CHALLENGES OF TRENDING TIME
SERIAL ECONOMETRICS

BY

PETER C. B. PHILLIPS

COWLES FOUNDATION PAPER NO. 1151

Challenges of trending time series econometrics[☆]

Peter C.B. Phillips[a,b,c]

[a] *Department of Economics, Cowles Foundation, Yale University, PO Box 208281, New Haven, 06520-8281 CT, USA*
[b] *University of Auckland, Auckland, New Zealand*
[c] *University of York, UK*

Abstract

We discuss some challenges presented by trending data in time series econometrics. To the empirical economist there is little guidance from theory about the source of trend behavior and even less guidance about practical formulations. Moreover, recent proximity theorems [W. Ploberger, P.C.B. Phillips, Empirical limits for time series econometric models, Econometrica 71 (2003) 627–673] reveal that trends are more elusive to model empirically than stationary processes, with the upshot that optimal forecasts are also harder to estimate when the data involve trends. These limitations are implicitly acknowledged in much practical modeling and forecasting work, where adaptive methods are often used to help keep models on track as trends evolve. The paper discusses these broader issues and limitations of econometrics and offers some thoughts on new practical possibilities for data analysis in the absence of good theory models for trends. In particular, a new concept of coordinate cointegration is introduced and some new econometric methodology is suggested for analyzing trends and co-movement and for producing forecasts in a general way that is agnostic about the specific nature of the trend process. Some simulation exercises are conducted and some long historical series on prices and yields on long securities are used to illustrate the methods.
© 2005 IMACS. Published by Elsevier B.V. All rights reserved.

"It is like giving the play of Hamlet without Hamlet to eliminate the secular trends of *i* and *P* from a study of long term relationships in which these very secular trends are most important and dominant influences." Fisher [5, p. 432]

1. Introduction

A distinguishing characteristic of most economic time series is trending behavior. Such time series often behave either in a wandering manner with long and erratic cycles (as in the case of interest rates and stock prices) or as if they were influenced by some secular drift over time (like many national income components). Much of modern time series econometrics is concerned with the statistical analysis of such properties, including the possible interconnectedness of the trends across different series. In spite of many decades of research in fields like monetary theory and economic growth, economics provides little guidance about the source of such trends and even less guidance concerning suitable formulations for practical work. Indeed, trend formulations that appear in economic theory models are often based on mathematical convenience and/or an appeal to some broadly acknowledged steady state characteristic, including the so-called 'great ratios of macroeconomics'. Such characteristics are themselves often based on simple long run data averages and reflect in a primitive way some commonality in the trending behavior of multiple series.

In practice, therefore, while many economists see trends in the data, the econometric modeling of such trends is a much more difficult task. It is also a task where failure has major implications in forecasting. The trend is often regarded as a dominant feature of the data (as in the headnote citation of Irving Fisher leading the article) and if the trend mechanism is poorly captured in an empirical model, we can expect forecasts from the model to carry forward the poor approximation. In practical work we are very accustomed to this phenomenon—as the forecast horizon is extended and observations are subsequently collected and calibrated against the forecasts, the data drift steadily away from the given model. In short, one of the laws of modern time series econometrics [20] is that '*no one understands trends, but everyone sees them in the data*'.

Fig. 1 graphs yields on long securities and logarithms of producer prices in the UK over 1720–2002. The figure is split into two periods, the first panel showing 1720–1939 and the second 1940–2002. These series show a remarkable commonality in their movement over this long historical period. Especially over 1720–1939, there are long subperiods in these two centuries where the series rise and fall together over time. This empirical phenomenon has long been noted. In his *Treatise on Money*, Keynes [10] described the apparent co-movement as

'one of the most completely established empirical facts within the whole field of quantitative economics, though theoretical economists have mostly ignored it'. Keynes [10, p. 198]

Keynes, Fisher [5] and many others since have put forward many possible theoretical explanations of the phenomenon. All explanations share the common ground that the phenomenon is dynamic. Economists agree that high price levels and high interest rate levels may well be associated, as they appear to be in Fig. 1, but they also agree that interest rates are not high because prices are at a high level and vice versa. It is, instead, the transitions between the levels that are important. As Fisher explained it,
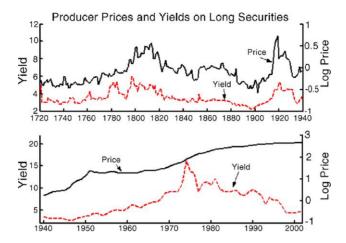
Fig. 1. UK producer prices and long security yields 1720–2002.

"at the peak of prices, interest is high, not because the price level is high, but because it has been rising and, at the valley of prices, interest is low, not because the price level is low, but because it has been falling." Fisher [5, p. 441]

Fisher noted the trending behavior of these series and considered the trends to be a vital element in their relationship. Over 1720–1939 the series appear to trend together in a systematic way. Since 1940, prices have risen persistently, but at very different rates of inflation over subperiods, whereas yields have both risen (to the mid-1970s) and fallen. Recent inflation targeting and cash rate management policies of the Bank of England and other monetary authorities give good reason to expect some change in the relationship between these variables, leading to stronger links between inflation and yields. Both empirical and institutional evidence therefore indicate that the long run historical relationship between prices and rates may have altered since 1940. This is a matter than can be tested empirically.

