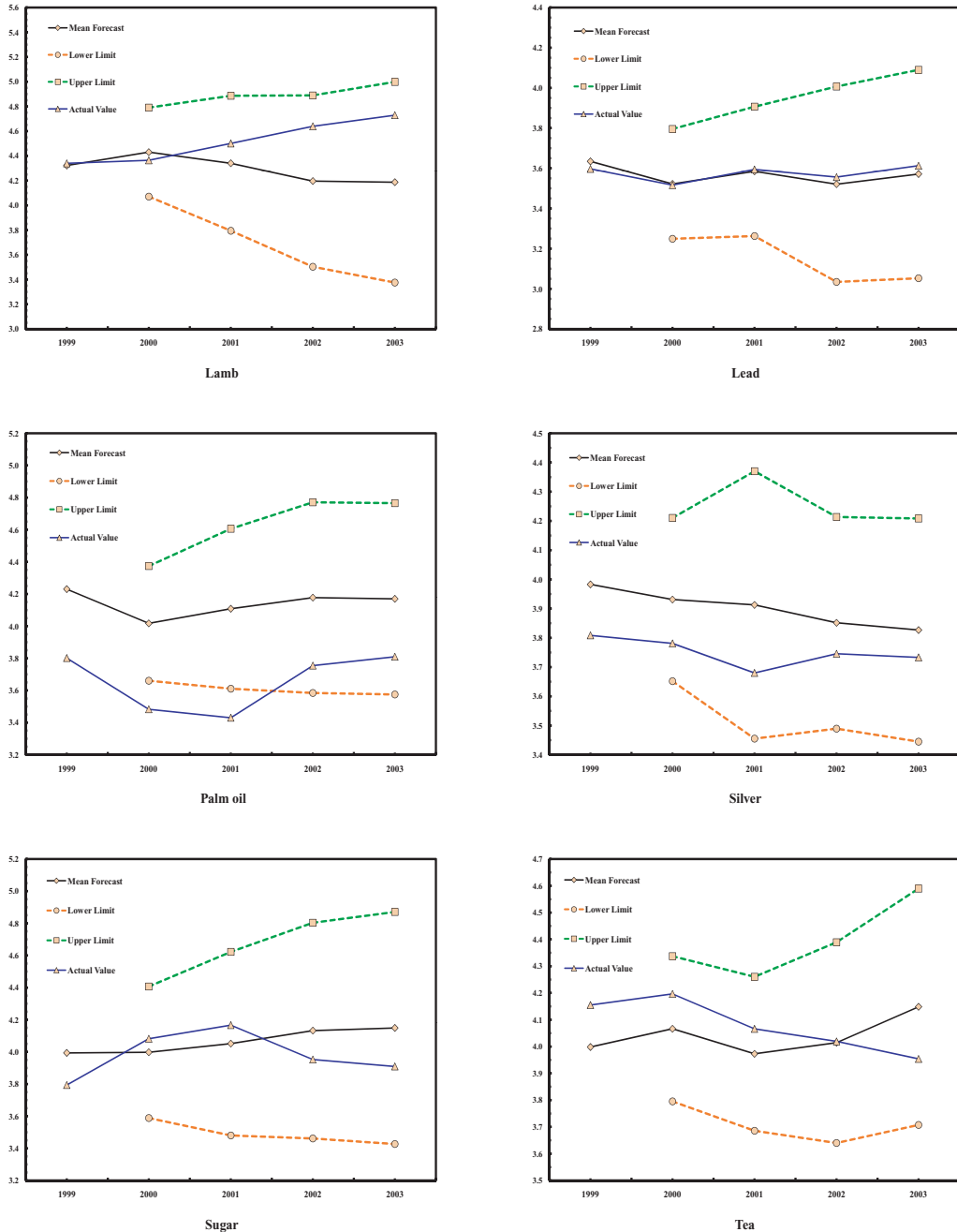


Figure 2. Continued

have real roots that are slightly greater than one. For wool, then, there is a gradual tendency for the forward simulations to drift downward, and for the confidence bands to gradually drift further apart.¹³ For all other commodities, the

forward simulations of the estimated STAR-type models provide no evidence of a continued deterioration of the terms of trade. In part these results are as expected given that the

¹³ It should be noted that this result is not inconsistent with the results reported in table 1. Importantly, the linear unit root model

may be rejected in favor of the alternative because the data are globally stationary, because the data are better characterized by nonlinearities, or both.

