Special Feature C

The Mysteries of Trend[1]

by Peter C. B. Phillips[2]

The Hamlet of Econometrics

"A statistician is a fellow that draws a line through a set of points based on unwarranted assumptions with a foregone conclusion."

"No one understands trends. Everyone sees them in data."

Trends are ubiquitous in economic discourse, they figure prominently in media commentary, they play a role in much economic theory, and they have been intensively studied in econometrics for three decades. Yet the empirical economist, forecaster, and policy maker have little guidance from theory about the source and nature of trend behaviour. They have even less guidance about practical formulations, and they are heavily reliant on a limited class of stochastic trend, deterministic drift, and structural break models to use in applications. A vast econometric literature has emerged but the nature of trend remains elusive.

In spite of being the dominant characteristic in much economic data, having a role in policy assessment that is often vital, and attracting intense academic and popular interest that extends well beyond the subject of economics, trends are little understood. Like the protagonist in Shakespeare's most famous play, trend remains unfathomable and inscrutable, the Hamlet of econometrics. No one knows what it will do next.

This Special Feature discusses some implications of these limitations, mentions some research opportunities, and briefly illustrates the extent of the difficulties in learning about trend phenomena even when the time series are far longer than those that are available in economics.

What is Trend?

Trend is a simple five letter word. Its use is ubiquitous in economics, dominating macroeconomic discourse on growth and productivity which, as Paul Krugman[3] once said, in the long run affect almost everything in economics. The concept is equally pervasive in modern microeconomics and all the applied subfields of economics, where intertemporal comparisons play a major role in economic theories of behaviour and in subsequent assessments of policy effectiveness, covering issues as sociologically diverse as the impact of abortion rights legislation on crime, schooling on earnings, and no-fault legislation on divorce statistics. In the world of finance, trend is just as vital and important because it is the drift in asset prices that provides the allure of long-term capital appreciation and rewards risky investment.

---

[1] A longer version of this Special Feature is available from the author's website (http://korora.econ.yale.edu). Partial research support is acknowledged from the NSF under Grant No. SES 09-56687.

[2] Peter C. B. Phillips is Sterling Professor of Economics and Professor of Statistics at Yale University and Distinguished Term Professor at Singapore Management University.

[3] "Productivity isn't everything, but in the long run it is almost everything" is the opening line in Krugman (1995).

The dictionary definition of the word trend originates from a nineteenth century usage[4] as "the general course of events or prevailing tendency" – a seemingly simple concept that is readily apprehended by all. Or is it? Is our apprehension of the concept so unambiguous that it needs no explicit definition beyond that of our general understanding derived from its dictionary meaning? Media commentaries as well as professional economic discussion frequently take the meaning of the term for granted and proceed to lever policy argument on the basis of this presumption. How often, for instance, do we hear senior public economists like the Chairman of the Federal Reserve, Governors of central banks and Treasury Secretaries describing the context for economic policy decisions by speaking of the data in expressions such as "if current trends continue ..." or "a newly-emergent trend is ..." or "long-term trends indicate ...".

It is one of the ironies of economics that while these commonly-used phrases appear to carry a measure of technical precision that lends professional import to discussion, that precision (and presumably some of the credibility that comes with it) is illusory. Leaving aside the issue of what is really meant or intended by the word "trend" (whose popular meaning has changed significantly over the last few centuries and whose scientific meaning is seldom given), the attendant epithets (such as current, emergent, or long-term in the usages cited) seem to lend precision to the concept, thereby creating a misleading impression of scientific import in their usage. Misleading, because it is impossible to have clarity in these expressions without making the component terms themselves unambiguous: what is a trend, and how are the terms current, newly emergent, and long term to be interpreted? In short, is it possible to measure and discuss with clarity any quantity that is undefined? Econometricians have been battling with these ideas over the last 30 years and know how imprecise the terms are. How is it then that such a fundamental concept as trend, whose use is so widespread in the elite quantitative journals and public economic forums, can be so imprecise in discourse, so little understood and so often misleading in practice?

The ubiquity of the word trend and its imprecision are by no means confined to economic discourse. Imprecisions in usage arise everywhere across the social, behavioural and business sciences to the natural sciences and from popular discussion in the media to scientific work. In some cases, as in the assessment of climate change (to which we turn below), trend measurement has major societal and planetary consequences, as well as economic and policy implications.

