NBER WORKING PAPER SERIES

NONRENEWABLE RESOURCE PRICES: DETERMINISTIC OR STOCHASTIC TRENDS?

Junsoo Lee
John A. List
Mark Strazicich

Working Paper 11487
http://www.nber.org/papers/w11487

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2005

Corresponding author: Mark Strazicich, Department of Economics, Appalachian State University, Boone, NC, 28608. Telephone: 828-262-6124. Fax: 828-262-6105. Email: strazicichmc@appstate.edu. We thank Ashley Ahrens and Vijaya Sharma for access to their data. Charles Mason provided valuable comments on an earlier version of this paper. We are grateful to the editor and two anonymous referees for helpful comments and suggestions. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Nonrenewable Resource Prices: Deterministic or Stochastic Trends?
Junsoo Lee, John A. List and Mark Strazicich
NBER Working Paper No. 11487
July 2005
JEL No. Q31, C12, C53

ABSTRACT

In this paper we examine temporal properties of eleven natural resource real price series from 1870-1990 by employing a Lagrangian Multiplier unit root test that allows for two endogenously determined structural breaks with and without a quadratic trend. Contrary to previous research, we find evidence against the unit root hypothesis for all price series. Our findings support characterizing natural resource prices as stationary around deterministic trends with structural breaks. This result is important in both a positive and normative sense. For example, without an appropriate understanding of the dynamics of a time series, empirical verification of theories, forecasting, and proper inference are potentially fruitless. More generally, we show that both pre-testing for unit roots with breaks and allowing for breaks in the forecast model can improve forecast accuracy.

Junsoo Lee
Department of Economics, Finance and Legal Studies
255 Alston Hall
Box 870224
University of Alabama
Tuscaloosa, AL 35487
jlee@cba.ua.edu

John A. List
Department of Economics and AREC
2200 Symons Hall
University of Maryland
College Park, MD 20742
and NBER
jlist@arec.umd.edu

Mark C. Strazicich
Department of Economics
Appalachian State University
Boone, NC 28608
strazicichmc@.appstate.edu

I. Introduction

An important literature has developed recently that empirically examines nonrenewable natural resource price paths and investigates whether they are trend- or difference-stationary. For example, Ahrens and Sharma (1997) use annual data on eleven commodity price series ranging from 1870-1990 and conclude that six of these series are stationary around a deterministic trend, while the remaining five display stochastic trends implying a unit root. In a related paper, Berck and Roberts (1996) use a subset of the same data and find overwhelming support for non-stationary unit roots. Slade (1988) and Abgeyegbe (1993), among others, find similar overwhelming support for non-stationary natural resource prices. A consistent theme across much of this literature is that most natural resource prices are non-stationary.

Our motivation in this paper is to further the boundaries of econometric methodologies and provide new insights into natural resource price time paths. Previous empirical work on natural resource prices typically neglects possible structural change in the time series. Since the seminal work of Perron (1989), it is well known that ignoring structural change in unit root tests will lead to a bias against rejecting the unit root null hypothesis when it should in fact be rejected (e.g., see also Amsler and Lee, 1995). In this study, we advance the literature on time paths of natural resource prices by endogenously determining structural breaks and extending the two-break Lagrangian Multiplier (LM) unit root test of Lee and Strazicich (2003) to include a quadratic trend. Given that a quadratic trend might exist in some natural resource price series we believe that allowing for a quadratic trend in conjunction with structural breaks may provide additional insights.

Understanding the nature of resource price time paths is important for several reasons. Theoretically speaking, Ahrens and Sharma (1997, page 61), for example, note that in regards to both a simple and more general Hotelling (1931) model as described in Slade (1988), "price movement is still systematic and may be modeled appropriately as a deterministic trend." In contrast, in a world with uncertainty "in which speculative motives drive the behavior of extracting firms or unanticipated events largely characterize the market, resource prices may be generated by a random walk process" (Ahrens and Sharma, 1997, pages 61-62). Thus, knowing the correct time series behavior of natural resource prices can be vital to distinguish among theories that most accurately describe observed behavior.

Knowledge of the time series properties of natural resource prices is also important for proper econometric estimation. For example, Ahrens and Sharma (1997) and Labson and Crompton (1993) note that conventional regression analysis and hypothesis testing cannot be correctly undertaken without first understanding characteristics of the time series. Otherwise, results from estimating regression models may be rendered invalid. Several additional examples can be found in the literature (see, e.g., Hamilton, 1996).

Finally, given that good policymaking typically depends on sound economic forecasts, appropriately modeling the nature of the time series can be invaluable to forecasters (Diebold and Senhadji, 1996). This task recently came to hand in Berck and Roberts (1996), who did not consider structural breaks. They found that each commodity price series had a unit root and hence initially paid more credence to their ARIMA forecasts. Berck and Roberts (1996), however, found that their ARMA forecasts outperformed their ARIMA counterparts. In the later part of our paper, we investigate whether pre-testing for unit roots with structural change can help to identify the most accurate forecasting model. This question has not been previously examined in the literature.

Our investigation begins by examining annual data comprised of eleven fuel and metal real price series ranging from 1870-1990.[1] After including structural breaks, we find evidence *against* a unit root in each of the eleven series. We then estimate different forecasting models and find that ARMA models with breaks generally outperform ARIMA models with breaks, which is consistent with our pre-test expectations. These results help to solve the puzzling results noted by Berck and Roberts (1996), and demonstrate the importance of considering structural breaks in economic forecasting. A second major finding is that our unit root test results are robust-with or without quadratic trends natural resource prices are stationary around deterministic trends with two structural breaks in intercept and trend slope. Thus, while nonlinear time trends are important in certain price series, rejection of the underlying unit root hypothesis and support for trend-break stationary price series does not depend on inclusion of such trends. Finally, we find that both pre-testing for unit roots and including structural breaks can improve the accuracy of forecasting natural resource prices.

The remainder of the paper is structured as follows. Section II provides a brief background on unit root tests and further describes the importance of understanding the time-series properties of exhaustible resource prices. Section III describes our empirical

[1] This data was utilized to most accurately compare our empirical findings to the previous works. In section V, we estimate a variety of forecasting models with updated data through 2002.

methodologies. Section IV presents our empirical results. In Section V, we examine how structural breaks affect the accuracy of forecasts. Section VI concludes.

II. Background

Recent years have witnessed an explosion of research that examines the time-series properties of economic data. An important stimulant in this time-series renaissance was Nelson and Plosser's (1982) study, which applied unit root analysis to test for the stationarity of macroeconomic and financial time series. Since this seminal study, many authors have analyzed data ranging from stock prices to air pollutant emissions, with the bulk of research having clear implications theoretically as well as from a policy perspective (see, e.g., List, 1999). Certainly, the empirical results from many of these studies have broad implications in numerous areas of research.

Within the natural resource literature, Slade (1988), Berck and Roberts (1996), and Ahrens and Sharma (1997), among others, discuss the notion that exogenous shocks can affect the time path of natural resource prices. These studies use a variety of unit root tests to examine time paths of commodity prices. Rejection of the unit root null supports the alternative hypothesis of a mean or trend reverting stationary series, implying that shock effects are transitory. Alternatively, failure to reject the unit root null implies a non-stationary series in which shocks have permanent effects; following a shock there is no tendency for commodity prices to revert to a stable mean or trend. The most recent of these studies—Ahrens and Sharma (1997)—assumes one known or exogenously given structural break common to all commodity price series in 1929, 1939, or 1945; the other studies mentioned do not consider structural breaks.

Natural resource theory has stark predictions on the time pattern of resource prices. In fact, in his treatise on the economics of exhaustible resources, Hotelling (1931) predicts that, under certain assumptions, the price of an exhaustible resource will rise at the rate of interest. But a quick glance at the majority of mineral commodity price series suggests that relative commodity prices have *declined*, rather than increased, for long periods of time. Barnett and Morse (1963), who observed this seemingly anomalous pattern, concluded that scarcity was not a real problem. Although these particular results are crucial evidence against the original premise of Hotelling, modifications of the simple Hotelling model can produce predictions of falling or stagnant prices over time. For example, environmental constraints and natural resource abundance may induce price declinations (Berck and Roberts, 1996). Resource prices may also decrease when a backstop technology is introduced, causing an inward shift in the demand for natural resources (e.g., Heal, 1976). Technical change and an endogenous change also could produce decreasing, or U-shaped paths for relative prices (e.g., Slade, 1982). Finally, modifications of initial informational assumptions, such as knowledge of the original resource stock, can induce the model to predict price decreases (see Pindyck, 1980, for one scenario).

III. A New Modeling Approach

In this paper we begin our testing approach by employing the recently developed two-break LM unit root test of Lee and Strazicich (2003). One important advantage of the LM unit root test is that it is free of spurious rejections. Nunes, Newbold, and Kuan (1997) and Lee and Strazicich (2001) showed that the endogenous break ADF-type unit root tests are subject to spurious rejections in the presence of a unit root with break.[2] Because these tests assume no breaks under the null and derive their critical values under this assumption, rejection of the null need not imply rejection of a unit root per se, but may imply rejection of a unit root *without* break. As such, researchers may incorrectly conclude that a time series is (trend) stationary with break(s) when in fact the series is non-stationary with break(s). In contrast, rejection of the null using the LM test is unaffected by breaks. Thus, rejection of the null using the LM test unambiguously indicates a trend-stationary series with break(s).[3]

To illustrate the underlying model and LM testing procedure, we consider the following data-generating process (DGP):

$$y_t = \delta' Z_t + e_t \qquad (1)$$

$$e_t = \beta e_{t-1} + u_t, \qquad (2)$$

where y_t is the commodity price in period t, δ is a vector of coefficients, Z_t is a matrix of exogenous variables, and u_t is an error term. We define $Z_t = [1, t, D_{1t}, D_{2t}, DT_{1t}, DT_{2t}]$ to allow for a constant term, linear time trend, and two structural breaks in level and trend, where T_{Bj} denotes the time period of the breaks. Under the trend-break stationary alternative, the D_{jt} terms describe an intercept shift in the deterministic trend, where D_{jt} =

[2] The endogenous break ADF-type unit root tests include the one-break minimum test of Zivot and Andrews (1992) and two-break minimum test of Lumsdaine and Papell (1997). These tests have been popular in the literature. However, as shown in Lee and Strazicich (2001, 2003), these tests tend to estimate the break point incorrectly at one period prior to the true break(s) where bias in estimating the unit root test coefficient is greatest. The outcome is a size distortion that increases with the magnitude of the break. Due to a different detrending procedure in the endogenous break LM test, this type of size distortion does not occur.

[3] As is common practice in the literature, we refer to a time series that is stationary around a breaking trend as "trend-break stationary."

1 for $t \geq T_{Bj} + 1, j = 1, 2$, and zero otherwise; DT_{jt} describes a change in slope of the deterministic trend, where $DT_{jt} = t$ for $t \geq T_{Bj} + 1, j = 1$, 2, and zero otherwise.