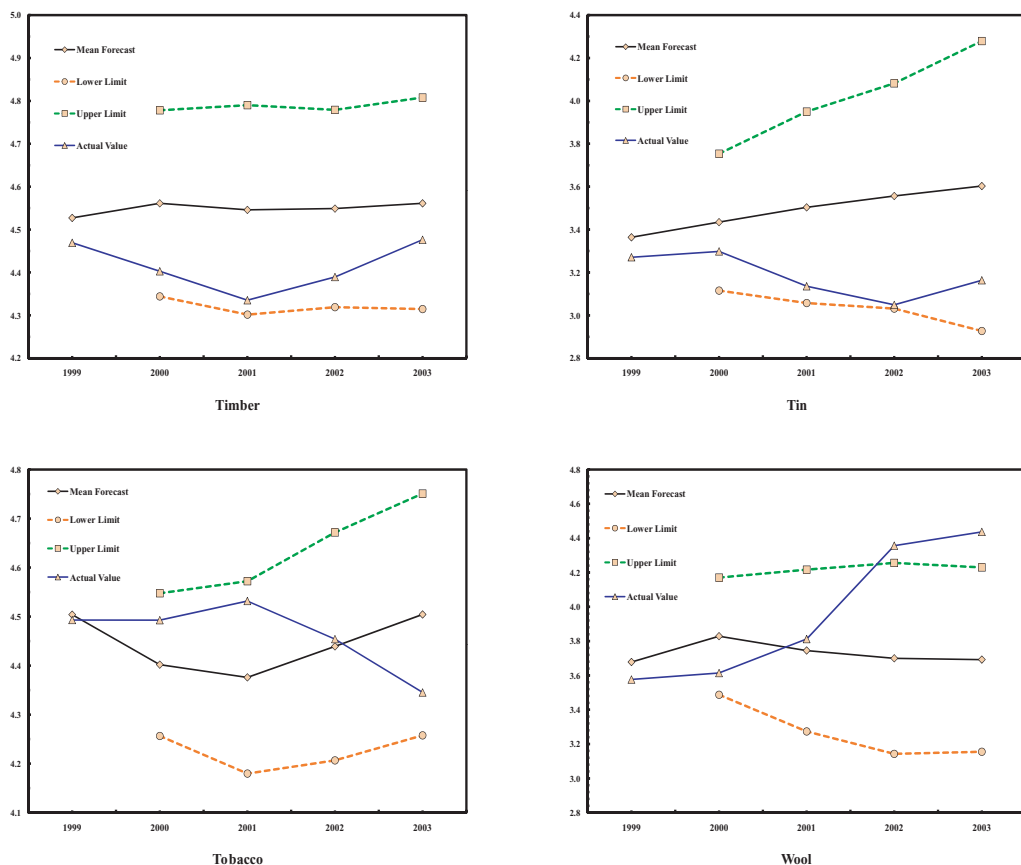


Figure 2. Continued

estimated models include lagged level terms. But even so, the TV-STAR models in particular do not require a priori that the parameter change be completed by the end of the sample period. And yet this seems to be the case for most commodities.