A primary limitation of empirical econometric work is that the true model for such data is unknown and probably unknowable. In the present case, modeling the apparent trending behavior is particularly difficult, more so because the trends seem to consist of long cycles of upswings and downswings punctuated by subperiods in which the movement is in different directions. A remarkable feature of commonality of the series over the period 1720–1939 is that the co-movement appears to include much of the subperiod behavior as well as the longer lived trends, a matter on which we will comment further below.

The present paper discusses the challenge presented by trending data when, as in the example just outlined, there is little guidance from theory about the source and nature of the trending behavior. In such cases, it is possible to represent the trend in terms of coordinate functions that capture both trend and cyclical behavior. Using this agnostic coordinate function approach, it is possible to analyze series co-movement such as that observed above between yields and prices. We present some new methods for doing so and discuss some recent research on the limitations of modeling and forecasting with trending data.

2. No one understands trends

The absence of a rich theoretical framework for modeling trends partly explains the rather impoverished class of trend formulations that appear in applied econometric work. The most commonly used models are polynomial time trends, simple trend break polynomials, and stochastic trends, which include unit root models, near unit root models and fractional processes. Occasionally, nonparametric trend specifications are used. These models are often stated in a components format as $X_t = k_t + u_t$ $(t = 1, \ldots, n)$, where X_t is the observed process, the component u_t is commonly taken to be stationary (or a heterogeneous short memory process) and k_t is the 'trend' process. No encompassing definition of the latter has yet been developed in the literature. Instead, it is customary in econometrics to work with specific formulations, which generally fall within the models just listed, and to assume that these are sufficiently general to include the true process. Another practice that is popular in applied econometric work is the use of smoothing filters to extract the trend and then proceed to analyze the residual time series.

The trend process k_t is often regarded as the component about which short or medium run fluctuations (like business cycles) occur. It may therefore be characterized in the frequency domain in terms of its (dominating) low frequency behavior. To ensure a convenient level of generality, it is possible to employ a definition that characterizes this dominating behavior in terms of the order of magnitude of the discrete Fourier transform (dft) of k_t as the sample size $n \to \infty$. Write the dft as $w_k(\lambda) = (2\pi n)^{-1/2} \sum_{t=1}^{n} k_t \, \mathrm{e}^{it\lambda}$ and let the fundamental frequencies be $\lambda_s = 2\pi s/n$ $(s = 1, \ldots, n)$. Then, if k_t is a regularly varying function at infinity of the form μt^α for some $\alpha > 0$ (representing a deterministic power function trend), we find that

$$w_k(\lambda_s) = \frac{\mu}{\mathrm{e}^{i\lambda_s} - 1} \frac{n^\alpha}{\sqrt{2\pi n}}[1 + o(1)] = O_p\left(\frac{n^{\alpha+1/2}}{s}\right) \text{ as } n \to \infty$$

(Lemma A of [3]). Or, if there is both power function drift and integration (of possibly fractional order $d > 1/2$) in the trend process $k_t = \mu t^\alpha + k_t^0$, where $k_t^0 = (1 - L)^{-d} u_t$ and L is the lag operator, then

$$w_k(\lambda_s) \simeq \frac{\mu}{\mathrm{e}^{i\lambda_s} - 1} \frac{n^\alpha}{\sqrt{2\pi n}} + \frac{1}{(1 - \mathrm{e}^{i\lambda_s})^d} w_u(\lambda_s) - \frac{\mathrm{e}^{i\lambda_s}}{1 - \mathrm{e}^{i\lambda_s}} \frac{k_n^0}{\sqrt{2\pi n}},$$

$$w_k(\lambda_s) = O_p(\lambda_s^{-1} n^{\alpha-1/2}) + O_p(\lambda_s^{-d}) + O_p(\lambda_s^{-1} n^{d-1}) \tag{1}$$

(Lemma A7 of [23]). In such cases as these, the dft of k_t has the general form

$$w_k(\lambda_s) = O_p\left(\frac{n^\beta}{s^\gamma}\right), \quad \text{for some } \beta > \tfrac{1}{2}, \ \gamma \in (0, 1] \tag{2}$$

uniformly for $s = 1, \ldots, m$, where $m = o(n)$, the latter condition focusing attention on low frequency behavior as $n \to \infty$. The limit behavior as $n \to \infty$ may be given for both the trend process and its dft. For instance, under certain regularity conditions, we will have asymptotic behavior of the form

$$n^{-\max(\alpha, d-1/2)} k_{[nr]} \overset{d}{\to} k(r) = \mu r^\alpha \mathbf{1}_{\{\alpha \geq d-1/2\}} + B_{d-1}(r)\mathbf{1}_{\{\alpha \leq d-1/2\}}, \tag{3}$$

where B_{d-1} is fractional Brownian motion, where [ ] represents the integer part of the argument and where $r \in [0, 1]$; and if s is fixed as $n \to \infty$, the dft behaves as follows:

$$n^{-\max(\alpha+1/2,d)} w_k(\lambda_s) \xrightarrow{d} (2\pi)^{-1/2} \int_0^1 e^{2\pi i s r} B_{d-1}(r) \, dr \mathbf{1}_{\{\alpha \leq d-1/2\}} + \frac{\mathbf{1}_{\{\alpha \geq d-1/2\}}}{(2\pi)^{3/2} s}$$

(Theorem 3.7c of [18]). Limit (3) indicates that a central characteristic of trend is that trajectory behavior present in $k_{t=[nr]}$ persists in the limit as the sample size $n \to \infty$. In particular, upon standardization by a suitable function of n, where that function itself depends on the nature of the trend process, the trajectories in the data converge to limiting random functions with properties analogous to those in the data. Thus, in the case above, r^{α} and $B_{d-1}(r)$ are deterministic and random functions, respectively, on the space $L_2[0, 1]$ of square integrable functions on the interval $[0, 1]$, whereas t^{α} and $k_t^0 = (1 - L)^{-d} u_t$ are deterministic and stochastic trends on the real line $\mathbb{R}$.