One explanation for the ubiquitous usage lies in a natural human desire to bring order to disorder when seeking to understand (or model) the world around us. When we see a cloud of data points plotted against time, our minds bring order to that disorder by drawing a line through the points – representing the data in a way that seeks to satisfy an innate need to understand its primary features. We want to know what has been, where we are now, and most of all, where we are going. A trend line satisfies these primitive requirements. It summarises where we have been, shows where we are now in relation to the past, and, most of all, reveals a hint of where we are going. The lines we draw in our minds, like those we draw on paper or fit by econometric methods, are typically smooth and the derivative is a direction vector for the future. Lines through the data reveal features like a long-run tendency to increase over time, a cyclical pattern, or turning points that can be associated with known events, thereby helping to reinforce their value to us. Parametric and nonparametric trend regression and smoothing techniques like Whittaker (1923) graduation (known in macroeconomics as Hodrick-Prescott filtering) are simply technical mechanisms that formalise this mental process of representation and *ex-post* discovery.

[4] According to the Online Etymology Dictionary (http://www.etymonline.com/) this usage in the sense of "general tendency" is recent and dates from 1884. The older meaning of the word (a verb) dates from 1598 – "to run or bend in a certain direction, as of a river or coastline" – and is based on the Middle English "trenden" which meant "to roll about, turn, or revolve" – certainly a different meaning from "the general course of events" as we presently understand the term. Given this etymology, the modern notion of a stochastic trend seems to possess an atavistic link with the earlier usage.

Whether the device is the eye, the hand, or the technical apparatus of econometrics, trend fitting leads to a curve through a set of points that is typically continuous and smooth, or at least piecewise so. These properties facilitate the exercise and they offer advantages in potential interpretation, suggesting the existence of a generating mechanism for which continuous differentiability is a basic feature, subject perhaps to an occasional structural shift. Such a "trend" is manufactured from the data and easily apprehended. But how realistic is such a heroically simplified representation of a mechanism that by its very nature resists understanding, when even the vocabulary of description defies scientific clarity? For when we speak of current trends continuing, do we simply assert that a line drawn through a given set of points continues into the future? If so, which line or curve and which set of current points? Do we mean the last three data points, the last five or the last ten? And how well does the proposition that emerges withstand these changes in formulation?

Technical market analysis, which is so common in the popular financial press, abounds with such lines, giving readers a visual directory of upper and lower trend support lines, long-term containment triangles, resistance levels, and many other data-manufactured lines, all purporting to represent some fundamental feature of a series and its evolution. As the definition of a statistician that heads this Special Feature implies, much data analysis of trending time series is of this kind, often resting on unstated and unwarranted assumptions that are not tested. How then are we to value and interpret such analysis? And what better alternatives do formal econometric methods offer the empirical researcher and policy maker whose decisions often rely on trend evaluation in relation to alternate policies?

A partial answer to these questions has been provided by the econometrics of stochastic trends, structural breaks, and nonstationary time series which has produced toolrooms of new methodology for analysing trends. This machinery allows practitioners to cope with trend processes that are inherently random or subject to random shifts, as well as many practical trend models that are misspecified.

Stochastic Trend

To the wide professional community of applied economists working in macroeconomics and international finance, the most influential and practically useful transformation in the last three decades in econometrics has been the unit root and cointegration revolution. This revolution changed the way the profession thought about trend by emphasizing the role of stochastic elements in the trend mechanism and by formulating a technically well-defined concept of long-run behaviour that did not remove randomness. In the mid 1980s, functional limit laws and integral functionals of Brownian motion took time series econometrics in a firestorm that swept through all the mainline economics journals.

The new thinking swiftly penetrated econometric teaching and empirical practice, creating a vast new literature of applied economics sophisticated in its use of modern econometric technology and nonstandard limit theory.

Beyond economics, the methods became a major export of econometrics to other social and business sciences. Their rapid acceptance and widespread use across many disciplines affirmed the importance of an idea whose time had come – a random trending mechanism that could be used to study commonality in movement over time among many series and deliver estimates of long-run linkages and adjustments, as well as transient dynamics.