Finally, we evaluate the estimated models by comparing plots of the k -step-ahead forecasts for 1999–2003 against the realized values. Figure 2 presents these comparisons. The overall picture emerging from figure 2 is that the actual values tend to fall within the 95% error bands.¹⁴ A more thorough evaluation of forecasting performance might consider a longer time horizon and would compare forecasting performance across alternative models. Space limitation and data availability prohibit us from doing so here. Even so, the models gen-

erally appear to do a reasonable job of predicting post-sample commodity prices, and thus in many instances capture many of the salient features of the data observed at the end of the available period.

Conclusion

An issue of continuing interest to development and international economists is the prediction, based on the Prebisch–Singer hypothesis, that commodity prices will continue to decline relative to the price of finished or manufactured goods. Numerous studies have sought to obtain empirical evidence either for or against this basic conjecture, with work in recent years allowing for stochastic trends, one or more trend breaks, and nonlinearity. The emerging evidence based on application of models that allow richer time series behavior often militates against the PSH.

This article contributes to this literature by conducting formal tests of the linear unit root model of relative commodity prices against

¹⁴ In several cases the models systematically over- or under-predict (beef, coffee, cotton, palm oil, silver, timber, tin). Indeed, this result is not surprising given that the models are not updated and thus do not incorporate information contained in the post 1998 data, that is, the forecasts are truly k -step-ahead forecasts.

nonlinear or time-varying alternatives, and by specifying, estimating, and testing STAR-type models where called for. Surprisingly, we find that the linear unit root model is rejected in favor of a STAR-type alternative for nineteen of the twenty-four commodities investigated. Somewhat surprisingly, for none of these was a TVAR model identified. Of interest is that, with the exception of maize, jute, and rice, all of the commodities associated with a rejection of the linear unit root model were also identified previously by Kellard and Wohar (2006) as being associated with stationary trend break models. Of course Kellard and Wohar (2006) did not consider nonlinear alternatives, as we do here. Among other things it seems that with the relatively small sample sizes that are available it may be difficult to distinguish between trend stationarity (when breaks are incorporated) and nonlinearity.

Of the nineteen commodities for which linearity was rejected, we were able to successfully fit STAR-type models in sixteen instances. Each commodity price series is best characterized by a STAR or additive TV-STAR model; none is associated exclusively with time-varying parameters. Using forward simulations of the estimated models to evaluate the long-run price behavior, we find very limited support for the PSH. Only in the case of wool do we find evidence of a secular decline in the terms of trade. These results suggest caution in prescribing industrial policy based on the PSH, or, for that matter, based on any long-run predictions of relative commodity prices. In part this is because the big peaks and valleys observed in many of the commodity prices seem to be adequately characterized by nonlinearity. That said, it is essentially impossible to predict the size and direction of any future shocks, and therefore impossible to know with any precision the precise trajectory that relative commodity prices might follow in future (Cashin and McDermott 2002). For this reason it will likely remain difficult for developing countries to anticipate how and when to intervene in primary commodity markets, as well as to know which policies to pursue to enhance export earnings.

[Received October 2006;
accepted April 2008.]