When the focus is on trend elimination (for instance, in the extraction of the cyclical component of a series for the study of business cycles), smoothing methods are common. The most prominent of these is based on ideas developed originally by Whittaker [30] on smoothing filters for graduating time series and his method has become commonly known in macroeconomics as the Hodrick–Prescott filter following work by Hodrick–Prescott [7] that utilized these techniques with macroeconomic data. These and other methods like spline smoothing [25,29] and band-pass filtering [2,3] all provide practical mechanisms for dealing with trends in data. But it is unrealistic to pretend that these formulations and filters explain the process by which trends actually occur in the real world. In short, no one really understands trends, even though most of us see trends when we look at economic data.

One nearly universal consequence of trends in the data is an empirical regression phenomena called 'spurious regression'. In effect, any trend function that we postulate in an econometric specification will turn out to be statistically significant in large samples provided the data do in fact have a trend, whether it is of the same form as that specified in the empirical regression or belongs to a general alternative such as (2). The history of this subject has early beginnings dating back to the formal discussion by Yule (1926) [31] and beyond. Over the last two decades, the case of multivariate regression of unit root processes (including independent random walks) was analyzed in [15] and that of polynomial trends fitted to stochastic trends by Durlauf and Phillips [4]. In both cases, the fitted regression coefficients turn out to be statistically significant with probability one asymptotically and this is true even when robust standard errors that account for residual autocorrelation are used to assess significance [17]. Some related results have been shown to hold for trend breaks, fractional processes and regressions among such variables (e.g., [1,11–14,28]) even when they are stochastically independent, the latter being the phenomenon originally examined in simulations by Granger and Newbold [6].

The nomenclature 'spurious regression' carries a pejorative connotation and empirical researchers are usually anxious to avoid such specifications and to show that fitted relationships are validated by some process such as cointegration testing. Of course, even when cointegration between series is empirically validated, the issue of modeling the common trend in the series usually remains, or it is dealt with only by indirect means, e.g., by fitting a reduced rank regression with some unit roots and possibly a drift.

In his early studies of the correlation between interest rates and prices, Fisher [5] recognized these potential difficulties, arguing that

"it is necessary to guard against the possibility that these coefficients are of the familiar nonsense type, and are spuriously high because of the presence of *secular trend* [Fisher's emphasis] forces that affect both *P* and *i*." Fisher [5, p. 431]

While acknowledging that for these variables 'it is rather doubtful that trend forces are involved which should be eliminated', Fisher extracted both linear and quadratic trends over subperiods of the data (effectively performing a trend break analysis of the data) and took correlations of the residuals, finding results that were 'interesting and amazing' giving correlations between the series that were still significantly high, corroborating co-movement of interest rates and prices. Fisher concluded:

"The elimination of the secular trends from the comparisons makes the relationship of *i* and *P* depend solely upon the similarity of fluctuations in the shorter or cyclical periods. Even without Hamlet the play proves to be astonishingly informing and interesting. It is quite definitely demonstrated that, in times of marked price changes, as in the World War period, the effects of price movements are felt rather quickly upon the rates of interest, even in the case of long term bond yields." Fisher [5, p. 438]

3. Co-movement

Presently available econometric methods make alternative methods of analysis possible. To begin, it is obviously of interest to test the apparent co-movement of the series directly. Tests of cointegration enable us to assess the evidence for co-movement directly without concern for short run dynamic effects. Figs. 2 and 3 show recursive calculations of the residual based Z_t and Z_α tests of cointegration [22] involving the long security yields and the logarithm of wholesale (producer) prices shown in Fig. 1.

The tests both confirm the presence of cointegration at the 5% between the variables and both statistics advance further into the tail of the distribution as we move through the data up to around 1940 at which point there is an abrupt change. By the 1970s the statistics reject cointegration, giving statistical confirmation of a change in the relationship between prices and yields in the post war period that was suggested earlier by inspection of the data.

Similarly, when a linear relationship between yields and prices is fitted recursively by fully modified least squares [21] over 1750–2000, we find that the slope coefficient, which is shown in Fig. 4,
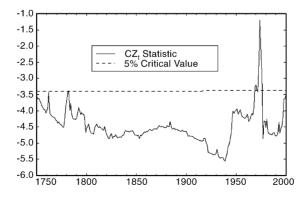

Fig. 2. Recursive values of Z_t test for cointegration of yields and log prices 1750–2000.
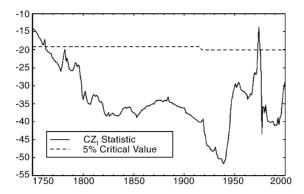
Fig. 3. Recursive values of Z_α test for cointegration of yields and log prices 1750–2000.

is quite stable over one and a half centuries from 1800–1940. Fig. 5 shows recursive estimates of the same slope coefficient obtained by low frequency band least squares [16], where the coefficients have a similar recursive form over the sample period. After 1940, the coefficient estimates (from both procedures) fluctuate considerably. There is also considerable coefficient fluctuation over 1750–1800, which can in part be explained by the shorter sample period and in part by substantial short term fluctuations in prices and yields over this period that are not always in concert, as is evident in the data. As might be expected, the low frequency coefficient estimates (which eliminate high frequency components in the regression) in Fig. 5 show less fluctuation in the early part of the sample. The evidence from the cointegration tests and the coefficient estimates therefore both point to a stable long term relationship (with a coefficient around 2.0–2.5) between yields and prices over the period 1800–1940.