In their limiting forms – Brownian motion, fractional motions, diffusions, and semi-martingales – these trends form continuous stochastic processes but they are not smooth and they have inherently unpredictable elements. Change and randomness form a critical element in their composition. In this respect they differ from the trend lines that our minds draw when we are confronted with a cloud of points. Correspondingly, when econometric time trend regressions or smoothing algorithms draw lines through data manifesting stochastic trends, we obtain spurious regressions which give a misleading view of the nature of the trend and its direction. The econometric methods developed in studying these phenomena enabled us to explain precisely what conventional trend line regressions do deliver in the context of such misspecification (Phillips, 1986; Durlauf and Phillips, 1988) and how comovement may be efficiently estimated (Johansen, 1988; Phillips and Hansen, 1990; Phillips, 1991). These methods opened the door to a new arena of empirical research and policy discussion that has been enormously productive and has created a new standard of professional econometric practice and empirical policy analysis.

In this new standard, spurious regressions have a well defined pejorative meaning, usually taken in contrast to cointegrating regression. But cointegrating regressions do not model or explain trends, they simply co-relate trending time series under given assumptions about the form of the trending mechanism. These assumptions are necessary for many econometric methods but they are inevitably approximations in view of the complex and poorly understood nature of the forces that determine trends in the data. The result is inescapable – trend misspecification and some degree of spurious regression.

Spurious modelling of trends may be inevitable but it is far from useless. If it were, then there would be little value in much applied macroeconomic work, where trend misspecification must be taken as universal. Here (and elsewhere in applied work) convenience is frequently a decisive factor heightening the appeal of devices like polynomial time trend regression and simple smoothing operations such as the Whittaker-Hodrick-Prescott filter. Like least squares regression, these methods still form the backbone of much empirical work and they do not yield their ground easily to more sophisticated alternatives such as various forms of nonparametric fitting using both time and frequency approaches (e.g., Corbae, Ouliaris and Phillips, 2003; Shimotsu and Phillips, 2005).

Nor do more sophisticated methods necessarily address the root issue of misspecification. But nonparametric approaches in the frequency domain can be helpful in that they distinguish the memory component in the data as an important individual feature and they permit general formulations of trending processes in terms of the asymptote of the spectral density in the immediate locality of the zero frequency. These asymptotic forms hold for many different classes of trend, both deterministic and random. So they appeal in terms of their generality. Correspondingly, general representations hold for the discrete Fourier transform of the time series in the region of the zero frequency and therefore furnish sample information about the nature and strength of the trend.

Recent research (Phillips, 1998; 2005) I have been pursuing has shown that trend misspecification need not be fatal, even when using smooth polynomials to model stochastic trends. In such cases, the regression coefficients remain random even in infinite samples and may be interpreted as the random coefficients that arise in projecting the limiting stochastic trend process on subspaces furnished by basis functions, such as the time polynomials used in regression. Similar properties hold in the case of breaking trend basis functions. We can, in fact, think of these models as coordinate approximations to an always more complex (and random) underlying trend function. In effect, the time polynomials or other regressors act as basis functions forming a sieve space (an approximating space using an infinite family of functions) for a stochastic process. The random coefficients then reflect the randomness in the trend process itself. It is also possible to use these coordinate regression functions in a meaningful way for prediction – in the limit, these predictors can even reproduce martingale like forms, as shown in Phillips (2005). In effect, smooth deterministic functions can represent nondifferentiable (unpredictable) martingales in the limit when we allow for random coefficients. The coordinate basis approach may also be used to model and capture co-movement among such time series in a very general way, extending the notion of reduced rank regression that is now commonly used in applied econometric modelling to a stochastic process context.

In practice, therefore, while economists and financial analysts frequently see trends in the data and wish to use estimates of these trends in policy projections, the econometric modelling of such trends is demanding and failure can have major implications for policy. When the trend-generating mechanism is poorly captured in an empirical model, forecasts carry forward the poor approximation. The phenomenon is familiar to empirical researchers and forecasters who see the incoming data drift away from their model projections as the horizon increases. Quick model adaptation to the random wandering, unpredictable element of trend (witness the original medieval meaning of the word) then becomes a critical feature in good applied modelling and needs to be accounted for in forecasting and policy analysis, as many experienced practitioners acknowledge.

