OXFORD
MARTIN
SCHOOL

UNIVERSITY OF
OXFORD

All Change!

The Implications of
Non-stationarity for
Empirical Modelling,
Forecasting and Policy

DAVID F. HENDRY AND FELIX PRETIS

The Oxford Martin School at the University of Oxford is a world-leading centre of pioneering research that addresses global challenges. We invest in research that cuts across disciplines to tackle a wide range of issues such as climate change, disease and inequality. We support novel, high risk and multidisciplinary projects that may not fit within conventional funding channels. We do this because breaking boundaries can produce results that could dramatically improve the wellbeing of this and future generations. We seek to make an impact by taking new approaches to global problems, through scientific and intellectual discovery, by developing policy recommendations and working with a wide range of stakeholders to translate them into action.

The Economic Modelling Programme (EMoD) aims to develop new methods of economic analysis and forecasting that are robust after crises. The 21st Century began with the largest global economic and financial crisis since the Great Depression 80 years ago. Many factors have been blamed for this disastrous outcome, but inadequate economic models leading to a failure to forecast the crash are partly at fault. These have been exacerbated by poor policy responses, precipitating the need for a paradigm shift. Consequently, researchers in EMoD are investigating the changes needed in economic analyses, policy, empirical modelling and forecasting when there are sudden, or very rapid, unanticipated changes in economies. EMoD researchers are also analysing the causes of economic inequality and the role of inequality in financial crises; and have played a key role in developing the World Top Incomes Database. EMoD is led by Professor Sir David Hendry (Director) and Professor Bent Nielsen (Co-Director) and is conducted in collaboration with the University of Oxford's Department of Economics and Nuffield College. EMoD is part of the Institute for New Economic Thinking at the Oxford Martin School (INET Oxford), a multidisciplinary research institute dedicated to applying leading-edge thinking from the social and physical sciences to global economic challenges.

The Climate Econometrics (CE) project concentrates on developing econometric methods to augment climate-economic research by helping disentangle complex relationships between human actions and climate responses and their associated economic effects, masked by non-stationarities in the form of stochastic trends and breaks. CE aims to: improve our understanding of the impact of humanity on climate and vice versa; improve our understanding of how modern econometrics can be used in climate-economic research; bring together researchers in the field of empirical climate-economic modelling; help create more accurate historical climate records; and reduce uncertainty in socio-economic scenarios for long-run predictions of the resulting climate damages. CE is directed by Professor Sir David Hendry and Dr Felix Pretis, and is conducted in collaboration between EMoD, the Environmental Defense Fund, Nuffield College, and the Department of Economics at the University of Oxford.

This paper reflects the views of the authors, and does not necessarily reflect the position of the Oxford Martin School or the University of Oxford. Any errors or omissions are those of the authors.

All Change!

The Implications of Non-stationarity for Empirical Modelling, Forecasting and Policy

DAVID F. HENDRY

FELIX PRETIS

Executive Summary

'I want a clean cup,' interrupted the Hatter: 'let's all move one place on.'
From Alice's Adventures in Wonderland, by Lewis Carroll (1865).

In an age of congested transport systems, everyone knows what it is like to be stationary: stuck motionless in a traffic jam; a train standing still at a station long after the due departure time; an aircraft sitting at the departure gate several hours delayed. The same word is used in a more technical sense in statistics: a stationary process is one where its mean and variance are constant over time.[1] As a corollary, a non-stationary process is one where the distribution of a variable does not stay the same at different points in time—the mean and/or variance may change for many reasons. Non-stationarity is like a statistical version of the changeover point in a relay race — as they all change, one team successfully transfers, while another drops the baton, and a third is reaching towards a future transfer with an unknown outcome.

A glance at most economic and related time series suffices to reveal the invalidity of the assumption of stationarity: economies evolve and change over time in both real and nominal terms, sometimes dramatically. Moreover, the historical track record of economic forecasting is littered with forecasts that went badly wrong, an outcome which should occur infrequently in a stationary process.

Many models used in empirical research, forecasting or for guiding policy have been predicated on treating observed data as stationary. But policy decisions, empirical research and forecasting must take non-stationarity into account if they are to deliver useful outcomes. We offer guidance for policymakers and researchers on identifying what forms of non-stationarity are prevalent, what hazards each form implies for empirical modelling and forecasting, and for any resulting policy decisions, and what tools are available to overcome such hazards. The behaviour of UK wages, prices and productivity over 1860–2014 illustrates our discussion.

Theories and models of human behaviour that do not account for non-stationarity will continually fail to explain outcomes. As one example, in a stationary world, the conditional expectation of an event tomorrow based on all the available information today will be the best predictor. But if the mean of the distribution can shift, the conditional expectation can be far from tomorrow's outcome. This will create a disequilibrium, and individuals who formed such expectations will need to adjust to their mistakes. In fact, the mathematical basis of much of 'modern' macroeconomics requires stationarity to be valid, and fails when distributions shift in unanticipated ways. As an analogy, continuing to use such mathematical tools in non-stationary worlds is akin to insisting on using Euclidian geometry to measure angles of triangles on a globe: then navigation can go seriously adrift.