It is also of interest to assess the degree of nonstationarity in the two variables. One way of doing so is to estimate the memory parameter (i.e., the index of fractional integration) in each case. Fig. 6 shows recursive estimates of the fractional integration parameter (d) for both series together with upper and lower 95% confidence limits. The estimates are obtained using the exact local Whittle estimator of [27] with a bandwidth of $m = n^{0.7}$ frequencies around the origin. This estimator is consistent and the (asymptotic) confidence interval is valid for all values of stationary and nonstationary d.
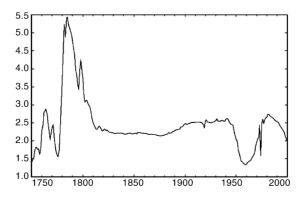

Fig. 4. Recursive values of fully modified least squares estimates of cointegrating coefficient of yields on prices 1750–2000.
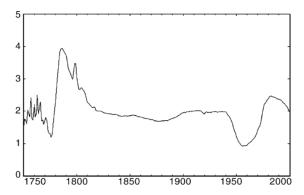
Fig. 5. Recursive values of low frequency band least squares estimates of cointegrating coefficient of yields on prices 1750–2000.

The estimates of *d* for yields are generally in the range 0.8–0.95 and appear to be fairly stable around 0.85 for much of the period 1820–1960. After 1960, the estimates of *d* rise to above unity and then fall back down again to around 0.85 by the year 2000. The (pointwise) confidence band for *d* includes unity, so we cannot reject the hypothesis of a unit root in yields throughout the period.

For prices, the estimates of *d* are in a similar range 0.75–1.0 and again are fairly stable around the level 0.8 over the period 1800–1940 with the exception of a small blip upwards during the first world war. After 1940, the long memory estimates for prices rise to around 0.9 and continue rising thereafter towards unity around the year 2000. The upper limit of the 95% confidence band is around unity for much of the period to 1940 and thereafter rises and includes unity from around 1950 onwards.

The differences in the memory properties of the two series therefore appear to be small over much of the period. In both cases the memory parameter appears to be generally less than unity. There are indications of some instability in the estimates over the final part of the period from 1940, corroborating the earlier evidence reported.
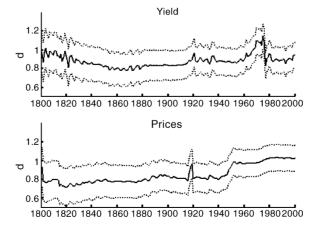

Fig. 6. Recursive exact local Whittle estimates of the long memory parameter *d*.

4. Coordinate cointegration

It has recently been suggested in [17] that deterministic trend functions can be used as a coordinate system for measuring the trend behavior of an observed variable, much as one set of functions can be used as a coordinate basis for studying another function. Thus, any function $f \in L_2[0, 1]$ can be written in terms of an orthonormal basis $\{\varphi_k\}_{k=1}^{\infty}$ as $f(x) = \sum_{k=1}^{\infty} c_k \varphi_k(x)$. Continuous stochastic processes such as Brownian motion and diffusions also have representations in terms of the functions φ_k but with coefficients c_k that are random variables rather than constant Fourier coefficients. Such formulations can be given a rigorous function space interpretation in terms of functional representations of the limiting stochastic processes or deterministic functions to which standardized versions of the trending data or trend functions converge.

To fix ideas suppose that X_t is a stochastic trend with $\Delta X_t = u_t$ and that partial sums of the stationary process u_t satisfy the functional law $n^{-1/2} \sum_{k=0}^{[n\cdot]} u_t \to_d B(\cdot)$, a limit Brownian motion process, where $[\ ]$ represents the integral part of the argument. If the initial condition $X_0 = o_p(n^{1/2})$, we then have $n^{-1/2} X_{[n\cdot]} \to_d B(\cdot)$. When X_t is a p-vector, B is a vector Brownian motion with covariance matrix $\Omega = \sum_{h=-\infty}^{\infty} E(u_t u_{t+h})$. If Ω is singular then X_t is cointegrated with cointegration matrix β, where β spans the null space of Ω, and $\beta' B(\cdot) = 0$ with probability 1.