While Berck and Roberts (1996) considered an LM unit root test with quadratic trend, they did not consider structural breaks. Ahrens and Sharma (1997) advance the literature by considering one structural break that is given *a priori*. To further advance the literature, we consider two structural breaks that are *endogenously determined* by the data. In addition, we extend the two-break LM unit root test to allow for a quadratic trend. Defining Z_t appropriately, the DGP in (1) with quadratic trend and two breaks in intercept and trend slope can be described by $Z_t = [1, t, t^2, D_{1t}, D_{2t}, DT_{1t}, DT_{2t}]$. To the best of our knowledge, this is the first paper in any field of research to consider a quadratic time trend in an endogenous break unit root test.[4]

Test statistics for the LM unit root test are obtained from the following regression:

$$\Delta y_t = \delta' \Delta Z_t + \phi \tilde{S}_{t-1} + u_t , \tag{3}$$

where $\tilde{S}_t = y_t - \tilde{\psi}_x - Z_t\tilde{\delta}$, $t=2,..,T$, and $\tilde{\psi}_x = y_1 - Z_1\tilde{\delta}$. $\tilde{S}_t$ is a detrended series of y_t using the coefficients in $\tilde{\delta}$, which are estimated from the regression in first differences of Δy_t on $\Delta Z_t = [1, \Delta D_{1t}, \Delta D_{2t}, \Delta DT_{1t}, \Delta DT_{2t}]$. This procedure follows from the LM (score) principle, which imposes the null restriction. Subtracting $\tilde{\psi}_x$ makes the detrended series begin at zero so that $\tilde{S}_1 = 0$. This detrending method based on $\tilde{\delta}$ differs from that adopted in the DF type tests, where the parameter δ is estimated from the regression of y_t on Z_t in

[4] There are technical difficulties in obtaining relevant asymptotic distributions and corresponding critical values of endogenous break unit root tests with three or more breaks. However, for our purposes, including three or more breaks does not appear necessary. First, visual inspection of the data reveals that more than two breaks are unlikely. Second, the two-break test is found to have sufficient ability to reject the unit root in all cases.

levels. Vougas (2003) has shown that the LM type test using the above detrending method is more powerful than the DF type test.

Under the null hypothesis of a unit root $\phi = 0$, and under the alternative $\phi < 0$. The LM unit root test statistic is denoted as:

$$\tilde{\tau} = t\text{-statistic testing the null hypothesis } \phi = 0. \quad (4)$$

The search for two breaks ($\lambda_j = T_{Bj}/T$, j=1,2) that minimize $\tilde{\tau}$ can be described by a grid search as follows:[5]

$$LM_\tau = \inf_{\lambda} \tilde{\tau}(\lambda). \quad (5)$$

Correction for autocorrelated errors in (3) is accomplished by including augmentation terms $\Delta\tilde{S}_{t-j}$, $j=1,..,k$, as in the standard ADF test. To determine k we follow the "general to specific" procedure described in Perron (1989). Beginning with a maximum number of $\Delta\tilde{S}_{t-i}$ terms, $maxk$ = 8, we examine the last augmented term $\Delta\tilde{S}_{t-8}$ for significance at the 10% level (asymptotic normal critical value is 1.645). If insignificant, the last augmented term is dropped from the regression and the model is re-estimated using k = 7 terms, etc., until the maximum lagged term is found or k = 0, at which point the procedure ends. We repeat this procedure at each combination of two break points $\lambda = (\lambda_1, \lambda_2)'$ over the time interval [.1T, .9T] (to eliminate end points).[6]

IV. Empirical Results

[5] Note that the break locations and unit root test statistic are jointly estimated in the minimum LM unit root test. Critical values were derived given this joint estimation (Lee and Strazicich, 2003).
[6] This "general to specific" procedure has been shown to perform well as compared to other data-dependent procedures in selecting the optimal number of augmented terms (see, e.g., Ng and Perron, 1995).

Our data, generously provided by W. Ashley Ahrens and Vijaya Sharma, are annual real price series for eleven fuels and metals. The series are typically available from 1870-1990 and include aluminum, bituminous coal, copper, iron, lead, nickel, petroleum, natural gas, silver, tin, and zinc.[7] To allow a direct comparison with previous research, we use commodity price series similar to those examined in Slade (1988) and Berck and Roberts (1996), and, of course, identical to the data used in Ahrens and Sharma (1997).[8] Visual inspection of each time series in Figure 1 suggests that most are stationary with one or two structural breaks. In a few series, there is evidence suggesting a possible nonlinear (quadratic) trend. Curvature of trend would be consistent with Hotelling's basic premise. Yet, in Figure 1, the trends are not typically convex with a positive first derivative—a majority of the figures show that relative commodity prices have *declined* rather than increased. This finding matches results in previous studies and casts doubt on some predictions of the simple Hotelling (1931) model.

Although casual inspection of Figure 1 provides some insights, formal econometric tests are necessary to determine the properties of each time series. In our analysis, we strive for the most general results possible. Accordingly, we employ five different unit root tests to analyze each price series. Our first test, also employed by Berck and Roberts (1996), is the no-break LM unit root test of Schmidt and Phillips (1992, SP hereafter) with an added quadratic time trend. Our second test is the two-break LM unit root test developed in Lee and Strazicich (2003). To check the robustness of our

[7] Price series for aluminum, iron, and petroleum are available through 1984, 1973, and 1989, respectively. Aluminum, gas, nickel, and tin begin in 1895, 1919, 1913, and 1885, respectively.

[8] These data were previously converted to real prices by deflating with the producer price index (1967 = 100). Most of the data comes originally from Manthy (1978) and Schurr (1960).

findings, and more clearly examine the effect of including two breaks instead of one, in our third test we utilize the one-break minimum LM unit root test of Lee and Strazicich (2004). In our most general empirical model we extend the two-break LM unit root test to include a quadratic time trend. Finally, we extend the one-break LM test to include a quadratic trend. Our combination of testing procedures permits us to compare our findings with previous studies, while simultaneously allowing us to isolate effects of allowing for breaks.

Quadratic Trend with No Break

Results using the no-break SP LM unit root test with linear and quadratic time trends are presented in Table 1. The second column provides sample sizes; the third column notes the starting year; the fourth column presents the number of lagged augmentation terms included in each regression. The resulting *t*-statistics testing the unit root null hypothesis are presented in the fifth column. The rightmost column contains summary estimation results of equation (3)—estimated coefficients of the augmented terms are omitted to save space. Critical values are contained in the Table notes.

The test results in Table 1 suggest that five of the eleven real price series reject the unit root null at the $p < .05$ level (coal, iron, lead, gas, and zinc). Interestingly, four of these five series (all except lead) have a quadratic time trend that is significantly different from zero at conventional levels, whereas the six series that do not reject the unit root have insignificant quadratic time trend terms. Thus, non-rejection of a unit root hypothesis for these six series may be due to lower power caused by including an insignificant quadratic time trend, or to bias from omitting breaks. Consistent with Ahrens and Sharma (1997; their Table IV), our results reject the unit root in nearly half the series; but our rejections do not overlap completely. One explanation for this finding is our inclusion of a quadratic time trend—Ahrens and Sharma (1997) do not include a quadratic trend in their ADF tests. Ahrens and Sharma (1997) also employ the test suggested by Ouliaris et al. (1989, OPP hereafter), which includes a quadratic time trend but no structural break. Our results still do not exactly replicate their findings. One possible explanation, which is noted by Ahrens and Sharma (1997), is that the OPP test is plagued by size distortions. Fortunately, similar size distortions do not occur in the no-break SP LM test and may help to explain the difference in results.

Comparing our results with those in Berck and Roberts (1996), which use the same *no-break* SP LM test with quadratic trend, we find that inferences across the two sets of empirical results differ significantly. Whereas we reject the unit root in five price series, they cannot reject the unit root in any of their nine series at the 5% level. This difference in results is most likely due to the longer time series in our tests, leading to greater power to reject the null—whereas we make use of the entire data set, Berck and Roberts (1996) use two smaller subsets of the data (i.e., 1940-76 and 1940-91). Using these shorter time series, we were able to replicate the results in Berck and Roberts (1996). We should note that replication was achieved even though they used a transformed test based on estimating the long-run variance, whereas we employed an augmented version of the no-break SP LM unit root test. This provides evidence that the difference in results is not due to different methods of correcting for autocorrelated errors.

Two Endogenous Breaks with a Linear Trend

Moving to our more flexible models, we present empirical results from the two-break LM unit root test in Table 2. Critical values are provided in the Table notes. Empirical results may well reflect the reduction in bias from including variables to model structural change. Compared to the no-break SP LM test results in Table 1, we find significantly more rejections of the unit root null. Whereas the no-break model in Table 1 rejects the unit root null for five of eleven series at the $p < .05$ level, the two-break model in Table 2 rejects the unit root null in eight of eleven series at $p < .05$. If one considers a slightly less stringent $p < .10$ test, the two-break model rejects the unit root null in all eleven series. These findings provide the strongest evidence to date against the unit root process as a description of natural resource prices.

Our test results in Table 2 strengthen the findings of Ahrens and Sharma (1997), and differ from the conclusions of Agbeyegbe (1993), Berck and Roberts (1996), and Slade (1988), among others. Whereas Ahrens and Sharma (1997) allow for one known break in 1929, 1939, or 1945 and reject the unit root in five of eight price series, after allowing for two endogenously determined breaks we reject the unit root in all eleven series. By including one (known) structural break Ahrens and Sharma (1997) can be seen as an extension of Berck and Roberts (1996), while our tests can be seen as a further extension to a more general model. [9]

[9] An anonymous referee notes that the shorter time span (1940-1991) examined in Berck and Roberts (1996) can be viewed as allowing for a structural break in 1940. However, as our results will demonstrate, while World War II is one of the major structural breaks identified by our tests, there are other breaks that we identify. By considering a longer time span of data, our testing procedure should benefit from greater power and a reduction in bias from including additional breaks. It is also possible that failure to reject the unit root null in Berck and Roberts (1996) might be due to their inclusion of a quadratic trend. If a quadratic trend does not occur, especially in the post-1940 time period, then inclusion would lower power. We thank an anonymous referee for noting this possibility.

To further examine the effect of allowing for two breaks instead of one, we additionally test all price series using the endogenous one-break LM test. The results are displayed in Table 3. Compared to the two-break results, we find fewer rejections of the unit root using the one-break test (seven of eleven series reject the null at $p < .05$ and $p < .10$). This difference in results is likely due to bias from omitting the second break variable, and demonstrates the impact of allowing for two breaks instead of one. [10]

Comparing the breaks identified by our two-break LM tests in Table 2 with those in Ahrens and Sharma (1997), only one of our estimated break years (1945, lead) coincides exactly with an assumed break year selected by Ahrens and Sharma (1997). However, five additional breaks identified by our two-break tests (1928, 1932, 1942, 1943, and 1944) are close to one of their assumed breaks (1929, 1939, and 1945). Overall, our identified break years in Table 2 are intuitively appealing as they broadly correspond to years associated with World War I, the Great Depression, World War II, and the energy crisis of the early 1970s. Fifteen of the twenty-two breaks (68%) correspond to these time periods, with the most common time period for breaks in the early 1970s. Besides these obvious break years, other estimated breaks occur concurrently with important events. For example, the model shows that a structural break occurred in 1902 for coal prices. This year coincides with the 100,000 United Mine Worker strike that crippled coal production in the U.S. for five months.

Two Endogenous Breaks with a Quadratic Trend

---

[10] For comparison purposes, we also performed unit root tests using the ADF-type one-break Zivot and Andrews (1992) test and two-break Lumsdaine and Papell (1997) test. While these tests tend to over-reject the null, their results are similar to ours. The two-break LP test rejects the null at the 10% level for all series except for gas and petroleum. Details of these results are available upon request.

In Table 4 we present results from our most general model, the two-break LM unit root test with a quadratic time trend. Since this test is new to the literature, we follow the procedure of Lee and Strazicich (2003) and simulate critical values, which are presented in the Table notes.[11] Compared to the two-break test results in Table 2, which include only a linear trend, we report a few significant differences. Whereas exclusion of the quadratic time trend in Table 2 leads to rejection of the unit root at the 10% level for all eleven series (8 of 11 at $p < .05$), after including quadratic time trends, unit roots are not rejected (at the 10% level) for aluminum and petroleum. However, the results for aluminum should be interpreted with care, as the coefficient on the quadratic time trend is not significantly different from zero at conventional levels, suggesting the linear time trend specification is more appropriate. Although we find that the quadratic time trend is significant at the 10% level in five of the eleven price series, other results are qualitatively unchanged. Overall, at the 5% level, the results in Table 4 suggest that we should reject the unit root in nine of the eleven series, providing additional evidence that commodity prices are stationary around a breaking deterministic trend.