Empirical modelling also faces important difficulties when time series are non-stationary. If two *unrelated* time series are non-stationary because they evolve by accumulating past shocks, their correlation will nevertheless appear to be significant about 70% of the time using a conventional 5% decision rule. Apocryphal examples during the Victorian era were the surprising high positive correlations between the numbers of births, and storks nesting, in Stockholm, and between murders and membership of the Church of England. As a consequence, these are called nonsense relations. This problem arises because uncertainty is seriously under-estimated if stationarity is wrongly assumed. During the 1980s, econometricians established solutions to this problem, and en-route also showed that the structure of economic behaviour virtually ensured that most economic data would be non-stationary.

Two common sources of non-stationarity affect the accuracy and precision of forecasts. The first source is the cumulation of past shocks, somewhat akin to changes in DNA cumulating over time to permanently change later characteristics. Evolution from cumulated shocks leads to far larger interval forecasts than would occur in stationary processes, so if a stationary model is incorrectly fitted, its calculated uncertainty will dramatically underestimate the true uncertainty. The second source is the occurrence of sudden unanticipated shifts in the level of a time series, called location shifts. These usually lead to forecast failure, where forecast errors are systematically much larger than would be expected in the absence of shifts, as happened during the Financial Crisis and Great Recession over 2008–2012. Consequently, the uncertainty of forecasts can be much greater than that calculated from past data, both because the sources of evolution in data cumulate over time, and also because 'unknown unknowns' can occur.

Scenarios based on outcomes produced by simulating empirical models are often used in economic policy, for example, by the Bank of England in deciding its interest-rate decisions. When the model is not a good representation of the non-stationarities prevalent in the economy, policy changes (such as interest-rate increases) can actually cause location shifts that lead to forecast failure, so after the event, what had seemed a good decision is seen to be badly based.

Thus, all four arenas of theory, modelling, forecasting and policy face serious hazards from non-stationarity unless it is appropriately handled. Fortunately, in each setting action can be taken, albeit providing palliative, rather than complete, solutions. Concerning theory derivations, there is an urgent need to develop approaches that allow for economic agents always facing disequilibrium settings, and needing error-correction strategies after suffering unanticipated location shifts. Empirical modelling can detect and remove location shifts that have happened: for example, statistical tools for dealing with shifts enabled Statistics Norway to revise their economic

forecasts within two weeks of the shock induced by the Lehman Brothers bankruptcy. Modelling can also avoid the 'nonsense relation' problem by checking for genuine long-run connections between variables (called cointegration, the development of which led to a Nobel Prize), and embody feedbacks that correct previous mistakes. Forecasting devices can allow for the ever-growing uncertainty arising from cumulating shocks. There are also methods for helping to robustify forecasts against systematic failure after unanticipated location shifts. Tests have been formulated to check for policy changes having caused location shifts in the available data, and warn against the use of those models for making future policy decisions.

Finally, although non-stationary time series data are harder to model and forecast, there are some important benefits deriving from non-stationarity. Long-run relationships are hard to isolate with stationary data: since all connections between variables persist unchanged over time, it is difficult to determine genuine causal links. However, cumulated shocks help reveal what relationships stay together (i.e., cointegrate) for long time periods. This is even more true of location shifts, where only connected variables will move together after a shift (called co-breaking). Such shifts also alter the correlations between variables, facilitating more accurate estimates of empirical models. Strong trends and location shifts can also highlight genuine connections, such as cointegration, through a fog of measurement errors in data series. Lastly, past location shifts allow the tests noted in the previous paragraph to be implemented before a wrong policy is adopted. Thus, non-stationarity is indeed a statistical version of 'All Change!' that can reveal important links, and need not just entail problems, like your train terminating unexpectedly at the wrong station.

1. Introduction

'It seems very pretty,' she said when she had finished it, 'but it's rather hard to understand.'
Alice, from Through the Looking-Glass and What Alice Found There, by Lewis Carroll (1899).

Many empirical models used in research and to guide policy in areas as diverse as economics to climate change analyse data by methods that assume observations come from stationary processes. A stationary time series is one where the distributions of outcomes remains constant over time. This requires that both the mean and the variance stay the same across years. Such processes are essentially ahistorical as the 'historical' time at which events happened is not a crucial attribute of the time series: the world always 'looks the same'.

However, most 'real world' time series are not stationary in that the means and variances of outcomes change over time. Present levels of knowledge, living standards, average age of death etc., are not highly unlikely draws from their distributions in medieval times, but come from distributions with very different means and variances. For example, the average age of death in London in the 1860s was around 45, whereas today it is closer to 80—a huge change in the mean. Moreover, some individuals in the 1860s lived twice the average, namely into their 90s, whereas today, no one lives twice the average age, so the relative variance has also changed.

Two main sources of non-stationarity are often visible in time series: evolution and sudden shifts. The former reflects slower changes, such as knowledge accumulation and its embodiment in capital equipment, whereas the latter occurs from (e.g.) wars, major geological events, and policy regime changes. We will explain the basic sources of non-stationarity, describe empirical modelling methods which handle non-stationarities, and discuss their importance for policymakers. If economic data are non-stationary, that will 'infect' other variables influenced by economics (e.g., CO_2 emissions), and so spread like a pandemic to most socio-economic and related variables. Many theories, most empirical models of

time series, and all forecasts will go awry when these two forms of non-stationarity are not tackled. For example, a key feature of processes where the distributions of outcomes shift over time is that probabilities of events calculated in one time period need not apply in another: 'once in a hundred years' can become 'once a decade'.