Econometric analysis of model adaptation mechanisms to capture changes and account for shifts in location and trend soon after they occur are becoming part of a new armoury for forecasters (Phillips and Ploberger, 1994; Andrews, 2003; Clements and Hendry, 2006; Castle *et al*, 2010). Recent analysis by Ploberger and Phillips (2003) provides a limit theory which explains how much harder it is to get closer to a true generating mechanism with nonstationary components than it is one with only stationary covariates. A corollary of this theory is that forecasting is an order of magnitude harder for trending data because the optimal forecast is harder to estimate even when the form of true trend model is known.

The moral is that if trend terms are present in our models we need to be sure that they are relevant, well estimated, and quickly adapted to change. Otherwise, they can be powerfully wrong in forecasting and mislead policy. As the second header to this Special Feature intimates, trends have an elusive quality: no one understands the mechanism, but everyone sees evidence of it in the data.

Economic Policy and Climate Trend

National economic policies are commonly motivated by long-term goals and correspondingly reflect perceived trends in various indicators of societal needs. Similar considerations drive global policy agreements on financial stability, trade, and economic cooperation. Trend assessment is inevitably part of all such policy decisions. Nowhere is this more evident at present than in the ongoing global discussion of policy on climate change.

Underlying all discussion and policy enactment is the science of climate change – understanding the natural processes, external forces and human activity that may affect long-term climate. There is broad scientific agreement about human impact on the level of greenhouse gases (GHG) in the atmosphere, manifested in the popular "hockey stick" graphic that shows the trend in GHG over the last two centuries as a sharp spike against the blade of little change over the previous two millennia. There is also agreement, but less unanimity, about the quantitative impact of GHG emissions on climate. Evidence available from ice core data[5] over the past half million years confirms a strong and persistent association but the causal mechanism and time lags involved are complex and little understood.

Economic policy analysis has to assess the cost of doing nothing or too little about climate change against the cost and potential gains of implementing GHG abatement strategies like emissions trading and carbon taxation. Caught up in this policy debate are major questions of trend determination: how GHG emissions will affect climate over the next century and what impact on the trend the different abatement measures may have. Economic analysis, national economic policy and successful global cooperation all rely on estimates of climate trend. The horizons cover everything from a few years to generations in the future.[6] The difficulties and uncertainties involved in these trend projections are simply enormous.

For comparison, climatological data extend over geologic time frames and are measured in thousand year or million year units. Against this time frame, economic time series seems woefully short, especially when it comes to studying trend behaviour. Yet many of the same problems (such as the inherent random elements in trend, shortfalls in theory guidance, and ambiguities between trend and cycle) continue to manifest themselves. Having more data, in effect, does not always lead to improvement in analysis or understanding. Sometimes, especially with trending time series, the advent of more data simply means more to explain. As in economics, it is the synergy of good theory, data, and statistical methodology that is most likely to enhance understanding.

No present climatological (or planetary) simulation models are capable of generating climate trajectories of the type that have been observed over long geologic periods. Neither do the models or methods currently in use in studying trends in econometrics measure up to the task of modelling these series. Paleoclimate data over many millions of years raise the difficulties of trend modelling to an entirely different level. Trend is a complex phenomenon with features that turn out to be endogenous to the sample size. As we lengthen the time span of observation, what first appears as a pattern of drift later becomes absorbed into a cycle with a longer period or even manifests as volatility. The pattern continues to repeat itself over different time scales, extending back with presently available data as far as half a billion years.

Is trend itself then a phenomenon that is relative to time scale? If so, when we model trend how do we take account of the wider picture presented by a longer time frame when that data is not available to us? And what form of asymptotic theory is appropriate in a finite sample where the trend form is random and endogenous to the sample size? These are hard questions that push the limits of present understanding. In the absence of data, the answers must lie in good theory, better econometrics and fast algorithms for adapting models that are inevitably misspecified.

To capture the random forces of change that drive a trending process, we need sound theory, appropriate methods, and relevant data. In practice, we have to manage under shortcomings in all of them. It is at least some comfort for the econometrician and economic policy maker to know that these manifold difficulties of modelling trend are not confined to economics.