To further examine the effect of allowing for two breaks instead of one, we additionally test all price series using an endogenous one-break LM test with quadratic

[11] Since the two-break LM unit root test with quadratic trend is a new test, we derive critical values by generating pseudo-iid $N(0,1)$ random numbers using the Gauss (version 6.0.5) RNDNS procedure. Critical values are derived using 5,000 replications involving a grid search of all combinations of the endogenous break points for minimum statistics in samples of $T = 100$. The asymptotic distribution of the two-break minimum LM test with quadratic trend can be shown by a simple modification of the expression for that of the two-break minimum LM test with linear trend. The expression of the demeaned and de-breaked Brownian bridge $\underline{V}(r, \lambda)$ in Lee and Strazicich (2003, equations 6c, Appendix A.6) must be changed to incorporate a detrended p-level Brownian bridge as given in SP (1992, equation 40). This expression is described as a residual process projected onto the subspace generated by $dz(\lambda,r) = [1, \underline{\mathbf{r}}, d_1(\lambda_1,r), d_2(\lambda_2,r)]$. The trend function is given as $\underline{\mathbf{r}} = \{r^j, j=0,1,..,p\text{-}1\}$, where $d_j(\lambda_j,r) = 1$ if $r > \lambda_j$, for $j=1,2$, and 0 otherwise.

trend.[12] The results are displayed in Table 5. Compared to the two-break with quadratic trend results in Table 4, we find the same number of rejections of the unit root null (nine of eleven series reject the null at $p < .05$ and $p < .10$). The only notable difference in results is with regards to the two series that do not reject the unit root. In Table 4 aluminum and petroleum did not reject the unit root, but in the one-break test results of Table 5 petroleum rejects the unit root (at $p < .05$) while silver does not. The difference in results for silver is likely due to bias from omitting the second break variable. The difference for petroleum may be due to including an unnecessary break in the two-break quadratic trend model, which would be expected to reduce power.

Again, the majority of identified break years correspond primarily to the two World Wars, the Great Depression, and the 1970s energy crisis: fifteen of the twenty-two breaks (68%) correspond to these time periods, with the most common time period for breaks in the early 1970s. In addition, other important events, such as the Lead and Zinc Stabilization Program in 1961, seem to also coincide with an identified break.

The most appropriate model in the literature to compare our results is, again, Ahrens and Sharma's (1997) OPP test, which allows for a quadratic trend but no structural break(s). Using the OPP test, Ahrens and Sharma (1997) reject the unit root null in five of eight series at $p < .05$. Using the two-break LM test with quadratic trend, we reject the unit root in nine of eleven series at the 5% level. This difference is likely due to allowing for structural breaks.

To examine further the effects of allowing for breaks, we present plots showing each commodity price series in Figure 1. Each panel shows the actual commodity price

---

[12] We thank an anonymous referee who suggested this robustness check.

series as well as fitted values from a linear time trend regression connecting the two breaks in intercept and trend slope identified in Table 2. The plots suggest that most price series are trend-stationary around a small number of breaks. More importantly, a visual inspection indicates that it may be more appropriate to specify the model with discrete shifts in a deterministic linear trend rather than a quadratic trend. Even in some series for which a quadratic trend appears to fit, the linear trend model with two structural breaks also appears to fit the data well. Overall, seven of the eleven series have downward trends in the most recent period.

In sum, the above results provide the strongest evidence to date against the unit root hypothesis and suggest that natural resource price series are stationary around deterministic trends with occasional changes in intercept and trend slope. Our findings are at odds with those in previous studies. In addition, our findings strengthen those of Ahrens and Sharma (1997), who concluded that the majority (six of eleven) of natural resource price series reject the unit root hypothesis. In certain respects, our unit root test results are consistent with the scores of theoretical models that suggest resource price paths are deterministic, but downward or U-shaped. In the next section, we examine how structural change affects forecasting.

V. Forecasting Issues

As aforementioned, we view one contribution of the above analysis as informing the forecaster. Whether, and to what extent, structural change in natural resource prices influences the selection and performance of forecasting models is largely unknown, however. In this section we fill this void by exploring whether pre-testing for a unit root with structural change and inclusion of breaks in the forecast model can lead to more accurate forecasts. Diebold and Kilian (2000) previously examined whether pre-testing for unit roots matters for forecasting, given that a priori information on the properties of a time series is not available. The issue that we examine is similar in spirit, yet we consider structural breaks. One may argue that allowing for breaks when pre-testing for unit roots can give potentially different results, since a unit root process with breaks can sometimes be viewed as observationally equivalent to, or hardly distinguishable from, a trend stationary process with breaks. Intuitively speaking, this can be the case because structural change implies persistent effects similar to those in a unit root process (see Phillips, 1988). Since the pioneering work of Perron (1989), however, it is well known that inference in unit root tests will be adversely affected by ignoring structural change. Under the alternative hypothesis of a trend-break stationary series, structural change would be described by occasional predetermined large shocks that permanently alter the level and slope of the deterministic trend. This is the common view in the literature as consistently advocated in numerous theoretical and empirical papers. Following this view, the existence of structural change is not considered to be sufficient support for the unit root hypothesis.

In this section, we seek to determine how structural change will affect the selection and accuracy of forecasting models when we do not know *a priori* if a time series is (trend) stationary (with breaks) or nonstationary (with breaks). Several authors have noted the deleterious effects of ignoring structural breaks in forecasting models (see, e.g., Pesaran and Timmermann, 2002, 2004, and Clark and McCracken, 2003). However, these previous papers deal with mostly stationary data and do not address the question of whether forecasts in levels perform better than forecasts in differences in the presence of breaks.

We begin our investigation by conducting Monte Carlo simulations to study the role of structural breaks in selecting and estimating the most accurate forecasting model. To do so, we compare the unconditional prediction mean squared error (PMSE) of the following five strategies:

(M1) Difference the data and include two breaks in intercept and trend slope.

(M2) Analyze the data in levels and include two breaks in intercept and trend slope.

(M3) Pre-test with unit root tests that allow for two breaks in intercept and trend slope and proceed as follows: if the unit root null is rejected, then select M2; if we cannot reject the unit root, then select M1.

(M4) Difference the data and do not allow for breaks.

(M5) Analyze the data in levels and do not allow for breaks.

Note that models M1-M3 are conditional on the existing breaks, while M4 and M5 are not. Thus, we can compare the performance of pre-testing for a unit root in addition to the effect of including known breaks in the forecasting model. In general, when comparing the accuracy of forecasts with different models the common approach is to compare out-of-sample forecasts. We follow this approach.[13]

[13] There does not appear to be a unanimous consensus on this issue, however. Inoue and Kilian (2004) suggest that comparing in-sample tests of forecast accuracy are more reliable than out-of-sample tests. The underlying issue is potential data mining or un-modeled structural change. Also, the accuracy of out-of-sample forecasts can depend on the selection of the holdout period.

For simplicity, we restrict our simulations to AR(1) or ARI(1,1) models as in Diebold and Kilian (2000). Thus, M1 is an ARI(1, 1) model with two breaks and can be described as follows:

$$\Delta y_t = \mu + b\Delta y_{t-1} + d_1 B_{1t} + d_2 B_{2t} + d_3 D_{1t} + d_4 D_{2t} + v_t , \quad (6)$$

where $B_{jt} = 1$ for $t = T_{Bj} + 1, j = 1, 2$, and zero otherwise. M2 is an AR(1) model with two breaks as follows:

$$y_t = c + by_{t-1} + d_1 D_{1t} + d_2 D_{2t} + \gamma t + d_3 DT_{1t} + d_4 DT_{2t} + w_t . \quad (7)$$

M3 would be either the model in (6) or (7) depending on the results from pre-testing for a unit root (with two breaks). M4 and M5 can be described by (6) and (7), respectively, without the dummy variables describing the breaks. We can rewrite the DGP in (1) to be as follows:

$$(y_t - c - \gamma t - d_1 D_{1t} - d_2 D_{2t} - d_3 DT_{1t} - d_4 DT_{2t}) = \beta(y_{t-1} - c - \gamma(t\text{-}1) - d_1 D_{1,t-1} -$$

$$d_2 D_{2,t-1} - d_3 DT_{1,t-1} - d_4 DT_{2,t-1}) + u_t , \quad (8)$$

where we let $c = 0$, $\gamma = 0.1$, and $\sigma_u^2 = 1$. Then, we vary the values of the persistent parameter β as follows. To generate nonstationary data we let $\beta = 1.0$; to generate stationary data we let $\beta = 0.95$ and $\beta = 0.5$. Another interesting question is how big should structural breaks be to be recognized as such? We address this question by varying the magnitude of the breaks as follows: $d = (d_1, d_2, d_3, d_4) = (0, 0, 0, 0)$, (2, 2, 0.2, 0.2), (6, 6, 0.6, 0.6), and (10, 10, 1.0, 1.0), where d_1 and d_2 denote intercept shifts, and d_3 and d_4 denote trend slope changes, respectively. To abbreviate, we denote these different magnitudes as $d = 0$, $d = 2$, $d = 6$, and $d = 10$, respectively, where $d = 0$ indicates the model without breaks. All break sizes are described using standardized units. We consider different break locations at $T_{B1}/T = 0.3$ and $T_{B2}/T = 0.6$.

To undertake our comparisons we examine the recursive one-step ahead predictions of y_{T+h} at forecast horizons h ranging from 1 to 100 periods, and employ sample sizes T = 100 and T = 500. Pre-testing is performed using the two-break LM unit root test and critical values at the 5% level of significance. To compare the forecast accuracy of each model, we calculate the ratio of the prediction mean squared error (PMSE) for each h in each model relative to the PMSE of the reference pre-test strategy M3. The calculated ratios are denoted as M1/M3, M2/M3, M4/M3, and M5/M3. For the models estimated in first-differences, the forecasted values are converted to levels prior to calculating their mean squared error. The forecast error is calculated as the difference between the (simulated) data and the predicted values (in levels), or $y_{T+h} - \hat{y}_{T+h}$, h = 1,…,100. Thus, PMSE(h) = $h^{-1}\sum_{j=1}^{h}(y_{T+j} - \hat{y}_{T+j})^2$ using forecast errors through $t = h$. When the selected forecasting model is more accurate than the model suggested by pre-testing, then the ratio will be less than one and vice versa. For parsimony, we report only the summary results for selected values of *h.*

We first examine the simulation results using nonstationary data (β = 1.0) in Table 6. When the DGP is nonstationary, the differencing strategy of M1 (ARI with breaks) has the lowest PMSE in nearly all cases. There are only very small differences in some cases. In nearly all cases the best model (M1) is also the model that is selected by the pre-testing. Any differences are approximately zero, and the results are unaffected by the size of the breaks or the sample size. Overall, these results clearly support the use of pre-testing for a unit root and including structural breaks. For the case of no breaks there is a small gain from omitting the breaks in the (differenced) forecasting model (M4). For the case of small breaks (d = 2), the results demonstrate little difference between including breaks (M1) and omitting breaks (M4). Thus, as one might expect, small breaks have only minimal effects on the PMSE and can be hardly considered as structural breaks. However, as the size of the breaks increases the benefit of including breaks increases especially at longer forecast horizons.

We next examine the simulation results using stationary data in Table 6. When the DGP is clearly stationary (β = 0.5), the model in levels with breaks (M2) is superior in virtually all cases. Any differences are so small that they are approximately zero and can be ignored. In addition, M2 is the model that is selected with pre-testing in nearly every case. Again, any differences are quite small, and the results are unaffected by the break size, forecast horizon, and sample size. When β = 0.95 and the unit root null is more difficult to reject, the most accurate forecasting model is in differences (M1) instead of levels (M2). However, the difference model is also the model most frequently selected by pre-testing. Thus, pre-testing for a unit root (with breaks) is again the preferred method to select the most accurate forecasting model. As in the nonstationary case, including breaks when they exist is always beneficial. Figures 2 and 3 summarize the simulation results by displaying spectrum with varying sample sizes of T = 20n, n = 1,..,50, over the continuum of forecasting horizons h = 1,..,100. These figures generalize the results in Table 6 and are consistent with the above discussion.