The limit stochastic process $B(r)$ has an almost sure unique representation in terms of deterministic functions over the interval $r \in [0, 1]$. It is particularly convenient to use the orthonormal functions corresponding to the covariance kernel of B and this leads to the Loève Karhunen representation

$$B(r) = \sqrt{2} \sum_{k=1}^{\infty} \frac{\sin[(k - 1/2)\pi r]}{(k - 1/2)\pi} \xi_k = \sum_{k=1}^{\infty} \lambda_k^{1/2} \varphi_k(r) \xi_k, \tag{4}$$

where the components ξ_k are iid $N(0, \Omega)$ and $\varphi_k(r) = \sqrt{2} \sin[(k - 1/2)\pi r]$. This series representation of $B(r)$ is convergent almost surely and uniformly in $r \in [0, 1]$. Let ξ_K, and $\varphi_K(r)$ be K-vectors of the first K elements of $\{\xi_k\}$ and $\{\varphi_k(r)\}$, respectively, and $\xi_\perp$, and $\varphi_\perp(r)$ be vectors of the remaining elements of these sequences. Then, we may write (4) as

$$B(r) = \Xi_K \Lambda_K^{1/2} \varphi_K(r) + \Xi_\perp \Lambda_\perp^{1/2} \varphi_\perp(r), \tag{5}$$

where $\Lambda_K = \mathrm{diag}(\lambda_1, \ldots, \lambda_K)$, $\Lambda_\perp = \mathrm{diag}(\lambda_{K+1}, \lambda_{K+2}, \ldots)$ and

$$\Xi_K = [\xi_1, \ldots, \xi_K], \qquad \Xi_\perp = [\xi_{K+1}, \xi_{K+2}, \ldots].$$

Note that the coefficient of the deterministic function $\varphi_k(r)$ in (4) is of order $O_p(1/k)$, so that the functions in the representation become less important as k gets large.

The relationship (5) can be fitted empirically using observations $\{X_t : t = 1, \ldots, n\}$ in the linear regression

$$X_t = \sum_{k=1}^{K} \hat{b}_k \varphi_k \left(\frac{t}{n} \right) + \hat{u}_{tK} \tag{6}$$

or, equivalently (with $\hat{a}_k = n^{-1/2}\hat{b}_k$),

$$\frac{X_t}{\sqrt{n}} = \sum_{k=1}^{K} \hat{a}_k \varphi_k \left(\frac{t}{n}\right) + \frac{\hat{u}_{tK}}{\sqrt{n}} := \hat{a}'_K \varphi_K \left(\frac{t}{n}\right) + \frac{\hat{u}_{tK}}{\sqrt{n}}. \tag{7}$$

Phillips [19, lemma 2.2] showed under weak regularity conditions that as $n \to \infty$

$$\hat{a}_K = \Xi_K \Lambda_K^{1/2} + o_{\text{a.s.}}(1), \tag{8}$$

so that the empirical estimates $\hat{a}_K$ asymptotically reproduce the coefficients in the LK representation (5).

In contrast to conventional regression asymptotics, the coefficients *and their limits* as $n \to \infty$ are random variables. In fact, if $K \to \infty$ as $n \to \infty$ with $K/n \to 0$, (7) succeeds in reproducing the entire series (5).

Now, if the elements of X_t are cointegrated with cointegrating vector β, B is degenerate in the sense that Ω is singular and $\beta'\Omega = 0$. Correspondingly, we have from (5) that

$$\beta'\Xi_K = 0, \qquad \beta'\Xi_\perp = 0. \tag{9}$$

We call (9) coordinate co-movement or coordinate cointegration because it implies that the coordinate coefficients in the LK representation satisfy the same linear equations when the series elements co-move or cointegrate. If only the first K of these relations held, we refer to it as K-coordinate co-movement.

Since the elements of $\Xi = [\Xi_K, \Xi_\perp]$ can be estimated empirically, as seen in (8), we can in fact attempt to estimate the cointegrating vector β from the fitted elements $\hat{a}_K$. A more direct approach is to build the hypothesis of cointegration into the structure of (7). The model then involves restrictions on the random coefficients and these can be dealt with by means of a reduced rank regression of the following form:

$$\frac{X_t}{\sqrt{n}} = \hat{\alpha}_K \hat{\gamma}'_K \varphi_K \left(\frac{t}{n}\right) + \frac{\hat{v}_{tK}}{\sqrt{n}}, \tag{10}$$

$$= \frac{\hat{X}_{tK}}{\sqrt{n}} + \frac{\hat{v}_{tK}}{\sqrt{n}}, \tag{11}$$

where $\hat{\alpha}_K$ is $p \times s$, $\hat{\gamma}'_K$ is $s \times K$ and $s = p - r$, where r is the cointegrating rank (the rank of the matrix Ω and the number of columns, or linearly independent cointegrating vectors, in β). $\hat{X}_{tK}$ in (11) is the fitted value of X_t using the K coordinate functions φ_K and allowing for co-movement by virtue of the reduced rank coefficient matrix in (10). The author has been able to show that under weak conditions

$$\hat{\alpha}_K \to_p \beta_\perp,$$

where $\beta_\perp$ is the orthogonal complement of β, from which a consistent estimate of β (subject to normalization) can be obtained.

In addition to the coordinate RRR approach, the author has developed a coordinate instrumental variable method of estimation that uses a coordinate basis such as $\{\varphi_k\}_{k=1}^{K}$ as instruments and allows $K \to \infty$ as $n \to \infty$. Analysis that will be reported elsewhere shows that such IV methods with 'large numbers' of instruments produce asymptotically efficient estimates of the cointegrating matrix β in the sense defined in [16] when the model is a Gaussian $I(1)/I(0)$ cointegrated system. This approach can be interpreted as a semiparametric IV regression and, interestingly, is a one step approach to efficient estimation. Simulations
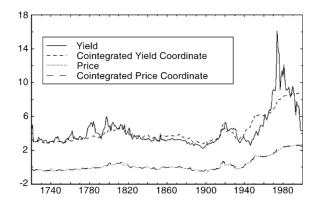
Fig. 7. Yields and prices against cointegrated coordinate system of deterministic functions 1718–2002.

indicate that this approach also seems to have better finite sample performance characteristics than VAR-based reduced rank regression (particularly when the order of the VAR is overspecified) and FM-OLS. Some simulations confirming this are reported later.