Consequently, both forms of non-stationarity affect theory models, empirical modelling, forecasting, and any policy decisions based on forecasts. Policy decisions have to take such non-stationarity into account: as an obvious example, with increasing longevity, pension payments and life insurance contracts are affected.
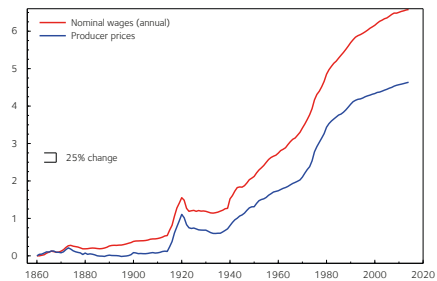

Figure 1: Indexes of UK wages and prices on a log scale.

Figure 1 graphs UK annual nominal wages and prices over the period 1860–2014. These have changed dramatically over the last 150 years, rising by more than **70,000%** and **1000%** respectively. Their rates of growth have also changed intermittently, as can be seen from the changing slopes of the graph lines. The magnitude of a 25% change is marked to clarify the scale. It is hard to imagine any 'revamping' of the statistical assumptions such that these outcomes could be construed as coming from stationary processes.[2] Figure 2 records real wages (in constant prices) with productivity, measured as output per person per year. Both trend strongly, but move closely together, albeit

with distinct slope changes and 'bumps' en route. The 'flat-lining' after the 'Great Recession' of 2008–2012 is highlighted by the ellipse. The wider 25% change marker highlights the reduced scale.
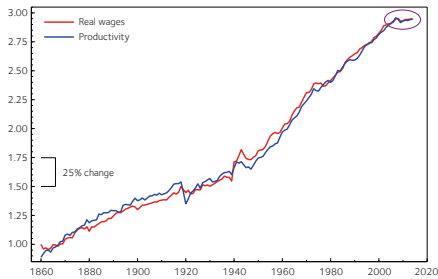

Figure 2: UK real wages and productivity.

Figure 3 (a) plots annual wage inflation (price inflation is similar) and (b) changes in real UK National Debt to emphasise that changes, or growth rates, also can be non-stationary, here from both major shifts in means (the black lines in (a)), as well as in variances. Compare the quiescent 50-year period before 1914 with the following 50 years, with scale markers of 5% in (a) and 20% in (b).
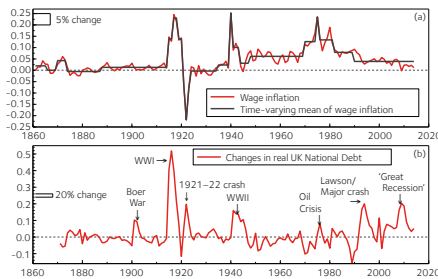

Figure 3: UK wage inflation and changes in real National Debt with major events shown.

None of these series has a constant mean or variance over time, so cannot be stationary. Two distinct features are exhibited, namely 'wandering' widely, most apparent in the first two figures, and suddenly shifting as in Figure 3, with the associated events shown in (b), features that will recur. Such phenomena are not limited to economic data. Figure 4 illustrates the non-stationary nature of climate time series, from cycles in the concentration of atmospheric CO_2 in pre-industrial times to the rapid annual increases in recent concentrations.
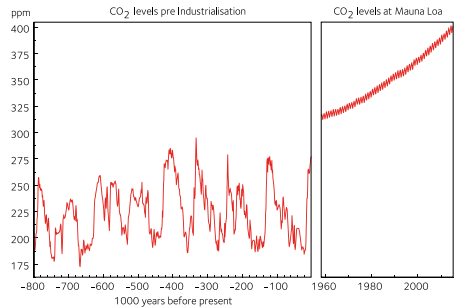

Figure 4: Levels of atmospheric CO_2.

Given the almost universal absence of stationarity in real-world time series, Hendry and Juselius (2000) delineated four issues with important consequences for research and policy:

- The central role of stationarity assumptions in empirical modelling and inference;
- The potentially hazardous impact on modelling, inference and policy of incorrectly assuming stationarity;
- The many sources of both main forms of non-stationarity (evolution and abrupt shifts);
- Yet fortunately, statistical analyses can often be transformed to eliminate most of the adverse effects of non-stationarity.

Section 2 discusses the properties that stationary time series should show in contrast to what is usually observed. Then Section 3 provides a historical review of our understanding of the first form of non-stationary data, called integrated processes (for reasons we will explain). Section 4 explains potential solutions to that form of non-stationarity, focusing on cointegration in section 4.1. Section 5 considers the second source of non-stationarity, namely location shifts, and Section 5.1 notes some solutions. Section 6 considers the impacts of non-stationarity on forecasting and Section 7 provides recommendations for forecasting when facing non-stationarity. Section 7.1 lists some potential benefits of non-stationarity, and Section 8 concludes.

2. Stationarity

'Now, here, you see, it takes all the running you can do, to keep in the same place.'
The Red Queen, from Lewis Carroll (1899).

Stationarity is the exception and non-stationarity is the norm for most time series. Anders Rahbek notes that looking for non-stationarity is analogous to going to the zoo and looking for non-elephants. The set of non-elephants is large and varied, as is the set of potential types of non-stationarity. To understand that, we begin by considering the special case of stationary time series.