To further our investigation, we next examine and compare the accuracy of the different forecasting strategies by using actual data on natural resource prices. To perform our investigation, we utilize more recent data on eight natural resource real prices (aluminum, copper, iron, lead, nickel, silver, tin, and zinc). These data come from the U.S. Geological Survey (Kelly et al., 2005), and are available annually from 1900-2002 in real (1998) dollars per ton.[14]

Our testing procedure is as follows. We first perform pre-testing for each real price series as before using the two-break LM unit root test with linear trend. In nearly all cases (except for nickel), the unit root null is rejected at the 10% level of significance. We also perform pre-testing using the two-break LM unit root test with quadratic trend and reject the unit root in all cases at the 10% level of significance.[15] Next, we estimate four forecasting models for each series with a linear trend as in (6) and (7) (M1, M2, M4, and M5), and then with a quadratic trend (M1Q, M2Q, M4Q, and M5Q).

Instead of the AR(1) and ARI(1, 1) models (with and without breaks) considered in the simulations of Table 6, we consider more general ARMA(p,q) models and ARIMA(p,1,q) models that might be more realistically considered in applications. To determine the order p and q of each ARMA and ARIMA model, and the location of two breaks, we utilize the SBC statistic. We consider all combinations of p and q over the range 0 to 5 at each combination of two breaks to find the model with the lowest SBC. As such, our estimation procedure allows for the location of breaks to be jointly determined with the best fitting model. In each case, we estimate the forecasting model using data from 1900-1990 and generate out-of-sample forecasts in levels for 1991-

[14] We would have preferred to update our data, but the new data uses a different PPI to deflate the data and different units of measurement. And, the new data is only available for eight of our previous eleven natural resource price series. However, besides using more recent data (past 1990), the new data has the advantage that each series is available for the same time period (1990-2002).
[15] These results are available upon request.

2002.[16] Then, the out-of-sample forecasts are compared to the actual data to calculate the relevant PMSE. To test the statistical significance of the break variables in the models that include breaks, F-tests were performed to test for the significance of the breaks. In each case, the null hypothesis that break coefficients are jointly equal to zero was rejected at the 1% level of significance. These results provide additional evidence that structural breaks are important.

The PMSE are displayed in Table 7 for each model type at forecast horizons h = 1, 3, 5, 8, and 12. The number in bold denotes the model with the most accurate forecasts (smallest PMSE), which is shown separately for the models with linear and quadratic trends. At each forecast horizon (h), the best overall model is denoted with an asterisk (*). Given that pre-testing implied that price series are trend-break stationary, we expect that M2 (ARMA model in levels with breaks) will have the lowest PMSE in nearly all cases.

We now examine the results in Table 7. Except for tin, and to a lesser extent iron, we see little evidence that including a quadratic trend improves the accuracy of the forecast, as the best model has a linear trend in all other cases. For aluminum, copper, and lead our pre-test expectations are realized as the most accurate model in every case but one (lead, h =1) is the ARMA model in levels with breaks (M2). While iron selects the quadratic trend model at three of the five forecast horizons, three of its five best models are in levels with breaks (M2, M2Q) as suggested by the pre-test. For the other four price series the results are mixed. In all cases but one (h = 1), the best forecasting

[16] As noted in footnote 14, the out-of-sample PMSE results might vary depending on how we define the out-of-sample period. We also considered 1996 as a starting period for the out-of-sample forecasts, but the main conclusions did not vary.

model for nickel is the ARMA model in levels without breaks (M5). For silver, the results are unexpected as four of the five best models are in differences (M1, M4), and three of these are without breaks (M4). Tin is the series where the quadratic trend model is most often selected (four of five best models). Although including a quadratic trend provides better forecasting accuracy for tin, the best model chosen in all cases is in differences and without breaks (M4, M4Q). Finally, zinc selects an ARMA model in levels in all cases as predicted by pre-testing, but the best model is without breaks (M5).

Overall, the results in Table 7 using actual natural resource data present a less clear picture than our simulation results in Table 6. However, as expected, the most consistently selected model is the same as selected in the pre-test; in four of the eight series the ARMA model with breaks (M2 or M2Q) is the most accurate in nearly all cases. For the other four series the results are less clear. In some of these cases, the best model is in levels as expected, but without breaks. In other cases the superior model is in differences. Thus, while the model selected by pre-testing is not always the best model it appears that pre-testing for a unit root and considering breaks are important. The gain from considering breaks is most obvious when comparing the PMSE of those superior models that include breaks to the other models.

In general, as is well known, forecasting future values out-of-sample is a difficult task. Therefore, it may not be desirable to draw any generalized conclusions regarding forecasting strategies from a particular data application to a particular out-of-sample period. Choosing different data and different out-of-sample periods may lead to different conclusions. Thus, we give stronger credence to our simulation results as a guide to more accurate forecasting. We conclude that it is better to follow a forecasting strategy of pre-testing for a unit root with breaks and including existing breaks in the forecast model.

VI. Concluding Remarks

Without an appropriate understanding of the dynamics of a time series, empirical verification of theories, forecasting, and proper inference are potentially misleading. In this paper we examine temporal properties of eleven nonrenewable natural resource real price series from 1870-1990. Previous efforts largely suggest that these series are difference-stationary with stochastic trends. We re-examine these data and advance the literature by considering unit root tests with two structural breaks that are endogenously determined by the data. In addition, we consider models with linear and quadratic trends.

Overall, we find evidence to indicate that natural resource prices are stationary around deterministic trends with structural breaks in intercept and trend slope. In particular, after controlling for breaks, the previous empirical findings of Agbeyegbe (1993), Berck and Roberts (1996), and Slade (1988) are essentially reversed. Our findings strengthen those of Ahrens and Sharma (1997), who allowed for one known break and concluded that the majority of natural resource price series (six of eleven) reject the unit root. Following our unit root tests, we examined forecasting models with breaks by employing both simulations and actual natural resource price data. From our simulations, we conclude that it is better to pre-test for a unit root (with breaks) prior to selecting the best forecast model, and to include breaks, if they exist, in all cases. While our forecasting results with actual data are less clear than our simulations, the most consistently selected model is the same as selected in the pre-test.

Given that accurate forecasting and empirical verification of theories can depend critically on understanding the appropriate nature of time-series, our results should have important implications for academics and policymakers alike. With the proliferation of environmental and natural resource data having a temporal dimension, researchers interested in dynamic issues as broad as environmental quality expenditures, environmental degradation, species extinction rates, and land-use patterns should have a firm grounding in time-series analysis. We hope that our paper will further the development of this literature.

References

Abgeyegbe, T. D., “The Stochastic Behavior of Mineral-Commodity Prices,” in “Models, Methods, and Applications in Econometrics: Essays in Honor of A. R. Bergstrom” (P. C. B. Phillips, Ed.), Blackwell Science, Oxford (1993).

Ahrens, W. A. and V. R. Sharma, “Trends in Natural Resource Commodity Prices: Deterministic or Stochastic?” *Journal of Environmental Economics and Management* 33, 59-74 (1997).

Amsler, C. and J. Lee, “An LM Test for a Unit-Root in the Presence of a Structural Change,” *Econometric Theory* 11, 359-368 (1995).

Barnett, H. J. and C. Morse, “Scarcity and Growth: The Economics of Natural Resource Availability,” John Hopkins Press, Baltimore (1963).

Berck, P. and M. Roberts, “Natural Resource Prices: Will They Ever Turn Up?” *Journal of Environmental Economics and Management* 31, 65-78 (1996).

Clark, T., and M. W. McCracken, “The Power of Tests of Predictive Ability in the Presence of Structural Breaks,” working paper, Federal Reserve Bank of Kansas City (2003).

Diebold, F. X., and L. Kilian, “Unit-Root Tests Are Useful for Selecting Forecasting Models,” *Journal of Business & Economic Statistics*, 18 (3), 265-273 (2000).

Diebold F.X. and A.S. Senhadji, “The Uncertain Unit Root in real GNP: Comment,” *American Economic Review* 86, 1291-1299 (1996).

Hamilton, J. D., “This Is What Happened to the Oil Price-Macroeconomy,” *Journal of Monetary Economics* 38(2), 215-20 (1996).

Heal, G. M., “The relationship between Price and Extraction Cost for Resource with a Backstop Technology,” *Bell Journal of Economics* 7, 317-378 (1976).

Hotelling, H., “The Economics of Exhaustible Resources,” *Journal of Political Economy* 39, 137-175 (1931).

Inoue, A. and L. Kilian, “In-sample or Out-of-sample Tests of Predictability: Which one Should We Use?” *Econometric Review, 23(4), 371-403* (2004).

Kelly, T., Buckingham, D., DiFrancesco, C., Porter, K., Goonan, T., Sznopek, J., Berry, C., and M. Crane, Historical Statistics for Mineral and Material Commodities in the United States, United States Geological Survey Open-File Report 01-006, Version 9.2, Washington, DC (2005).

Labson, B. S., and P. L. Crompton, "Common Trends in Economics Activity and Metals Demand: Cointegration and the Intensity of Use Debate," *Journal of Environmental Economics and Management* 25, 147-161 (1993).

Lee, J. and M.C. Strazicich, "Break Point Estimation and Spurious Rejections with Endogenous Unit Root Tests," *Oxford Bulletin of Economics and Statistics* 63, 535-558 (2001).

Lee, J., and M.C. Strazicich, "Minimum LM Unit Root Test with Two Structural Breaks," *Review of Economics and Statistics* 85(4), 1082-1089 (2003).

Lee, J., and M.C. Strazicich. "Minimum LM Unit Root Test with One Structural Break." Working Paper, Department of Economics, Appalachian State University, (2004).

List, J., "Have Air Pollutant Emissions Converged Amongst US Regions? Evidence from Unit-Root Tests," *Southern Economic Journal* 66 (1), 144-155 (1999).

Lumsdaine, R. and D. Papell, "Multiple Trend Breaks and the Unit-Root Hypothesis," *The Review of Economics and Statistics*, 79(2), 212-218 (1997).

Manthy, R. S., "Natural Resource Commodities: A Century of Statistics," Johns Hopkins Press, Baltimore (1978).

Nelson, C. R. and C. I. Plosser, "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics* 10, 139-162 (1982).

Ng, S. and P. Perron, "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag," *Journal of the American Statistical Association* 90, 269-281 (1995).

Nunes, L., P. Newbold, and C. Kuan, "Testing for Unit Roots with Breaks: Evidence on the Great Crash and the Unit Root Hypothesis Reconsidered," *Oxford Bulletin of Economics and Statistics* 59, 435-448 (1997).

Ouliaris, S., J. Y. Park, and P. C. Phillips, "Testing for a Unit Root in the Presence of a Maintained Trend," in *Advances in Econometrics* (B. Raj, Ed.), Kluwer Academic, Dordrecht/Norwell, MA (1989).

Perron, P., "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica* 57, 1361-1401 (1989).

Pesaran, H. M., and A. Timmermann, "Small Sample Properties of forecasts from Autoregressive Models Under Structural Breaks," working paper, University of California, San Diego (2004).

Pesaran, H. M., and A. Timmermann, "How Costly Is It to Ignore Breaks When Forecasting the Direction of a Time Series?" working paper, University of California, San Diego (2002).

Pindyck, R.S., "Uncertainty and Exhaustible Resource Markets," *Journal of Political Economy* 88, 1203-1225 (1980).

Schmidt, P. and P. C. B. Phillips, "LM Tests for a Unit Root in the Presence of Deterministic Trends," *Oxford Bulletin of Economics and Statistics* 54, 257-287 (1992).

Schurr, S. H., "Historical Statistics of Minerals in the United States," Resources for the Future, Inc. (1960).

Slade, M. E., "Trends in Natural-Resource Commodity Prices: An Analysis of the Time Domain," *Journal of Environmental Economics and Management* 9, 122-137 (1982).