Fig. 7 shows the results of applying coordinate RRR regression to the series of yields and prices, allowing for the presence of cointegration ($r = 1$). Here K was chosen to be $K = 175$. The degree of co-movement that is picked up by these coordinate functions is remarkable. Movement in the price series is captured closely because the number of deterministic regressors (K) is large. The corresponding trend behavior of the yield series is also followed quite closely even though only one extra parameter is fitted. As is clear from the figure, the fitted values $\hat{X}_{tK}$ reveal the common trending component in the series rather well over most of the historical period. It is only in the period post 1940 that the co-movement appears to break down, corroborating findings from earlier in the paper.

This approach of trend coordinatization has many advantages. The method is not restricted to series that have unit roots and may be applied to any series that are nonstationary. While the deterministic coordinate functions used here are convenient, others can be used if they are more suited to the series under study. It is also not necessary that the series have the same stochastic order or rates of convergence to a limiting stochastic process. In addition, it is possible to use trend coordinatization for trend extraction purposes. Fitted trend functions obtained by this method may then be used in conventional tests for co-movement and estimation. Finally, one can limit the coordinate function co-movement to any finite number (K) of coordinates. Full coordinate co-movement occurs when the reduced rank structure applies for $K \to \infty$ as $n \to \infty$. Partial coordinate co-movement occurs when the reduced rank structure applies only for finite K.

It is appealing that the approach provides a mechanism for relating variables of different stochastic order (like time polynomials and random walks) so that it can be used to justify relationships between observed variables which have differing memory characteristics, overcoming the apparent 'problem' of relationships between stochastically imbalanced variables. Interestingly, the approach also gives consistent estimates of cointegrating structural coefficients even though it is based on empirical regressions that are typically thought of as being 'spurious' such as the regression of stochastic trends on deterministic trends. Hence, relationships between trending variables that are often deemed spurious actually carry a great deal of useful information and can be used directly for consistent estimation.

5. Trend coordinates and prediction

To study the evolution of a trend process over time, we may extend some of the above ideas and use an orthogonal representation over a subinterval of $[0, 1]$ such as $[0, r]$ for some arbitrary fraction $r < 1$ of the sample data. Working with unit root data where the limiting trend process is the Brownian motion B, we find, instead of (4),

$$B(s) = \sum_{k=1}^{\infty} \varphi_k \left(\frac{s}{r} \right) \eta_k(r), \quad \text{for } s \in [0, r], \tag{12}$$

where $\eta_k(r) = (1/r) \int_0^r B(s) \varphi_k(s/r) \, \mathrm{d}s$. It follows that $B(r) = \sum_{k=1}^{\infty} \varphi_k(1) \eta_k(r)$, where $B(r)$ and $\eta_k(r)$ are both measurable with respect to the natural filtration $\mathcal{F}_r$ of the Brownian motion B. The process $\eta_k(r)$ describes the evolution over time of the coefficient of the coordinate basis φ_k. The evolution of these trend coordinates can be estimated by recursively regressing the sample data on the functions φ_k and the resulting estimates deliver direct information on how individual trend coordinates have evolved over time. They may, for instance, be used in subsequent regressions (e.g., on relevant economic variables like R&D expenditures) to help explain changes in trend over time.

We may also use these representations for forecasting. In particular, setting $s = r$ in (12), the optimal predictor of $B(r)$ given $\mathcal{F}_p$ and coordinates up to K is

$$E\{B(r)|\mathcal{F}_p; K\} = \sum_{k=1}^{K} \varphi_k(1) E\{\eta_k(r)|\mathcal{F}_p\}. \tag{13}$$

Upon evaluation, the conditional trend component appearing in (13) is

$$E\{\eta_k(r)|\mathcal{F}_p\} = \frac{1}{r} \int_0^p B(s) \varphi_k \left(\frac{s}{r} \right) \mathrm{d}s + B(p) \frac{\sqrt{2} \cos\{(k - 1/2)\pi(p/r)\}}{(k - 1/2)\pi}. \tag{14}$$

Summing over $k = 1, \ldots, K$ we get

$$E\{B(r)|\mathcal{F}_p; K\} = \sum_{k=1}^{K} \varphi_k(1) \left\{ \frac{1}{r} \int_0^p B(s) \varphi_k \left(\frac{s}{r} \right) \mathrm{d}s + B(p) \frac{\sqrt{2} \cos\{(k - 1/2)\pi(p/r)\}}{(k - 1/2)\pi} \right\}. \tag{15}$$

If we summed to infinity in this expression, we would end up with the (martingale) model predictor $E\{B(r)|\mathcal{F}_p\} = B(p)$ corresponding to the long run behavior suggested by a unit root model. Other long run models and orthogonal representations would lead, of course, to different classes of long run predictors.