References

- Ardeni, P.G., and B. Wright. 1992. "The Prebisch-Singer Hypothesis: A Reappraisal Independent of Stationarity Hypotheses." *Economic Journal* 102:803–12.
- Balagtas, J.V., and M.T. Holt. 2008. "AJAE Appendix: The Commodity Terms of Trade, Unit Roots, and Nonlinear Alternatives: A Smooth Transition Approach." Unpublished manuscript. Available at <http://agecon.lib.umn.edu/>.
- Banerjee, A., R.L. Lumsdaine, and J.H. Stoich. 1992. "Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence." *Journal of Business and Economic Statistics* 10:271–87.
- Bleaney, M., and D. Greenway. 1993. "Long-term Trends in the Relative Price of Commodities and in the Terms of Trade of Developing Countries." *Oxford Economic Papers* 45:349–63.
- Caner, M., and B.E. Hansen. 2001. "Threshold Autoregression With a Unit Root." *Econometrica* 69:1555–96.
- Cashin, P., and C.J. McDermott. 2002. "The Long-Run Behavior of Commodity Prices: Small Trends and Big Variability." *IMF Staff Papers* 49(2):175–99.
- Cuddington, J.T. 1992. "Long-Run Trends in 26 Primary Commodity Prices." *Journal of Economic Development* 39:207–27.
- Cuddington, J.T., R. Ludema, and S.A. Jayasuriya. 2006. "Prebisch-Singer Redux." In D. Lederman and W.F. Maloney, eds. *Natural Resources: Neither Curse nor Destiny*, pp. 103–40. Stanford University Press.
- Cuddington, J.T., and C.M. Urzua. 1989. "Trends and Cycles in the Net Barter Terms of Trade: A New Approach." *Economic Journal* 99:426–42.
- Davies, R.B. 1977. "Hypothesis Testing When a Nuisance Parameter Is Present Only Under the Alternative." *Biometrika* 64:247–54.
- . 1987. "Hypothesis Testing When a Nuisance Parameter Is Present Only Under the Alternative." *Biometrika* 74:33–43.
- Deaton, A. 1999. "Commodity Prices and Growth in Africa." *Journal of Economic Perspectives* 13:23–40.
- Deaton, A., and G. Laroque. 1995. "Estimating a Nonlinear Rational Expectations Commodity Price Model with Unobservable State Variables." *Journal of Applied Econometrics* 10:S9–S40.
- Dickey, D.A., and W.A. Fuller. 1979. "Distributions of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74:427–31.
- . 1981. "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root." *Econometrica* 49:1057–72.