–––"Grade Selection Under Uncertainty: Least Cost Last and Other Anomalies," *Journal of Environmental Economics and Management* 15, 189-205 (1988).

Vougas D, "Reconsidering LM Unit Root Testing," *Journal of Applied Statistics*, 30 (7), 727-741 (2003).

Zivot, E. and D. W. K. Andrews, "Further Evidence on the Great Crash, the Oil-Price Shock and the Unit Root Hypothesis," *Journal of Business and Economic Statistics* 10, 251-270 (1992).

Table 1. LM Unit Root Test with a Quadratic Trend

Series	*T*	*Starting Year*	$\hat{k}$	*Test Stat.*	Estimation Results [a]
Aluminum	90	1895	7	0.88	$\Delta y_t = .02\tilde{S}_{t-1} + .76 - .09t$ + lags (0.88) (0.25) (-0.80)
Coal	121	1870	7	-3.75**	$\Delta y_t = -.02\tilde{S}_{t-1} - .30 + 0.004t$ + lags (-3.75) (-2.06) (1.91) +
Copper	121	1870	3	-3.06	$\Delta y_t = -.21\tilde{S}_{t-1} + .62 - .02t$ + lags (-3.06) (0.49) (-1.06)
Iron	104	1870	1	-4.83**	$\Delta y_t = -.40\tilde{S}_{t-1} - .73 - 0.17t$ + lags (-4.83) (-0.29) (-2.97) ++
Lead	121	1870	0	-4.69**	$\Delta y_t = -.32\tilde{S}_{t-1} - .12 + .01t$ + lags (-4.69) (-0.34) (1.21)
Gas	72	1919	5	-3.87**	$\Delta y_t = -.13\tilde{S}_{t-1} + .77 - .04t$ + lags (-3.87) (1.14) (-1.95) +
Nickel	78	1913	2	-2.00	$\Delta y_t = -.16\tilde{S}_{t-1} - 1.07 - .23t$ + lags (-2.00) (-0.24) (-1.27)
Petroleum	120	1870	2	-1.77	$\Delta y_t = -.08\tilde{S}_{t-1} -.11 - .01t$ + lags (-1.77) (-0.59) (-1.13)
Silver	121	1870	2	-2.94	$\Delta y_t = -.24\tilde{S}_{t-1} - 6.59 - 0.05t$ + lags (-2.94) (-0.42) (-.22)
Tin	106	1885	0	-2.20	$\Delta y_t = -.08\tilde{S}_{t-1} - 4.70 + .23t$ + lags (-2.20) (-0.86) (1.58)
Zinc	121	1870	2	-3.74**	$\Delta y_t = -.41\tilde{S}_{t-1} + 2.85 - .09t$ + lags (-3.74) (2.61) (-3.40) ++

Notes: Critical values of the LM test with quadratic trend are –4.16, –3.60, and –3.31 at the 1%, 5%, and 10% levels, respectively, for samples of size T = 100. Critical values are –4.28, –3.65, and –3.34 for T = 50, and –4.12, –3.55, and –3.28 for T = 200, respectively. t-statistics are given in parentheses. a: Result is based on equation (3). Estimated coefficients of the augmented terms (lags) are omitted to conserve space. Note that under the stationary alternative hypothesis the coefficient of *t* corresponds to a quadratic trend, and the constant term corresponds to a linear trend, in Z_t, respectively * and ** denote significant at the 10% and 5% levels, respectively. + and ++ denote significance of the quadratic trend term at the 10% and 5% levels, respectively.

Table 2. Two Break Minimum LM Unit Root Test

Series	T	*Starting Year*	$\hat{k}$	$\hat{T}_B$	*Test Stat.*	Estimation Results [a]
Aluminum	90	1895	8	1913, 1944	-5.39*	$\Delta y_t = -1.04\tilde{S}_{t-1} - 40.05 - 32.55B_{1t} + 6.17B_{2t} + 51.63D_{1t} - 12.98D_{2t}$ +lags (-5.39) (-5.07) (-2.73) (0.66) (5.17) (-4.01)
Coal	121	1870	7	1902, 1972	-7.75**	$\Delta y_t = -.61\tilde{S}_{t-1} - .24 + .69B_{1t} - 1.95B_{2t} - .16D_{1t} + 1.81D_{2t}$ + lags (-7.75) (-2.44) (1.48) (-3.63) (-1.46) (6.60)
Copper	121	1870	1	1918, 1969	-6.01**	$\Delta y_t = -.53\tilde{S}_{t-1} + 3.93 - 5.41B_{1t} + 12.91B_{2t} - 7.14D_{1t} - 2.65D_{2t}$ + lags (-6.01) (3.60) (-0.91) (2.16) (-4.10) (-1.71)
Iron	104	1870	3	1932, 1951	-6.72**	$\Delta y_t = -.72\tilde{S}_{t-1} - 4.49 + 4.92B_{1t} - 6.66B_{2t} + 2.95D_{1t} + 11.43D_{2t}$ + lags (-6.72) (-3.20) (0.47) (-0.63) (1.07) (3.02)
Lead	121	1870	5	1945, 1961	-6.91**	$\Delta y_t = -.90\tilde{S}_{t-1} + .46 - 2.19B_{1t} + 10.46B_{2t} + 3.08D_{1t} - 5.73D_{2t}$ + lags (-6.91) (2.07) (-1.16) (5.46) (4.55) (-5.81)
Gas	72	1919	7	1943, 1975	-6.46**	$\Delta y_t = -.41\tilde{S}_{t-1} + 1.86 + .92B_{1t} - 4.51B_{2t} - 2.59D_{1t} + 6.30D_{2t}$ + lags (-6.46) (3.08) (0.43) (-1.93) (-3.15) (5.06)
Nickel	78	1913	1	1928, 1942	-5.44*	$\Delta y_t = -.58\tilde{S}_{t-1} - 13.83 - 15.05B_{1t} + 1.71B_{2t} + 23.94D_{1t} - 9.38D_{2t}$ + lags (-5.44) (-2.89) (-0.88) (0.10) (3.22) (-1.71)
Petroleum	120	1870	6	1896, 1971	-5.65*	$\Delta y_t = -.48\tilde{S}_{t-1} - 1.59 - 1.82B_{1t} - 1.59B_{2t} + 1.58D_{1t} + 1.43D_{2t}$ + lags (-5.65) (-5.02) (-2.18) (-1.85) (4.82) (4.51)
Silver	121	1870	0	1899, 1972	-7.22**	$\Delta y_t = -.63\tilde{S}_{t-1} + 32.43 + 21.90B_{1t} - 38.82B_{2t} - 64.03D_{1t} + 117.82D_{2t}$ +lags (-7.23) (2.11) (0.28) (-0.48) (-3.29) (4.38)
Tin	106	1885	6	1963, 1974	-6.66**	$\Delta y_t = -.67\tilde{S}_{t-1} + 3.61 + 51.76B_{1t} - 115.19B_{2t} - 16.23D_{1t} + 31.46D_{2t}$ + lags (-6.66) (1.68) (2.72) (-5.45) (-2.26) (3.00)
Zinc	121	1870	0	1917, 1952	-10.4**	$\Delta y_t = -.97\tilde{S}_{t-1} + 1.46 - 1.24B_{1t} - 3.89B_{2t} - 2.92D_{1t} - 0.65D_{2t}$ + lags (-10.40) (2.69) (-0.34) (-1.07) (-3.38) (-0.76)

Notes: Critical values of the two-break LM test with structural breaks in level and trend vary depending on the location of breaks (λ). Critical values are shown at the 1%, 5%, and 10% levels, respectively, for samples of size T = 100. Critical values are –6.16, –5.59, and –5.28 for λ = (.2,.4); –6.40, –5.74, and –5.32 for λ = (.2,.6); –6.33, –5.71, and –5.33 for λ = (.2,.8); –6.46, –5.67, and –5.31 for λ = (.4,.6); –6.42, –5.65, and –5.32 for λ = (.4,.8); –6.32, –5.73, and –5.32 for λ = (.6,.8), where λ indicates the relative location of the breaks. Other critical values can be interpolated. The critical values are symmetric around λ_1 and λ_2. B_{1t}, B_{2t} and D_{1t}, D_{2t} correspond to changes in intercept and trend slope under the alternative, respectively. B_{1t} and B_{2t} denote one-period jumps in level and D_{1t} and D_{2t} correspond to changes in drift under the null hypothesis, respectively. Note that the constant term corresponds to a linear trend in Z_t under the alternative. $\hat{k}$ denotes the estimated optimal number of first-differenced lagged terms included to correct for serial correlation. *t*-statistics are shown in parentheses. a: See notes in Table 1. * and ** denote significant at the 10% and 5% levels, respectively. Critical values come from Lee and Strazicich (2003).

Table 3. One Break Minimum LM Unit Root Test

Series	T	*Starting Year*	$\hat{k}$	$\hat{T}_B$	*Test Stat.*	Estimation Results [a]
Aluminum	90	1895	3	1914	-3.45	$\Delta y_t = -.26\tilde{S}_{t-1} - 13.87 + 47.79B_{1t} + 6.94D_{1t}$ + lags (-3.45) (-3.72) (4.73) (2.34)
Coal	121	1870	7	1915	-4.08	$\Delta y_t = -.20\tilde{S}_{t-1} - .03 - .27B_{1t} + .33D_{1t}$ + lags (-4.08) (-0.41) (-0.52) (2.69)
Copper	121	1870	0	1918	-5.18**	$\Delta y_t = -.37\tilde{S}_{t-1} + 2.73 - 10.45B_{1t} - 2.48D_{1t}$ + lags (-5.18) (2.69) (-1.76) (-2.03)
Iron	104	1870	1	1927	-5.49**	$\Delta y_t = -.47\tilde{S}_{t-1} + .98 - 15.93B_{1t} + 9.81D_{1t}$ + lags (-5.49) (0.61) (-1.34) (3.60)
Lead	121	1870	5	1946	-5.22**	$\Delta y_t = -.67\tilde{S}_{t-1} - .03 + 7.34B_{1t} - 1.28D_{1t}$ + lags (-5.22) (-0.16) (3.97) (-3.11)
Gas	72	1919	5	1973	-5.24**	$\Delta y_t = -.22\tilde{S}_{t-1} + .03 - 2.04B_{1t} + 5.38D_{1t}$ + lags (-5.24) (0.09) (-0.91) (5.39)
Nickel	78	1913	1	1969	-5.08**	$\Delta y_t = -.48\tilde{S}_{t-1} - 15.33 + 21.66B_{1t} + 9.24D_{1t}$ + lags (-5.08) (-4.19) (1.32) (2.02)
Petroleum	120	1870	6	1896	-3.82	$\Delta y_t = -.22\tilde{S}_{t-1} - .79 - 1.36B_{1t} + .69D_{1t}$ + lags (-3.82) (-3.09) (-1.60) (2.78)
Silver	121	1870	2	1960	-3.84	$\Delta y_t = -.37\tilde{S}_{t-1} - 6.07 - 45.44B_{1t} + 53.93D_{1t}$ + lags (-3.84) (-0.71) (-0.56) (2.54)
Tin	106	1885	6	1969	-4.97**	$\Delta y_t = -.44\tilde{S}_{t-1} - 2.49 - 29.73B_{1t} + 27.08D_{1t}$ + lags (-4.97) (-1.05) (-1.36) (3.31)
Zinc	121	1870	1	1955	-5.96**	$\Delta y_t = -.71\tilde{S}_{t-1} + .32 + 3.32B_{1t} - 2.53D_{1t}$ + lags (-5.96) (0.78) (0.88) (-2.86)

Notes: Critical values of the one-break LM test with a structural break in level and trend vary depending on the location of the break (λ). Critical values are shown at the 1%, 5%, and 10% levels, respectively, for samples of size T = 100. Critical values are –5.11, –4.50, and –4.21 for λ = .1; –5.07, –4.47, and –4.20 for λ = .2; –5.15, –4.45, and –4.18 for λ = .3; –5.05, –4.50, and –4.18 for λ = .4; –5.11, –4.51, and –4.17 for λ = .5, where λ indicates the relative location of the breaks. Other critical values can be interpolated. The critical values are symmetric around λ. B_{1t} and D_{1t} correspond to changes in intercept and trend slope under the alternative, respectively. B_{1t} denotes a one-period jump in level, and D_{1t} corresponds to a change in drift under the null hypothesis, respectively. Note that the constant term corresponds to a linear trend in Z_t under the alternative. $\hat{k}$ denotes the estimated optimal number of first-differenced lagged terms included to correct for serial correlation. t-statistics are shown in parentheses. a: See notes in Table 1. * and ** denote significant at the 10% and 5% levels, respectively. Critical values come from Lee and Strazicich (2004).