A time series is (weakly) stationary when its first two moments, namely the mean and variance, are finite and constant over time. In a stationary process, the influence of past shocks must die out, as otherwise the variance could not be constant. Since past shocks do not accumulate (integrate), such a stationary time series is said to be integrated of order zero, denoted I(0). Observations on the process will centre around the mean, with a spread determined by the magnitude of its constant variance. Consequently, any sample of a stationary process will 'look like' any other, making it *ahistorical*. If an economy were stationary, we would not need to know the historical dates of the observations: whether it was 1860–1895 or 1960–1995 would be irrelevant. However, as seen in Figure 3b, specific events can matter greatly, including major wars, pandemics, and massive volcanic eruptions; financial innovation; key discoveries like vaccination, antibiotics and birth control; inventions like the steam engine and dynamo; etc. These can cause persistent shifts in the means and variances of the data, thereby violating stationarity. Figure 5 shows the large drop in UK birth rates following the introduction of oral contraception, and increasing longevity, with the declines in death rates since 1960.
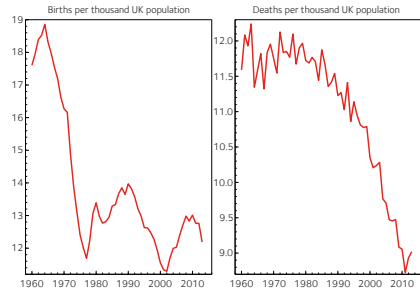

Figure 5: Births and deaths per thousand of the UK population.

3. Historical review

'The time has come,' the Walrus said, 'to talk of many things.'
From Lewis Carroll (1899).

Developing a viable analysis of non-stationarity in economics really commenced with the discovery of the problem of 'nonsense correlations'.[3] These are high correlations found between variables that should be unrelated: for example, that between the price level in the UK and cumulative annual rainfall shown in Hendry (1980). Yule (1897) had considered the possibility that both variables in a correlation calculation might be related to a third variable (e.g., population growth), inducing a spuriously high correlation. But by Yule (1926), he recognised the problem was indeed 'nonsense correlations'. He suspected that high correlations between successive values of each variable, called serial, or auto, correlation, might affect the correlation between the variables, and investigated that in a manual simulation experiment, randomly drawing from a hat pieces of paper with digits written on them. He then calculated correlations between pairs of draws for many samples of those numbers. He also calculated correlations between pairs after the numbers for each variable were cumulated once, and finally cumulated twice. For example, if the digits for the first variable went 5, 9, 1, 4, ..., the cumulative numbers would be 5, 14, 15, 19, ... and so on. Yule found that in the purely random case, the correlation coefficient was almost normally distributed around zero, but after the digits were cumulated once, was surprised to find the correlation coefficient was nearly uniformly distributed, so almost all correlation values were equally likely despite there being no genuine relation between the variables. Thus, he found 'significant', though not very high, correlations far more often than for non-cumulated samples. Yule was even more startled to discover that the correlation coefficient had a U-shaped distribution when the numbers were doubly cumulated, so the correct hypothesis of no relation between the genuinely unrelated variables was virtually always rejected due to a near-perfect, yet nonsense, correlation of ±1.

Granger and Newbold (1974) emphasised that an apparently 'significant relation' between variables, but where there remained substantial serial correlation in the residuals from that relation, was a symptom associated with nonsense regressions. Phillips (1986) provided a technical analysis of the sources and symptoms of nonsense regressions. Today, Yule's three types of time series are called integrated of order zero, one, and two respectively, usually denoted I(0), I(1), and I(2), as the number of times the series integrate (i.e., cumulate) past values. Conversely, differencing successive values of an I(1) series delivers an I(0) time series, etc.

4. Potential solutions to I(1) non-stationarity

'What am I to do?' exclaimed Alice.
From Lewis Carroll (1899).

Yule created integrated processes deliberately, but there are many economic, social and natural mechanisms that induce integratedness in data. Perhaps the best known example of an I(1) process is a random walk, first proposed by Bachelier (1900) to describe the behaviour of prices set in speculative markets, but such processes also occur in demography (see Lee and Carter, 1992), because the stock of a variable, like inventories or population, cumulates the net inflow (see Figure 5). A natural integrated process is the concentration of atmospheric CO_2, as emissions cumulate due to CO_2's long atmospheric lifetime as in Figure 4. Since the Industrial Revolution, such emissions have been mainly anthropogenic. When the inflows to an integrated process are random, because it cumulates past perturbations, the variance will grow over time, violating stationarity. Thus, unlike an I(0) process which varies around a constant mean with a constant variance, an I(1) process has an increasing variance, usually called a stochastic trend, and may also 'drift' in a general direction over time to induce an actual trend.

Figure 6: Twenty successive serial correlations for (a) wages, (b) real wages, (c) wage inflation and (d) real wage growth.

Cumulating many past random shocks should make the resulting time series relatively smooth. Since successive observations share a large number of past inputs, the correlation between them will be high, and only decline slowly as their distance apart increases. Figure 6 (a), (b) illustrates for the logs of wages and real wages, where the sequence of successive correlations shown is called a correlogram. Taking wages in the top left panel (a) as an example, the outcome in any year is still correlated 0.97 with the outcome 20 years previously, and similar high correlations between variables 20 years apart hold for real wages. Values above the green dashed line are significantly different from zero at 5%.

Differencing is the opposite of integration, so an I(1) process has first differences that are I(0). Thus, despite its non-stationarity, an I(1) process can be reduced to I(0) by differencing, an idea that underlies the approach in Box and Jenkins (1976). Now successive values in the correlogram should decline quite quickly, as Figure 6 (c) & (d) illustrates for the differences of these time series. Differences should also be approximately normally distributed when the shocks are nearly normal.