More generally, for every value of K we get from (15) a class of predictors corresponding to the leading K trend components in the process. Let $N = [np]$, and $N + h = [nr]$ so that (14) and (15) effectively provide h-step ahead optimal predictors of these components. $E\{\eta_k(r)|\mathcal{F}_p\}$ may be estimated from sample data by

$$\hat{\eta}_k(r|p) = \frac{1}{N + h} \sum_{t=1}^{N} X_t \varphi_k \left(\frac{t}{N + h} \right) + X_N \frac{\sqrt{2} \cos\{(k - 1/2)\pi(N/(N + h))\}}{(k - 1/2)\pi},$$

which leads to the following *h*-step ahead predictor of the trend in the data

$$\hat{X}_{N+h,N} = \sum_{k=1}^{K} \varphi_k(1) \left\{ \frac{1}{N+h} \sum_{t=1}^{N} X_t \varphi_k \left(\frac{t}{N+h} \right) + X_N \frac{\sqrt{2} \cos\{(k-1/2)\pi(N/(N+h))\}}{(k-1/2)\pi} \right\}. \quad (16)$$

This approach can be pursued further to construct formulae for trend components and trend predictors corresponding to a variety of long run models for the data, such as those appearing in (2) and (3). Such formulae enable trend analysis and prediction in a way that captures the main features of the trend for *K* small and which can be related back to specific long term predictive models for large *K*. The approach therefore helps to provide a foundation for studying trends in a general way, covering most of the trend models that are presently used with economic data.

6. Simulation evidence

As a brief illustration we simulated the following cointegrated model:

$$X_{1t} = bX_{2t} + u_{1t}, \qquad X_{2t} = X_{2t-1} + u_{2t}, \qquad \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim \text{iid } N(0, \Omega), \quad (17)$$

with

$$\Omega = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

The cointegrating coefficient *b* was estimated using ordinary least squares (OLS), fully modified least squares (FM-OLS), vector autoregression reduced rank regression (VAR-RRR) [9], and coordinate reduced rank regression as in (10) above. Kernel estimates of the probability densities of these estimates are shown in Fig. 8 for the case where $\rho = 0.8$, $b = 2.0$, $n = 100$, and $K = 75$. The VAR-RRR was estimated with a single lag, consonant with (17).

The results show that coordinate RRR performs well. It is less biased than OLS and has greater concentration than FM-OLS. But, it is more biased and a little less concentrated than VAR-RRR. Of
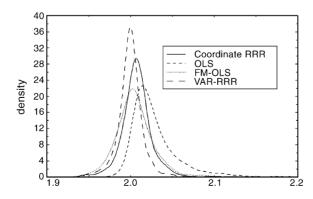

Fig. 8. Densities of cointegrating coefficient estimators: true coefficient $b = 2$, $n = 100$, $K = 75$, $\rho = 0.8$.
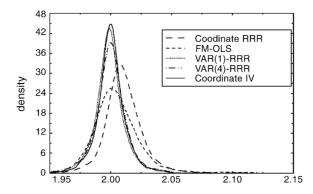
Fig. 9. Densities of cointegrating coefficient estimators.

course, VAR-RRR is performed with the correct specification corresponding to (17), including a single lag, and therefore has some obvious advantages in this case, especially over FM-OLS (which is designed to allow for general error specifications).

As a final illustration, Fig. 9 shows how these procedures perform against the coordinate IV procedure discussed earlier using the same parameter configuration for data generation. The results indicate that coordinate IV seems to have better finite sample performance characteristics than VAR-based reduced rank regression (particularly when the order of the VAR is overspecified), FM-OLS and coordinate RRR methods.

7. Conclusion

The probabilistic foundation of econometrics conventionally presumes that the observed process can be faithfully represented in terms of a probability space with quantifiable economic variables defined as random elements on that space. This conceptualization has proved to be a useful approach to formal modeling and is so much part of our conventional wisdom that it is very easy to accept that there must be an underlying true data generating process on the defined space. However, the actual process of data generation may not fit faithfully into this framework without an extraordinary level of complexity that belies the notion of modeling as we presently know it. This view may initially appear heretical but it becomes reasonable upon serious reflection.

When the data involve trends, as most macroeconomic time series do, even a little empirical experience is sufficient to show the inadequacy of commonly used trend formulations. The alternative perspective suggested here is that, while we may not understand the trending mechanism itself, we still have the opportunity to coordinatize a trend in terms of simple deterministic functions, just as we can coordinatize a function in a space of functions using a simple set of basis functions. As we have shown, this coordinatization provides rather a general framework for thinking about trending time series, one that is not restricted to a particular class of time series generating mechanisms. This framework gives rise to a new concept of coordinate co-movement which can be used to study patterns of common behavior in time series. Rather remarkably, while empirical regression estimates of such coordinate representations have in the past been considered 'spurious', econometric estimates of these coordinate systems can be used to produce consistent estimates of the cointegration space when there is co-movement in the data. The

empirical and simulation evidence given here indicates that the approach holds some promise in practical applications and at least provides a new way of looking at trending data.

8. Data

The producer price series has two main sources. It is constructed from a historical Wholesale Price Index (WPI) series used in Shiller and Siegel [26] and a Producer Price Index (PPI) series that was obtained from the UK Government Statistical Service on the website: http://www.statistics.gov.uk. The Shiller and Siegel price series covers the period 1718–1973 and is a Wholesale Price Index for the UK constructed by splicing several other constituent series, as explained in the appendix of their article. The PPI series is the (annual) producer price index for output prices (Series Identifier: PPLU all manufacturing 1974–2002), downloaded from http://www.statistics.gov.uk. The PPI data was spliced to the WPI series by first splicing the WPI to the retail prices index (also obtained from http://www.statistics.gov.uk) for 1974 using the common year 1973 for those series and then splicing the PPI series to this series by multiplying the PPI series by the ratio of the two series for the overlapping year 1974.