Table 4. Two Break Minimum LM Unit Root Test with Quadratic Trend

Series	*T*	*Starting Year*	$\hat{k}$	$\hat{T}_B$	*Test Stat.*	*Estimation Results* [a]
Aluminum	90	1895	8	1913, 1942	-5.15	$\Delta y_t = -.99\tilde{S}_{t-1} - 38.51 + .06\,t - 31.56B_{1t} + 6.96B_{2t} + 48.35D_{1t} - 14.44D_{2t}$ + lags (-5.15) (-4.63) (1.13) (-2.62) (0.70) (4.95) (-2.70)
Coal	121	1870	7	1949, 1972	-8.51**	$\Delta y_t = -.69\tilde{S}_{t-1} + .01 - .002t - .06B_{1t} - 2.26B_{2t} - .24D_{1t} + 2.35D_{2t}$ + lags (-8.51) (0.09) (-1.65)[+] (-0.14) (-4.25) (-1.49) (7.20)
Copper	121	1870	2	1897, 1970	-6.70**	$\Delta y_t = -.65\tilde{S}_{t-1} + 5.07 - 0.08t - 5.22B_{1t} - 6.68B_{2t} + 5.12D_{1t} + 3.39D_{2t}$ + lags (-6.70) (3.28) (-3.77)[++] (-0.93) (-1.21) (2.41) (1.58)
Iron	104	1870	3	1933, 1951	-6.66**	$\Delta y_t = -.75\tilde{S}_{t-1} - 1.26 - .06\,t - 1.17B_{1t} - 8.53B_{2t} + 7.61D_{1t} + 14.34D_{2t}$ + lags (-6.65) (-0.43) (-1.36) (-0.11) (-0.79) (1.79) (3.18)
Lead	121	1870	5	1945, 1961	-7.30**	$\Delta y_t = -.97\tilde{S}_{t-1} + .81 - .01t - 2.64B_{1t} + 10.57B_{2t} + 4.58D_{1t} - 5.81D_{2t}$ + lags (-7.30) (1.71) (-1.80)[+] (-1.41) (5.59) (5.01) (-5.92)
Gas	72	1919	7	1929, 1979	-9.54**	$\Delta y_t = -.57\tilde{S}_{t-1} + .79 - .05t + 4.19B_{1t} - 1.06B_{2t} + 2.20D_{1t} + 4.11D_{2t}$ + lags (-9.54) (0.76) (-3.92)[++] (2.43) (-0.54) (1.94) (2.50)
Nickel	78	1913	1	1930, 1974	-6.22**	$\Delta y_t = -.71\tilde{S}_{t-1} - 13.62 + .04t + 6.66B_{1t} - 17.51B_{2t} + 16.54D_{1t} - 14.04D_{2t}$ + lags (-6.22) (-3.05) (0.44) (0.41) (-1.10) (2.10) (-1.96)
Petroleum	120	1870	1	1914, 1926	-5.74	$\Delta y_t = -.34\tilde{S}_{t-1} - 1.25 + .01t - .72B_{1t} - .92B_{2t} + .42D_{1t} - .92D_{2t}$ + lags (-5.74) (-4.89) (3.65)[++] (-0.74) (-0.97) (1.19) (-2.40)
Silver	121	1870	0	1892, 1972	-7.49**	$\Delta y_t = -.66\tilde{S}_{t-1} + 12.70 + .17 \cdot t - 2.24B_{1t} - 42.01B_{2t} - 60.46D_{1t} + 102.80D_{2t}$ + lags (-7.49) (0.74) (0.88) (-0.03) (-0.52) (-2.17) (3.29)
Tin	106	1885	6	1963, 1974	-6.97**	$\Delta y_t = -.72\tilde{S}_{t-1} + 4.70 - .08t + 51.75B_{1t} - 119.67B_{2t} - 6.07D_{1t} + 37.27D_{2t}$ + lags (-6.97) (0.99) (-1.54) (2.75) (-5.69) (-0.79) (3.44)
Zinc	121	1870	0	1917, 1952	-10.4**	$\Delta y_t = -.98\tilde{S}_{t-1} + 2.08 + .01t - 1.51B_{1t} - 4.23B_{2t} - 3.73D_{1t} - .91D_{2t}$ + lags (-10.4) (2.32) (0.33) (-0.41) (-1.14) (-2.53) (-0.68)

Notes: Critical values of the two-break minimum LM test with quadratic trend vary depending on the location of breaks (λ). Critical values are shown at the 1%, 5%, and 10% levels, respectively, for samples of size T = 100. Critical values are –6.80, –6.24, and –5.92 for λ = (.2,.4); –6.79, –6.19, and –5.89 for λ = (.2,.6); –6.68, –6.19, and –5.91 for λ = (.2,.8); –6.91, –6.27, and –5.89 for λ = (.4,.6); –7.01, –6.24, and –5.88 for λ = (.4,.8); –6.81, –6.19, and –5.88 for λ = (.6,.8), where λ indicates the relative location of the breaks. Other critical values can be interpolated. The critical values are symmetric around λ_1 and λ_2. B_{1t}, B_{2t} and D_{1t}, D_{2t} correspond to changes in intercept and trend slope under the alternative, respectively. B_{1t} and B_{2t} denote one-period jumps in level and D_{1t} and D_{2t} correspond to changes in drift under the null hypothesis, respectively. Note that the coefficient of t corresponds to a quadratic trend, and the constant term corresponds to a linear trend in Z_t, respectively, under the alternative. $\hat{k}$ denotes the estimated optimal number of first-differenced lagged terms included to correct for serial correlation. t-statistics are shown in parentheses. a: see notes in Table 1. * and ** denote significant at the 10% and 5% levels, respectively. [+] and [++] denote significance of the quadratic trend term at the 10% and 5% levels, respectively.

Table 5. One Break Minimum LM Unit Root Test with Quadratic Trend

Series	T	*Starting Year*	$\hat{k}$	$\hat{T}_B$	*Test Stat.*	Estimation Results [a]
Aluminum	90	1895	8	1913	-4.38	$\Delta y_t = -.70\tilde{S}_{t-1} - 24.07 - .08t - 26.40B_{1t} + 38.45D_{1t}$ + lags (-4.38) (-3.74) (-2.41) [++] (-2.15) (4.18)
Coal	121	1870	7	1972	-6.95**	$\Delta y_t = -.54\tilde{S}_{t-1} - .28 + .001t - 1.69B_{1t} + 1.39D_{1t}$ + lags (-6.95) (-2.33) (0.70) (-3.08) (5.16)
Copper	121	1870	1	1934	-5.33**	$\Delta y_t = -.43\tilde{S}_{t-1} + 2.34 - .03t - 1.12B_{1t} + 2.55D_{1t}$ + lags (-5.33) (1.69) (-1.78) [+] (-0.18) (1.14)
Iron	104	1870	3	1951	-6.31**	$\Delta y_t = -.70\tilde{S}_{t-1} + .63 - .06t - 8.42B_{1t} + 19.89D_{1t}$ + lags (-6.31) (0.24) (-1.83) [+] (-0.77) (3.85)
Lead	121	1870	5	1946	-5.30**	$\Delta y_t = -.75\tilde{S}_{t-1} + 1.59 - .01t + 6.81B_{1t} - .59D_{1t}$ + lags (-5.30) (2.83) (-2.08) [++] (3.55) (-0.86)
Gas	72	1919	7	1978	-7.33**	$\Delta y_t = -.46\tilde{S}_{t-1} + 5.32 - .09t - 4.78B_{1t} + 10.63D_{1t}$ + lags (-7.33) (4.88) (-5.05) [++] (-1.83) (4.68)
Nickel	78	1913	1	1977	-5.27**	$\Delta y_t = -.54\tilde{S}_{t-1} - 12.09 + .18t - 6.72B_{1t} - 14.30D_{1t}$ + lags (-5.27) (-2.62) (2.89) [++] (-0.39) (-2.01)
Petroleum	120	1870	6	1919	-5.30**	$\Delta y_t = -.47\tilde{S}_{t-1} - 1.81 + .02t + 1.95B_{1t} - .67D_{1t}$ + lags (-5.30) (-5.15) (4.84) [++] (2.31) (-2.17)
Silver	121	1870	2	1968	-4.31	$\Delta y_t = -.47\tilde{S}_{t-1} + 32.04 - .44t - 60.90B_{1t} + 114.58D_{1t}$ + lags (-4.31) (1.68) (-2.41) [++] (-0.72) (3.05)
Tin	106	1885	1	1973	-5.56**	$\Delta y_t = -.43\tilde{S}_{t-1} - 6.20 - .08t - 84.33B_{1t} + 7.28D_{1t}$ + lags (-5.56) (1.52) (-1.86) [+] (4.55) (1.00)
Zinc	121	1870	0	1917	-9.38**	$\Delta y_t = -.87\tilde{S}_{t-1} + 2.54 - .02t - 2.97B_{1t} - .59D_{1t}$ + lags (-9.38) (3.29) (-1.78) [+] (-0.77) (-0.44)

Notes: Critical values of the one-break LM test with a structural break in level and trend vary depending on the location of the break (λ). Critical values are shown at the 1%, 5%, and 10% levels, respectively, for samples of size T = 100. Critical values are –5.39, –4.86, and –4.57 for λ = .1; –5.31, –4.74, and –4.46 for λ = .2; –5.29, –4.78, and –4.50 for λ = .3; –5.35, –4.80, and –4.51 for λ = .4; –5.31, –4.77, and –4.49 for λ = .5, where λ indicates the relative location of the breaks. Other critical values can be interpolated. The critical values are symmetric around λ. B_{1t} and D_{1t} correspond to changes in intercept and trend slope under the alternative, respectively. B_{1t} denotes a one-period jump in level, and D_{1t} corresponds to a change in drift under the null hypothesis, respectively. Note that the constant term corresponds to a linear trend in Z_t under the alternative. $\hat{k}$ denotes the estimated optimal number of first-differenced lagged terms included to correct for serial correlation. *t*-statistics are shown in parentheses. a: See notes in Table 1. * and ** denote significant at the 10% and 5% levels, respectively. [+] and [++] denote significance of the quadratic trend term at the 10% and 5% levels, respectively.