- Eitrheim, Ø., and T. Teräsvirta. 1996. "Testing the Adequacy of Smooth Transition Autoregressive Models." *Journal of Econometrics* 74:59–75.
- Eklund, B. 2003. "Testing the Unit Root Hypothesis Against the Logistic Smooth Transition Autoregressive Model." *Stockholm School of Economics, SSE/EFI Working Paper Series in Economics and Finance*, No. 546.
- Enders, W., and C.W.J. Granger. 1998. "Unit Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of the Interest Rate." *Journal of Business and Economic Statistics* 16:304–11.
- Grilli, E.R., and 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* 2:1–47.
- Hall, A. 1994. "Testing for a Unit Root in Time Series With Pretest Data-Based Model Selection." *Journal of Business and Economic Statistics* 12:461–70.
- Hall, A.D., J. Skalin, and T. Teräsvirta. 2001. "A Nonlinear Time Series Model of El Niño." *Environmental Modelling & Software* 16:139–46.
- Holt, M.T., and L.A. Craig. 2006. "Nonlinear Dynamics and Structural Change in the U.S. Hog–Corn Cycle: A Time-Varying STAR Approach." *American Journal of Agricultural Economics* 88:215–33.
- Kellard, N., and S. Wohar. 2006. "On the Prevalence of Trends in Commodity Prices." *Journal of Development Economics* 79:146–67.
- Kim, T.-H., N.E.S. Pfaffenzeller, T. Rayner, and P. Newbold. 2003. "Testing for Linear Trend With Application to Relative Primary Commodity Prices." *Journal of Time Series Analysis* 24:539–51.
- Leon, J., and R. Soto. 1997. "Structural Breaks and Long-Run Trends in Commodity Prices." *Journal of International Development* 3:44–57.
- Li, H., and G.S. Maddala. 1996. "Bootstrapping Time Series Models." *Econometric Reviews* 15:115–58.
- Lin, C.-F.J., and T. Teräsvirta. 1994. "Testing the Constancy of Structural Parameters Against Continuous Structural Change." *Journal of Econometrics* 62:211–28.
- Lipsey, R.E. 1994. "Quality Change and Other Influences on Measures of Export Prices of Manufactured Goods and the Terms of Trade Between Primary Products and Manufacturers." NBER Working Paper #4671.
- Lumsdaine, R.L., and D. Papell. 1997. "Multiple Trend Breaks and the Unit Root Hypothesis." *Review of Economics and Statistics* 79:212–18.
- Lundbergh, S., T. Teräsvirta, and D. van Dijk. 2003. "Time-Varying, Smooth Transition Autoregressive Models." *Journal of Business and Economic Statistics* 21:104–21.
- Luukkonen, R., P. Saikkonen, and T. Teräsvirta. 1988. "Testing Linearity Against Smooth Transition Autoregressive Models." *Biometrika* 75:491–99.
- MacKinnon, J. 2002. "Bootstrap Inference in Econometrics." *Canadian Journal of Economics* 35:615–45.
- Newbold, P., and D. Vougas. 1996. "Drift in the Relative Price of Primary Commodities: A Case Where We Care About Unit Roots." *Applied Economics* 28:654–61.
- Ocampo, J.A., and M.A. Parra. 2004. "The Terms of Trade for Commodities in the Twentieth Century." *Econ Working Paper Archive*, International Trade Series No. 0402006.
- Perron, P. 1989. "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis." *Econometrica* 57:1361–401.
- . 1997. "Further Evidence on Breaking Trend Functions in Macroeconomic Variables." *Journal of Econometrics* 80:355–85.
- Persson, A., and T. Teräsvirta. 2003. "Net Barter Terms of Trade: A Smooth Transition Approach." *Journal of Financial Economics* 8:81–97.
- Pfaffenzeller, N.E., P. Newbold, and A. Rayner. 2007. "A Short Note on Updating the Grill and Yang Commodity Price Index." *World Bank Economic Review* 21(1):151–63.
- Powell, A. 1991. "Commodity and Developing Country Terms of Trade: What Does the Long Run Show?" *Economic Journal* 101:1485–96.
- Prebisch, R. 1950. *The Economic Development of Latin America and its Principal Problems*. New York: United Nations Publications.
- Sapsford, D. 1985. "The Statistical Debate on the Net Barter Terms of Trade Between Primary Commodities and Manufacturers: A Comment and Some Additional Evidence." *Economic Journal* 95:781–88.
- Singer, H. 1950. "The Distributions of Gains Between Investing and Borrowing Countries." *American Economic Review* 40:473–85.
- Spraos, J. 1980. "The Statistical Debate on the Net Barter Terms of Trade Between Primary Commodities and Manufacturers." *Economic Journal* 90:107–28.
- Teräsvirta, T. 1994. "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models." *Journal of the American Statistical Association* 89:208–17.
- Teräsvirta, T., D. van Dijk, and M.C. Medeiros. 2005. "Linear Models, Smooth Transition

- Autoregressions, and Neural Networks for Forecasting Macroeconomic Time Series: A Re-examination." *International Journal of Forecasting* 21:755–44.
- Thirlwall, A.P., and J. Bergevin. 1985. "Trends, Cycles, and Asymmetries in the Terms of Trade of Primary Commodities from Developed and Less Developed Countries." *World Development* 13:805–17.
- Tong, H. 1990. *Nonlinear Time Series: A Dynamical Systems Approach*. Oxford: Oxford University Press.
- van Dijk, D., and P.H. Franses. 1999. "Modeling Multiple Regimes in the Business Cycle." *Macroeconomic Dynamics* 3:311–40.
- van Dijk, D., T. Teräsvirta, and P.H. Franses. 2002. "Smooth Transition Autoregressive Models—A Survey of Recent Developments." *Econometric Reviews* 21:1–47.
- Wang, D., and W.G. Tomek. 2007. "Commodity Prices and Unit Root Tests." *American Journal of Agricultural Economics* 89:873–89.
- Zanias, G.P. 2005. "Testing for Trends in the Terms of Trade Between Primary Commodities and Manufactured Goods." *Journal of Development Economics* 78:49–59.
- Zivot, E., and D.W.K. Andrews. 1992. "Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." *Journal of Business and Economic Statistics* 10:251–70.