The yield series also comes from two sources. It is constructed from the yield series used in Shiller and Siegel [26] and the yield on 3.5% War Loan securities obtained from the Bank of England (http://www.bankofengland.co.uk). The Shiller and Siegel series covers the period 1718–1973 and is based on series compiled by Homer [8] representing the yield on perpetual annuities and various consols, as explained in the appendix of their article. This series was spliced to a series for the yield on 3.5% War Loan securities – Series WRLN in Table 22.4 of the Bank of England Monetary and Financial Statistics site (http://www.bankofengland.co.uk/mfsd/abst/part1.htm) – by multiplying the series with the geometric mean of the ratios of the rates for the overlapping years 1970–1973.

Acknowledgments

References

[1] J. Bai, A note on spurious breaks, Economet. Theory 14 (1998) 663–669.
[2] M. Baxter, R.G. King, Measuring business cycles: approximate band-pass filters for economic time series, Rev. Econ. Stat. 81 (1999) 575–593.
[3] D. Corbae, S. Ouliaris, P.C.B. Phillips, Band spectral regression with trending data, Econometrica 70 (2002) 1067–1110.
[4] S.N. Durlauf, P.C.P. Phillips, Trends versus random walks in time series analysis, Econometrica 56 (1988) 1333–1354.
[5] I. Fisher, The Theory of Interest, Macmillan, New York, 1930.
[6] C.W.J. Granger, P. Newbold, Spurious regressions in econometrics, J. Economet. 74 (1974) 111–120.
[7] R.J. Hodrick, E.C. Prescott, Postwar US business cycles. An empirical investigation, J. Money Credit Banking 29 (1997) 1–16.
[8] S.A. Homer, History of Interest Rates, Rutgers University Press, New Jersey, 1996.
[9] S. Johansen, Statistical analysis of cointegration vectors, J. Econ. Dyn. Contr. 12 (1988) 231–254.

[10] J.M. Keynes, A Treatise on Money, Macmillan, New York, 1930.
[11] F. Marmol, Spurious regression between $I(d)$ processes, J. Time Series Anal. 16 (1995) 313–321.
[12] F. Marmol, Nonsense regressions between integrated processes of different orders, Oxford Bull. Econ. Stat. 58 (1996) 525–536.
[13] F. Marmol, Spurious regression theory with nonstationary fractionally integrated processes, J. Economet. 84 (1998) 233–250.
[14] L.C. Nunes, C.M. Kuan, P. Newbold, Spurious break, Economet. Theory 11 (1995) 736–749.
[15] P.C.B. Phillips, Understanding spurious regressions in econometrics, J. Economet. 33 (1986) 311–340.
[16] P.C.B. Phillips, Optimal inference in cointegrated systems, Econometrica 59 (1991) 283–306.
[17] P.C.B. Phillips, New tools for understanding spurious regressions, Econometrica 66 (1998) 1299–1326.
[18] P.C.B. Phillips, Discrete Fourier transforms of fractional processes, Cowles Foundation Discussion Paper No. 1243, Yale University, 1999.
[19] P.C.B. Phillips, New unit root asymptotics in the presence of deterministic trends, J. Economet. 11 (2001) 323–353.
[20] P.C.B. Phillips, Laws and limits of econometrics, Econ. J. 113 (2003) C26–C52.
[21] P.C.B. Phillips, B.E. Hansen, Statistical inference in instrumental variables regression with $I(1)$ processes, Rev. Econ. Stud. 57 (1990) 99–125.
[22] P.C.B. Phillips, S. Ouliaris, Asymptotic properties of residual based tests for cointegration, Econometrica 58 (1990) 165–193.
[23] P.C.B. Phillips, K. Shimotsu, Local Whittle estimation in nonstationary and unit root cases, Ann. Stat. 32 (2004) 656–692.
[24] W. Ploberger, P.C.B. Phillips, Empirical limits for time series econometric models, Econometrica 71 (2003) 627–673.
[25] I.J. Schoenberg, Spline functions and the problem of graduation, Proc. Natl. Acad. Sci. USA 52 (1964) 333–343.
[26] R.J. Shiller, J.J. Spiegel, The Gibson paradox and historical movements in real interest rates, J. Polit. Econ. 85 (1977) 891–907.
[27] K. Shimotsu, P.C.B. Phillips, Exact local Whittle estimation of fractional integration, Cowles Foundation Discussion Paper # 1367, 2002.
[28] W.-J. Tsay, C.F. Chung, The spurious regression of fractionally integrated processes, J. Economet. 96 (2000) 155–182.
[29] G. Wahba, Improper priors, spline smoothing, and the problem of guarding against model errors in regression, J. Roy. Stat. Soc., Ser. B 40 (1978) 364–372.
[30] E.T. Whittaker, On a new method of graduation, Proc. Edinburgh Math. Assoc. 78 (1923) 81–89.
[31] G.U. Yule, Why do we sometimes get nonsense-correlations between time series?—A study in sampling and the nature of time series, J. R. Stat. Soc. 89 (1926) 1–69.