Table 6. Simulation Results Comparing the Forecast Accuracy of Different Strategies

d	h	T = 100												T = 500							
		β = 1.0				β = 0.95				β = 0.5				β = 1.0				β = 0.5			
		M1/M3	M2/M3	M4/M3	M5/M3	M1/M3	M2/M3	M4/M3	M5/M3	M1/M3	M2/M3	M4/M3	M5/M3	M1/M3	M2/M3	M4/M3	M5/M3	M1/M3	M2/M3	M4/M3	M5/M3
0	1	0.99	1.10	0.98	1.02	0.99	1.07	0.97	0.98	1.14	0.99	1.12	0.94	1.00	1.02	1.00	1.01	1.23	1.00	1.23	0.99
	2	0.99	1.16	0.96	1.04	0.99	1.13	0.94	0.98	1.28	0.98	1.24	0.88	1.00	1.04	0.99	1.01	1.45	1.00	1.44	0.98
	5	0.99	1.21	0.92	1.07	0.99	1.19	0.89	0.95	1.46	0.97	1.35	0.80	0.99	1.09	0.98	1.03	1.73	1.00	1.71	0.97
	10	1.00	1.21	0.87	1.06	0.99	1.19	0.81	0.87	1.51	0.97	1.29	0.72	0.99	1.14	0.97	1.05	1.82	1.00	1.77	0.96
	50	1.00	1.27	0.66	0.97	1.00	1.34	0.46	0.56	1.74	0.94	0.81	0.36	0.99	1.21	0.88	1.05	2.13	1.00	1.82	0.93
	100	1.00	1.34	0.57	1.31	0.99	1.45	0.30	0.47	1.83	0.93	0.56	0.19	1.00	1.22	0.80	1.01	2.60	1.00	1.87	0.87
2	1	0.99	1.14	0.99	1.01	0.99	1.12	0.98	0.99	1.13	1.00	1.14	1.18	1.00	1.02	1.02	1.00	1.23	1.00	1.27	1.32
	2	0.98	1.22	1.00	1.03	0.98	1.18	0.97	1.00	1.28	0.99	1.29	1.37	1.00	1.05	1.05	1.00	1.47	1.00	1.57	1.59
	5	0.99	1.27	1.00	1.07	0.99	1.22	0.97	1.02	1.48	0.97	1.62	1.95	0.99	1.10	1.13	1.02	1.74	1.00	2.29	1.94
	10	0.99	1.25	1.02	1.09	0.99	1.20	0.97	1.00	1.54	0.97	2.20	2.97	0.99	1.16	1.25	1.05	1.82	1.00	3.92	2.14
	50	1.00	1.25	1.06	1.27	1.00	1.32	1.11	1.00	1.76	0.93	9.38	6.29	0.99	1.21	2.03	1.30	2.15	1.00	51.7	4.94
	100	1.00	1.30	1.07	2.33	1.00	1.43	1.28	1.42	1.88	0.92	14.75	7.25	1.00	1.21	2.74	1.68	2.69	1.00	188.2	10.4
6	1	0.98	1.47	1.10	0.99	0.97	1.42	1.09	0.98	1.10	0.99	1.26	1.14	1.00	1.04	1.19	1.03	1.24	1.00	1.57	1.35
	2	0.98	1.59	1.21	1.00	0.97	1.51	1.21	0.99	1.30	0.99	1.72	1.35	0.99	1.09	1.42	1.06	1.45	1.00	2.39	1.64
	5	0.99	1.52	1.60	1.07	0.98	1.43	1.60	1.04	1.53	0.97	3.66	1.60	0.99	1.17	2.13	1.15	1.71	1.00	6.50	2.29
	10	0.99	1.36	2.12	1.15	0.99	1.29	2.22	1.11	1.58	0.97	9.17	1.69	0.99	1.26	3.35	1.33	1.82	1.00	20.61	3.75
	50	1.00	1.21	4.21	1.88	1.00	1.24	6.26	1.84	2.00	0.94	77.61	2.30	0.99	1.27	11.21	3.08	2.21	1.00	449.6	61.3
	100	1.00	1.22	5.07	2.83	0.99	1.30	9.00	2.87	2.21	0.92	127.1	2.86	1.00	1.24	18.09	5.98	2.81	1.00	1663.	318.1
10	1	0.95	2.17	1.26	0.97	0.95	1.99	1.26	0.96	1.04	1.00	1.41	1.08	0.99	1.09	1.41	1.08	1.20	1.00	1.90	1.35
	2	0.96	2.18	1.63	1.02	0.96	2.01	1.65	1.01	1.27	0.99	2.38	1.34	0.99	1.17	1.94	1.16	1.42	1.00	3.54	1.76
	5	0.98	1.77	2.75	1.12	0.99	1.62	2.79	1.11	1.52	0.98	7.43	1.68	0.98	1.32	3.85	1.43	1.76	1.00	14.18	3.37
	10	0.99	1.45	4.27	1.30	0.99	1.35	4.57	1.26	1.63	0.97	22.79	2.10	0.98	1.43	7.10	1.89	1.85	1.00	52.15	8.21
	50	1.00	1.20	10.22	2.66	1.00	1.24	16.43	3.04	2.11	0.94	215.1	11.11	0.99	1.32	29.37	6.70	2.23	1.00	1217.	210.9
	100	1.00	1.20	12.93	4.11	1.00	1.28	24.70	5.36	2.34	0.93	350.1	23.07	1.00	1.25	48.39	14.68	3.00	1.00	4490.	1110.

Table 7. Comparison of Post-1990 PMSE using Different Models with Actual Data

(All prices are in 1998$ per ton)	*h*	Linear trend M1	M2	M4	M5	Quadratic trend M1Q	M2Q	M4Q	M5Q
Aluminum ($000)	1	688	**472***	656	497	995	**716**	2,368	2,168
	3	432	**192***	277	236	1,883	**1,235**	3,642	3,108
	5	299	**226***	386	383	1,989	**1,018**	3,522	2,908
	8	339	**215***	532	588	3,289	**1,182**	4,402	3,583
	12	718	**215***	1,055	948	6,095	**1,588**	6,111	4,966
Copper ($000)	1	278	**1***	135	69	158	265	**2**	437
	3	768	**8***	539	477	931	498	**385**	2,273
	5	554	**285***	377	419	773	333	**303**	2,000
	8	1,085	**246***	583	1,003	1,289	**550**	837	4,114
	12	1,958	**228***	921	1,960	2,334	**774**	1,899	7,051
Iron ($)	1	13	**0***	24	88	0	**1**	18	25
	3	23	**1***	148	195	3	**1**	46	118
	5	87	**26**	281	315	43	**15***	60	206
	8	347	**155**	382	401	186	87	**53***	295
	12	758	**342**	476	631	371	176	**77***	508
Lead ($000)	1	245	99	366	**32***	156	152	194	**126**
	3	355	**74***	704	204	**166**	171	251	186
	5	346	**48***	853	179	**136**	156	232	214
	8	293	**33***	962	164	**102**	166	194	167
	12	291	**22***	1,207	210	**120**	314	220	165
Nickel ($000,000)	1	5	19	**1***	3	**1**	4	22	36
	3	23	89	46	**9***	**18**	25	108	106
	5	107	310	50	**9***	83	104	**77**	97
	8	287	816	72	**9***	228	278	**110**	114
	12	685	1,922	98	**9***	550	657	**135**	141
Silver ($000,000)	1	666	**359***	4,959	5,885	**392**	752	21,031	46,651
	3	**4,832***	5,481	9,743	13,276	5,967	**5,854**	61,788	196,423
	5	16,589	19,602	**8,236***	13,270	21,256	**19,586**	86,698	334,376
	8	42,002	45,291	**7,936***	15,868	49,476	**43,546**	120,814	446,318
	12	84,740	87,582	**9,185***	22,059	96,442	**81,707**	167,586	371,394
Tin ($000,000)	1	2	19	**0***	2	16	1	**0**	10
	3	6	69	**3**	16	68	16	**2***	16
	5	19	103	**4**	20	150	45	**1***	39
	8	47	199	**7**	35	332	81	**1***	69
	12	99	348	**15**	60	668	155	**1***	122
Zinc ($000)	1	305	246	271	**28***	239	295	**170**	523
	3	345	475	435	**50***	362	452	**303**	1,120
	5	358	447	518	**80***	457	524	**375**	1,386
	8	366	474	592	**69***	561	584	**468**	1,576
	12	487	640	814	**139***	854	783	**753**	1,892

Notes: PMSE is the predicted mean squared error of the post-1990 out-of-sample forecast period (h). M1 = differences with breaks, M2 = levels with breaks, M4 = differences without breaks, M5 = levels without breaks, and Q denotes that the model includes a quadratic trend instead of linear trend. Structural breaks were jointly significant in all models that included breaks at $p < 0.01$. * denotes the model with the lowest PMSE.