5 Petit *et al.* (1999) provide a record and statistical analysis of various GHG levels as well as temperature and dust particles obtained from ice core samples covering the past 420,000 years from the Vostok station in Antarctica.

6 Over such time frames even the choice of the discount factor can have major implications (Nordhaus, 2007).

Trends and Truth

Picasso once said that art is a lie that tells the truth. Even the most ardent proponent of the merits of economic theory could hardly claim the same of economic models. Good economic models are lies that may reveal a kernel of insight about reality. Recognition of this shortcoming is as important as apprehending the truth that no one understands trends. The role of econometrics is to find that kernel of insight in the data and put it to work to aid forecasting and policy. If we are fortunate, some of the mysteries of trend, including its inherent random nature, may be revealed in the process.

References

Andrews, D W K (2003), "End-of-Sample Instability Tests", *Econometrica*, Vol. 71(6), pp. 1661-1694.

Castle, J L, Fawcett, N W P and Hendry, D F (2010), "Forecasting with Equilibrium-Correction Models during Structural Breaks", *Journal of Econometrics,* Vol. 158(1), pp. 25-36.

Corbae, D, Ouliaris, S and Phillips, P C B (2002), "Band Spectral Regression with Trending Data", *Econometrica,* Vol. 70(3), pp. 1067-1110.

Clements, M P and Hendry, D F (2006), "Forecasting with Breaks", pp. 605-657, in G. Elliott, C.W.J. Granger and A. Timmermann (eds), *Handbook of Economic Forecasting*, Elsevier.

Durlauf, S N and Phillips, P C B (1988), "Trends versus Random Walks in Time Series Analysis", *Econometrica*, Vol. 56(6), pp. 1333-1354.

Johansen, S (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, Vol. 12(2-3), pp. 231-254.

Krugman, P (1995), *The Age of Diminished Expectations,* The MIT Press, Cambridge, Massachusetts.

Nordhaus, W (2007), "Critical Assumptions in the Stern Review on Climate Change", *Science*, Vol. 317, pp. 201-202.

Petit, J R, Jouzel, J, Raynaud, D, Barkov, N I, Barnola, J M, Basile, I, Bender, M, Chappellaz, J, Davis, J, Delaygue, G, Delmotte, M, Kotlyakov, V M, Legrand, M, Lipenkov, V, Lorius, C, Pépin, L, Ritz, C, Saltzman, E, Stievenard, M (1999), "Climate and Atmospheric History of the Past 420,000 years from the Vostok Ice Core, Antarctica", *Nature*, Vol. 399, pp. 429-436.

Phillips, P C B (1986), "Understanding Spurious Regressions in Econometrics", *Journal of Econometrics,* Vol. 33(3), pp. 311-340.

Phillips, P C B (1991), "Optimal Inference in Cointegrated Systems", *Econometrica,* Vol. 59(2), pp. 283-306.

Phillips, P C B (1998), “New Tools for Understanding Spurious Regressions”, *Econometrica*, Vol. 66(6), pp. 1299-1326.

Phillips, P C B (2003), “Laws and Limits of Econometrics”, *Economic Journal*, Vol. 113(486), pp. C26-C52.

Phillips, P C B (2005), “Challenges of Trending Time Series Econometrics”, *Mathematics and Computers in Simulation*, Vol. 68(5-6), pp. 401-416.

Phillips, P C B and Hansen, B E (1990), “Statistical Inference in Instrumental Variables Regression with I(1) Processes”, *Review of Economic Studies*, Vol. 57(1), pp. 99-125.

Phillips, P C B and Ploberger, W (1994), “Posterior Odds Testing for a Unit Root with Data–Based Model Selection”, *Econometric Theory,* Vol. 10(3-4), pp. 774-808.

Ploberger, W and Phillips, P C B (2003), “Empirical Limits for Time Series Econometric Models”, *Econometrica*, Vol. 71(2), pp. 627-673.

Shimotsu, K and Phillips, P C B (2005), “Exact Local Whittle Estimation of Fractional Integration”, *Annals of Statistics,* Vol. 33(4), pp. 1890-1933.

Whittaker, E T (1923), “On a new Method of Graduation”, *Proceedings of the Edinburgh Mathematical Association*, Vol. 78, pp. 81-89.