To summarise, both the mean and the variance of I(1) processes change over time, and successive values are highly interdependent. As Yule (1926) showed, this can lead to nonsense regression problems. Moreover, the conventional forms of distributions assumed for estimates of parameters in empirical models under stationarity no longer hold.
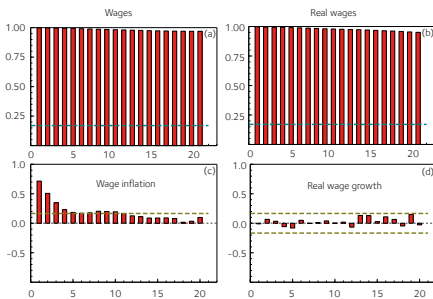
4.1 Cointegration between I(1) processes

Linear combinations of several I(1) processes are usually I(1) as well. However, stochastic trends can cancel between series to yield an

I(0) outcome. This is called cointegration. Cointegrated relationships define a 'long-run equilibrium trajectory', departures from which induce 'equilibrium correction' that move the relevant system back towards that path.[4] Real wages and productivity, shown in Figure 2, are each I(1), but their differential, which is the wage share shown in Figure 7, could be I(0). The wage share cancels the separate stochastic trends in real wages and productivity to create a possible cointegrating relation where the strong trends have been removed, but there also seem to be long swings and perhaps location shifts, an issue we consider in Section 5.
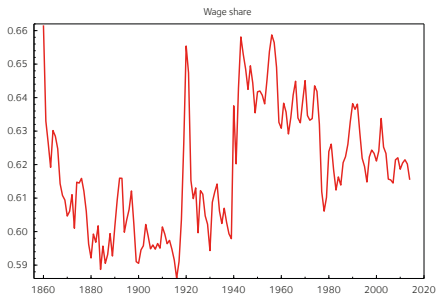

Figure 7: Time series for the wage share.

To illustrate pairs of variables that are (i) unrelated I(0) but autocorrelated, (ii) unrelated I(1), and (iii) cointegrated, Figure 8 shows 500 observations on computer-generated data. The very different behaviours are marked, and although rarely so obvious in practice, the close trajectories of real wages and productivity in Figure 2 over 150 years resembles the bottom panel, with the opposite trend.

In economics, integrated-cointegrated data seem almost inevitable because of the Granger (1981) Representation Theorem, for which he received the Sveriges Riksbank Prize in Economic Science in Memory of Alfred Nobel in 2003. His result shows that cointegration between variables must occur if there are fewer decision variables (e.g., your income and bank account balance) than the number of decisions (e.g., hundreds of shopping items: see Hendry, 2004, for an explanation). If that setting was the only source of non-

stationarity, there would be two ways of bringing an analysis involving integrated processes back to I(0): differencing to remove cumulative inputs (which always achieves that aim), or finding linear combinations that form cointegrating relations. There must always be fewer such relations than the total number of variables, as otherwise the system would be stationary.

Cointegration is not exclusive to economic time series. The radiative forcing of greenhouse gases and other variables affecting global climate cointegrate with surface temperatures, consistent with models from physics (see Kaufmann, Kauppi, Mann, and Stock, 2013, and Pretis, 2015). Thus, cointegration occurs naturally, and is consistent with many existing theories in the natural sciences where systems of differential equations in non-stationary time series can be written as a cointegrating model.

Other sources of non-stationarity also matter however, especially shifts in the means of data distributions of I(0) variables, including equilibrium corrections and growth rates, so we turn to this second main source of non-stationarity. There is a tendency in the econometrics literature to identify 'non-stationarity' with integrated data (unit roots), and so incorrectly claim that differencing a time series induces stationarity. There are many sources of non-stationarity, and for clarity we refer to the general case as wide-sense non-stationarity.
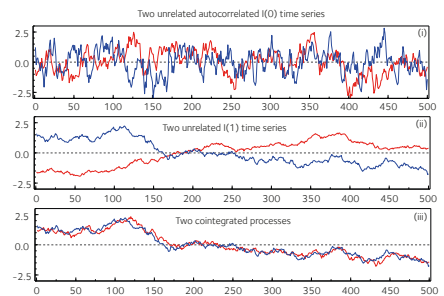

Figure 8: Pairs of artificial time series: (i) unrelated I(0); (ii) unrelated I(1); (iii) cointegrated.

5. Location shifts

'You're travelling the wrong way.'
The train guard to Alice, from Lewis Carroll (1899).

Location shifts are changes from the previous mean of an I(0) variable. There have been enormous historical changes since 1860 in hours of work, real incomes, disease prevalence, sanitation, infant mortality, and average age of death, among many other facets of life: see http://ourworldindata.org/ for comprehensive coverage. Figure 1 showed how greatly log wages and prices had increased over 1860–2014 with real wages rising seven-fold. Such huge increases could not have been envisaged in 1860. Uncertainty abounds, both in the real world and in our knowledge thereof. However, some events are so uncertain that probabilities of their happening cannot be sensibly assigned. We call such irreducible uncertainty 'extrinsic unpredictability', corresponding to unknown unknowns: see Hendry and Mizon (2014). A pernicious form of extrinsic unpredictability affecting modelling and forecasting is that of unanticipated location shifts, as these can occur at unanticipated times, changing by unexpected magnitudes and directions.