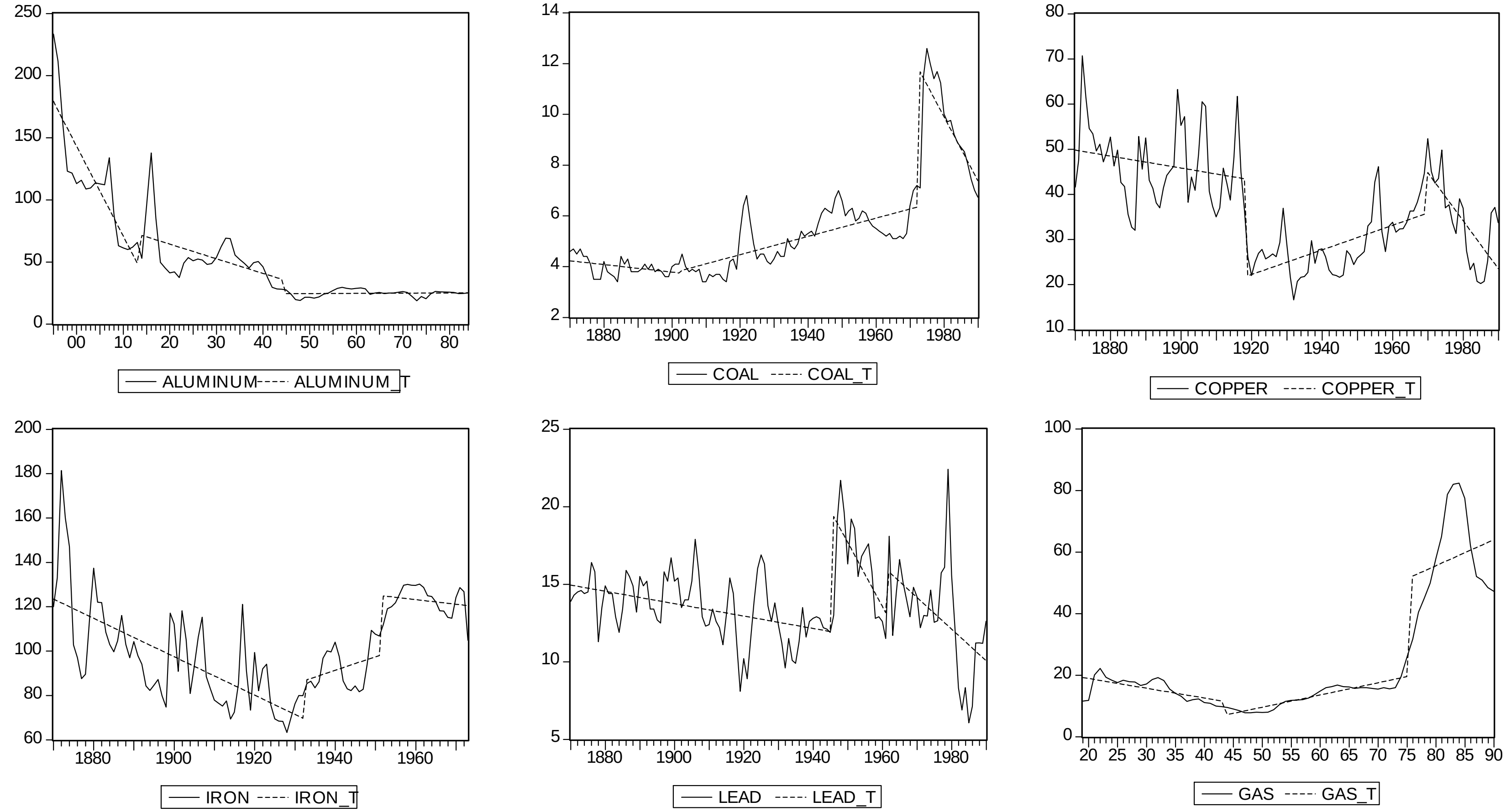


Figure 1. Plot of Data with Structural Breaks and Fitted Trends

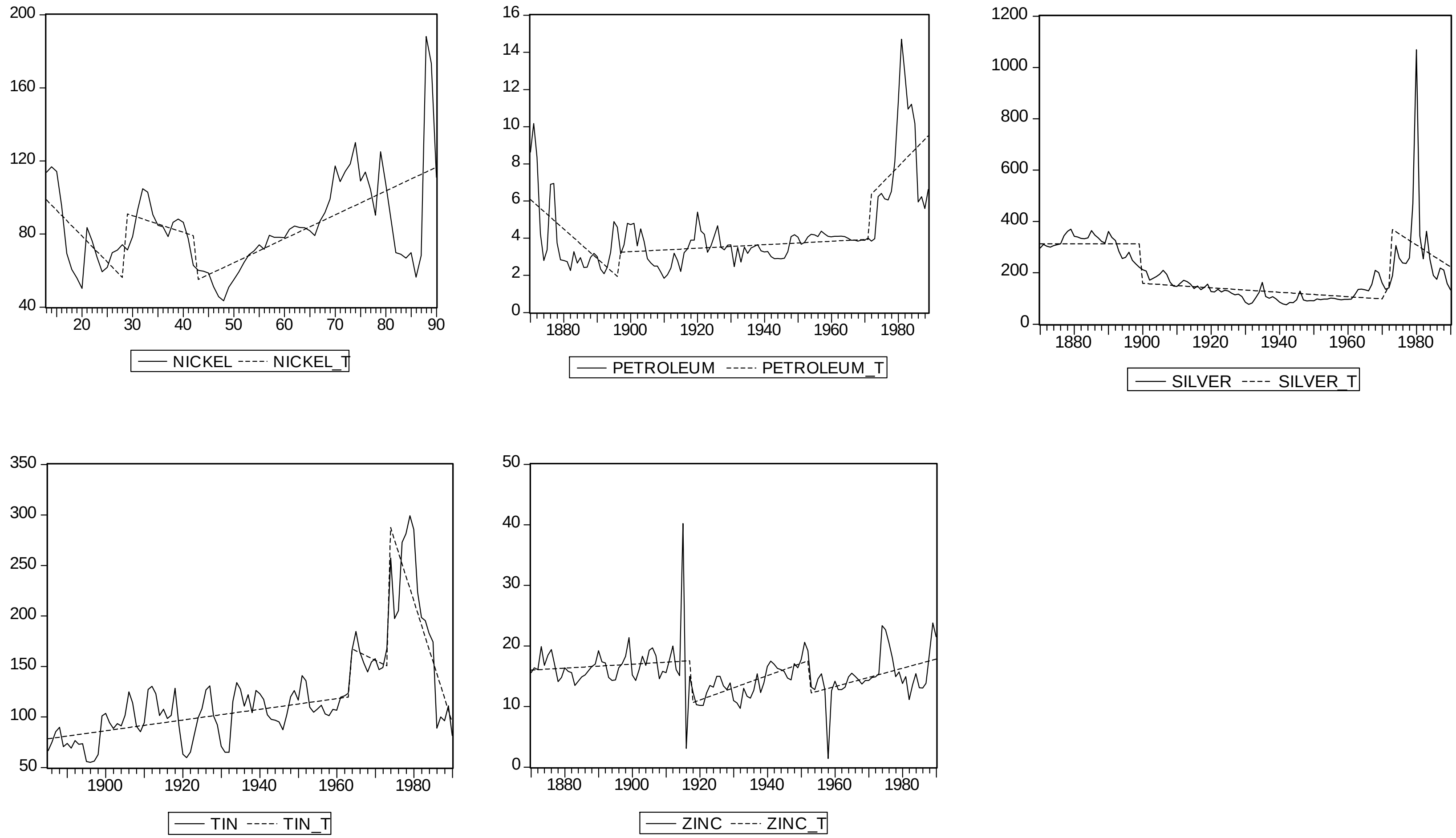
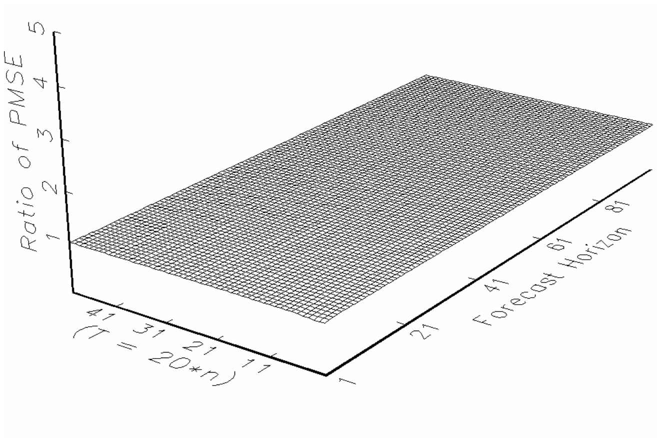
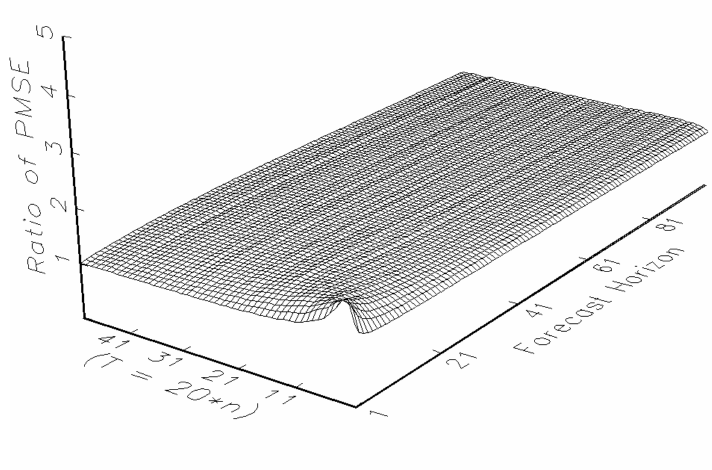


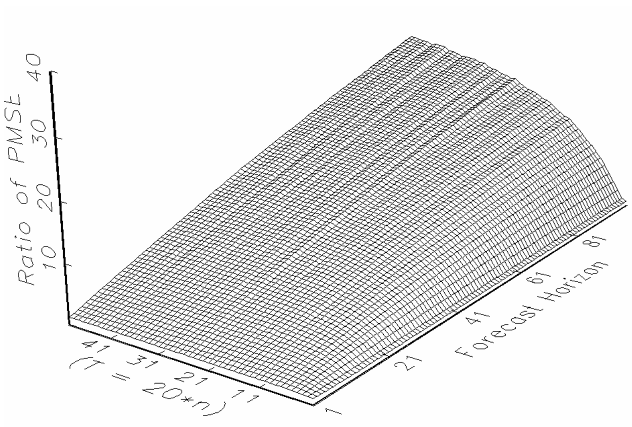
Figure 1 (continued). Plot of Data with Structural Breaks and Fitted Trends



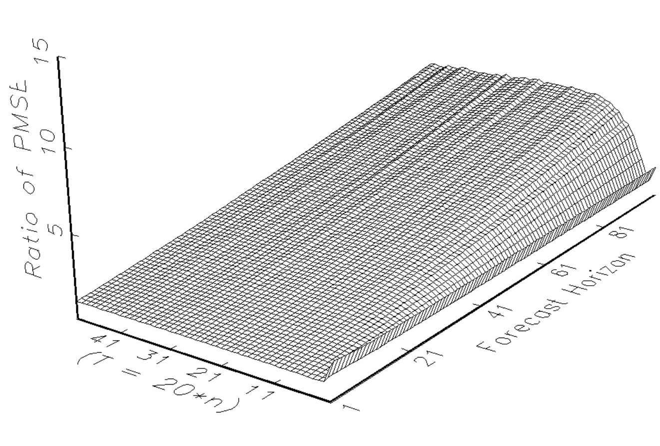
(a) Difference with Breaks (M1 / M3)



(b) Level with Breaks (M2 / M3)

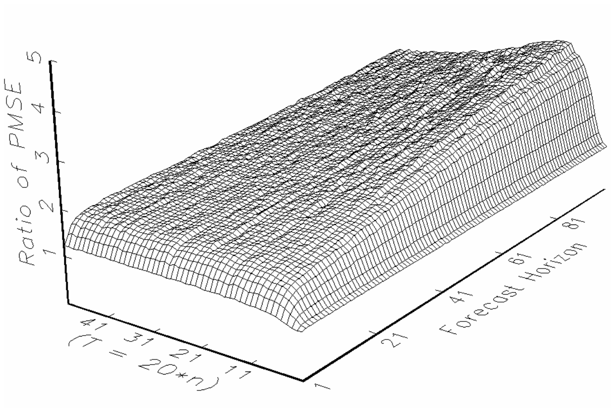


(c) Difference without Breaks (M4 / M3)

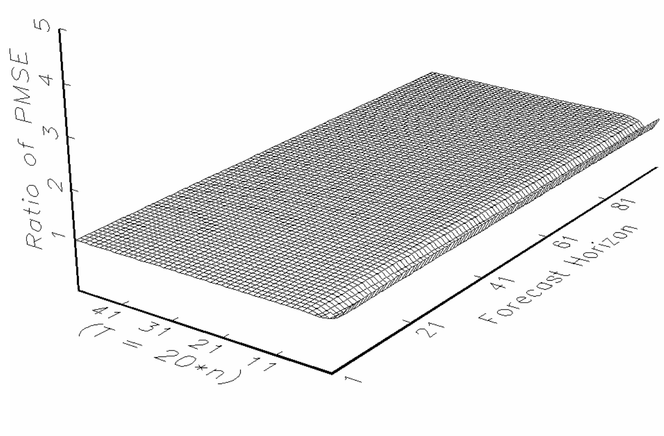


(d) Level without Breaks (M5 / M3)

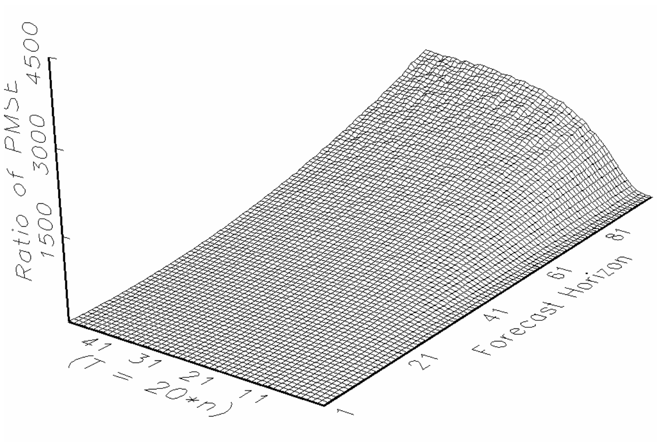
Figure 2. Comparison of Forecast Accuracy using Different Strategies (DGP: $\beta = 1$, $d = 6$)



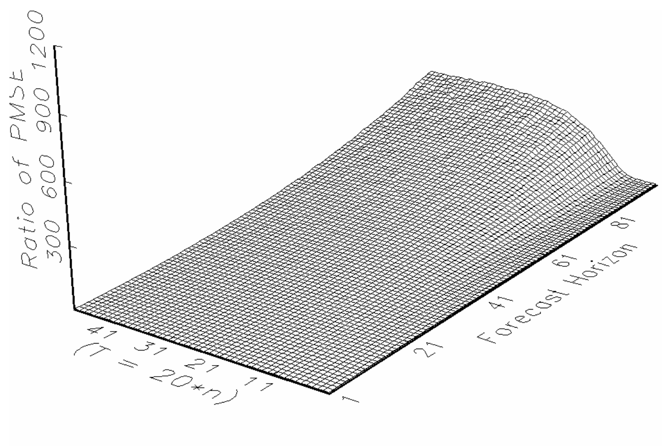
(a) Difference with Breaks (M1 / M3)



(b) Level with Breaks (M2 / M3)



(c) Difference without Breaks (M4 / M3)



(d) Level without Breaks (M5 / M3)

Figure 3. Comparison of Forecast Accuracy using Different Strategies (DGP: $\beta = 0.5$, $d = 6$)