Figure 9 illustrates a hypothetical setting where the initial distribution is either a standard Normal in red with mean zero and variance unity, or a 'fat-tailed' distribution in blue, which has a high probability of generating 'outliers' at unknown times and of unknown magnitudes and signs (sometimes called 'black swan events' as in Taleb, 2007). As I(1) time series can be transformed back to I(0) by differencing or cointegration, the Normal distribution often remains the basis for calculating probabilities for statistical inference, as in random sampling from a known distribution. Hendry and Mizon (2014) call this 'intrinsic unpredictability', because the uncertainty in the outcome is intrinsic to the properties of the random variables. Large outliers provide examples of 'instance unpredictability' since their timings, magnitudes and signs are uncertain, even when they are expected to occur in general, as in speculative asset markets.
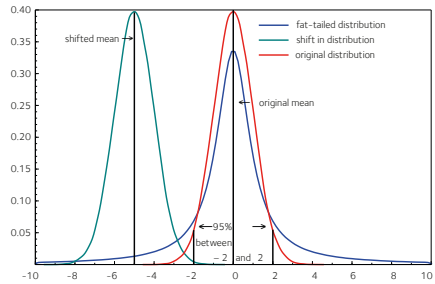

Figure 9: Location shifts in a Normal distribution.

However, in Figure 9 the baseline distribution experiences a location shift to a new Normal distribution in green with a mean of −5. There are many potential causes for such shifts, and many shifts have occurred historically, precipitated by changes in legislation, wars, financial innovation, science and technology, medical advances, climate change, social mores, evolving beliefs, and different political and economic regimes. Extrinsically unpredictable location shifts can make the new ordinary seem highly unusual relative to the past. In Figure 9, after the shift, outcomes will now usually lie between 3 and 7 standard deviations from the previous mean, generating an apparent 'flock' of black swans, which could never happen with independent sampling from the baseline distribution, even when fat-tails are assumed. During the Financial Crisis in 2008, the possibility of location shifts generating many extremely unlikely bad draws does not seem to have been included in risk models. But extrinsic unpredictability happens in the real world (see e.g., Soros, 2008): current outcomes are not highly discrepant draws from the distributions prevalent in the Middle Ages, but 'normal' draws from present distributions. Moreover, the distributions of many data differences are not stationary: for example, real growth per capita in the UK has increased intermittently since the Industrial Revolution, and most nominal differences have

experienced location shifts. Hendry (2015) provides dozens of examples.

5.1. Handling location shifts

At first sight, such shifts seem highly problematic for econometric modelling, but as with stochastic trends, there are several potential solutions. Differencing a time series will also inadvertently convert a location shift to an impulse (an impulse in the first difference is equivalent to a step-shift in the level). Secondly, time series can co-break, analogous to cointegration, in that location shifts cancel between series. Indeed, time series can be combined to remove some or all of the individual shifts. While individual series may exhibit multiple shifts, when modelling one series by another, co-breaking implies that fewer shifts will be detected when the series break together.
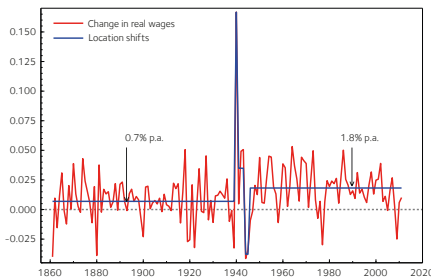

Figure 10: Partial co-breaking between wages and prices.

Figure 1 showed the divergent strong but changing trends in nominal wages and prices, and Figure 3(a) recorded the many shifts in wage inflation. Nevertheless, as shown by the time series of real wage growth in Figure 10, almost all the shifts in wage inflation and price inflation cancelled over 1860–2014. The huge 'spike' in 1940 was a key step in the UK's war effort, to encourage new workers to replace army recruits. The graph also reveals that the latter half of the 20th Century had substantively higher mean real-wage growth at 1.8% p.a. post-1945 versus 0.7% p.a. pre, and 1.3% overall. Real wages would have increased 16-fold at 1.8% p.a. from 1860, rather than just three-fold at 0.7%

p.a., and seven-fold in practice: 'small' changes in growth rates can dramatically alter living standards.

So how were the location shift lines drawn on Figures 3 and 10 chosen? This is the third possible solution: find all the location shifts and outliers whatever their magnitudes and signs and include them in the model. To do so requires us to solve the apparently impossible problem of selecting from more candidate variables in a model than observations. Hendry (1999) accidently stumbled over a solution. Most contributors to Magnus and Morgan (1999) had found that models of food demand were non-constant over the sample 1929–1952, so dropped that earlier data. To investigate why, yet replicate their models, Hendry added impulse indicators (which are zero everywhere except for unity at one data point) for all observations pre-1952, which revealed three large outliers corresponding to a US Great Depression food programme and post-war de-rationing. To check that his model was constant from 1953 onwards, he later added impulse indicators for that period, thereby including more variables plus indicators than observations, but only entered in his model in two large blocks, each much smaller than the number of observations. This has led to a statistical theory for modelling multiple outliers and location shifts (see e.g., Johansen and Nielsen, 2009, and Castle, Doornik, Hendry, and Pretis, 2015), available in our computational tool *Autometrics* (Doornik, 2009) as well as in the package *gets* (Pretis, Reade, and Sucarrat, 2016) in the statistical software environment *R*. This approach, called indicator saturation, considers a possible shift at every point in time, but only retains significant indicators.

Location shifts are of particular importance in policy as a policy change inevitably creates a location shift in the system of which it is a part. Consequently, a necessary condition for the policy to have its intended effect is that the parameters in its empirical models of the target variables must remain invariant to that policy shift. Thus, prior to implementing a policy, invariance should be tested, and that can be done automatically as described in Hendry and Santos (2010).

6. Forecasting non-stationary data

Alice began to remember that she was a Pawn, and that it would soon be time for her to move.
From Lewis Carroll (1899).

While empirical modelling is primarily concerned with understanding the interaction between variables to recover the underlying 'truth', the aim of forecasts is to generate predictions about the future regardless of the underlying structure of any forecasting model used. When forecasting in a stationary world, many famous theorems about how to forecast optimally can be rigorously proved (summarised in Clements and Hendry, 1998):

1. Causal models will outperform non-causal (i.e., models without any relevant variables);
2. The conditional expectation of the future value delivers the minimum mean-square forecast error (MSFE);
3. Mis-specified models have higher forecast-error variances than correctly specified ones;
4. Long-run interval forecasts are bounded above by the unconditional variance of the process;
5. Neither parameter estimation uncertainty nor high correlations between variables greatly increase forecast-error variances.

Unfortunately, when the variable to be forecast suffers from location shifts and stochastic trends, and the forecasting model is mis-specified, then:

1. Non-causal models can outperform correct in-sample causal relationships;
2. Conditional expectations of future values can be badly biased if later outcomes are drawn from different distributions (see Figure 9);
3. The correct in-sample model need not outperform in forecasting, and can be worse than the average of several devices;
4. Long-run interval forecasts are unbounded;
5. Parameter estimation uncertainty can substantively increase interval forecasts; as can
6. Changes in correlations between variables at or near the forecast origin.

The problem for empirical econometrics is not a plethora of excellent forecasting models from which to choose, but to find any relationships that survive long enough to be useful: as we have emphasised, the stationarity assumption must be jettisoned for observable variables in economics. Location shifts and stochastic trend non-stationarities can have pernicious impacts on forecast accuracy and its measurement.
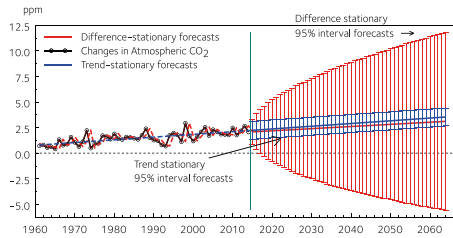

Figure 11: Forecasts of year-on-year changes in atmospheric CO_2 concentrations using a non-stationary stochastic-trend model (red) and a trend-stationary model (blue) with their associated 95% interval forecasts.

Because I(1) processes cumulate shocks, even using the correct in-sample model leads to much higher forecast uncertainty than would be anticipated on I(0) data. This is exemplified in Figure 11 showing forecasts of the year-on-year change in atmospheric CO_2 concentrations. Constant-change, or difference stationary, forecasts (in red) and deterministic trend forecasts (in blue) usually make closely similar central forecasts (solid lines) as can be seen here. But deterministic linear trends do not cumulate shocks, so irrespective of the data properties, and hence even when the data are actually I(1), their uncertainty is measured as if the data were stationary around the trend. Although the data properties are the same for the two models in Figure 11, their estimated forecast

uncertainties differ dramatically (red and blue bars), increasingly so as the horizon increases, due to the linear trend model assuming stable changes over time. Thus, model choice has key implications for measuring forecast uncertainty, where mis-specifications such as incorrectly imposing linear trends, can lead to understating the actual uncertainty in forecasts. As noted earlier, the assumption of a constant linear trend is rarely satisfactory. Caution is also advisable when forecasting integrated time series for long periods into the future, especially from comparatively short samples.
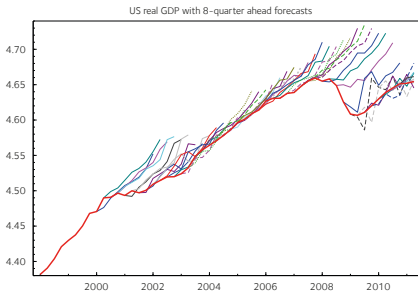
US real GDP with 8-quarter ahead forecasts

Figure 12: US real GDP with many successive 8-quarter ahead forecasts.

Next, forecasting in the presence of location shifts can induce systematic forecast failure (a significant deterioration in forecast performance relative to the anticipated outcome), unless the models used for forecasting account for the shifts. Figure 12 shows recent failures for a number of 8-quarter ahead forecasts of US log real GDP. There are huge forecast errors (measured by the vertical distance between the forecast and the outcome), especially at the start of the 'Great Recession', which are not corrected until near the trough. Almost irrespective of the forecasting device used, forecast failure would be rare in a stationary process, so intermittent episodes of forecast failure confirm that many time series are not stationary.

Figure 12 illustrated the difficulties facing forecasting deriving from wide-sense non-stationarity. Similar problems afflict the formation of expectations by economic actors: in theory models, today's expectation of tomorrow's outcome is often based on the 'most likely outcome', namely the conditional expectation of today's distribution of possible outcomes. In processes that are non-stationary from location shifts, previous expectations can be poor estimates of the next period's outcome, as Figure 9 illustrated, with adverse implications for economic theories of expectations based on so-called 'rational' expectations.

7. Recommendations for forecasting facing non-stationarity

Everything was happening so oddly that she didn't feel a bit surprised.
A reference to Alice from Lewis Carroll (1899).

Given the hazards of forecasting non-stationary variables, what can be done? First, be wary of forecasting I(1) processes over long time horizons. Modellers and policymakers must establish when they are dealing with integrated series, and acknowledge that forecasts then entail high uncertainty. The danger is that uncertainty can be masked by using mis-specified models which enforce stationarity, as seen in Figure 11, greatly reducing the measured uncertainty without reducing the actual, a recipe for poor policy and intermittent forecast failure. As Sir Alex Cairncross worried in the 1960s: 'A trend is a trend is a trend, but the question is, will it bend? Will it alter its course through some unforeseen force, and come to a premature end?'

Second, once forecast failure has been experienced, detection of location shifts (see §5.1) can be used to correct forecasts even with few observations, or alternatively it is possible to switch to more robust forecasting devices that adjust quickly to location shifts, removing much of any systematic forecast biases, but at the cost of wider interval forecasts (see e.g., Clements and Hendry, 1999). In turbulent times, such devices are an example of a method with no necessary verisimilitude that can outperform the in-sample correct representation. Figure 13 illustrates the substantial improvement in the one-step ahead forecasts of the log of UK GDP over 2008–2012 using a robust forecasting device, compared to a 'conventional' method. The robust device has a much smaller bias and MSFE, but as it is knowingly mis-specified, clearly does not justify selecting it as an economic model.

Thus, it is important to refrain from linking out-of-sample forecast performance of models to their 'quality' or verisimilitude. When unpredictable location shifts occur, there is no necessary link between forecast performance

and how close the underlying model is to the truth. Both good and poor models can forecast well or badly depending on unanticipated shifts.

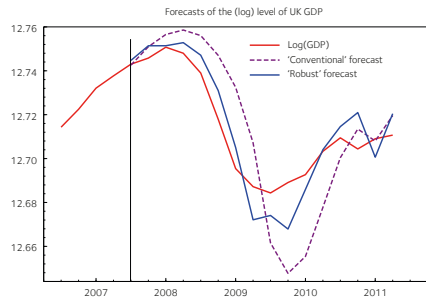Forecasts of the (log) level of UK GDP

Figure 13: One-step ahead forecasts of the log of UK GDP over 2008–2012 by 'conventional' and robust methods.

Third, the huge class of equilibrium-correction models includes almost all regression models for time series, autoregressive equations, vector autoregressive systems, cointegrated systems, dynamic-stochastic general equilibrium (DSGE) models, and many of the popular forms of model for autoregressive heteroskedasticity (see Engle, 1982). Unfortunately, all of these suffer from systematic forecast failure after shifts in their long-run, or equilibrium, means. Indeed, because they have in-built constant equilibria, their forecasts tend to go up (down) when outcomes go down (up), as they converge back to previous equilibria. Consequently, while cointegration captures equilibrium correction, care is required when using such models for genuine out-of-sample forecasts after any forecast failure has been experienced.

7.1. Some benefits of non-stationarity

Non-stationarity is pervasive, and as we have documented, needs to be handled carefully to produce viable empirical models, but its occurrence is not all bad news. When time series are I(1), their variance grows over time, which can help establish long-run relationships. Some economists believe that so-called 'observational equivalence'—where several different theories look alike on all data—is an important problem. While that worry could be true in a stationary world, cointegration can only hold between I(1) variables that are genuinely linked. 'Observational equivalence' is also unlikely facing location shifts: no matter how many co-breaking relations exist, there must always be fewer than the number of variables, as some must shift to change others, separating the sheep from the goats.

When I(1) variables also trend, or drift, that can reveal the underlying links between variables even when measurement errors are quite large (see Duffy and Hendry, 2015). Those authors also establish the benefits of location shifts that co-break in identifying causal links between mis-measured variables: intuitively, simultaneous jumps in both variables clarify any 'fog' from measurement errors surrounding their relationship.

Moreover, empirical economics is plagued by very high correlations between variables (as well as over time), but location shifts can substantively reduce such collinearity.

Finally, location shifts also enable powerful tests of the invariance of policy models to policy changes before new policies are implemented, potentially avoiding poor policy outcomes. Thus, while wide-sense non-stationarity certainly poses problems for economic theories, empirical modelling and forecasting, there are benefits to be gained as well.

8. Conclusions

'The Eighth Square at last!' she cried as she bounded across.
Alice, from Lewis Carroll (1899).

Non-stationary time series are the norm in many disciplines including economics, climatology, and demography as illustrated in Figures 1–5: the world changes, often in unanticipated ways. Research, and especially policy, must acknowledge the hazards of modelling what we have called wide-sense non-stationary time series, where distributions of outcomes change, as illustrated in Figure 9. The two most common sources of non-stationarity in practice are stochastic trends and location shifts. Individually and together when not addressed, they can distort in-sample inferences, lead to systematic forecast failure out-of-sample, and substantively increase forecast uncertainty as seen in Figure 11. However, both forms can be tamed in part using the methods of cointegration and modelling location shifts respectively, as Figure 10 showed.

Non-stationarity has important implications for inter-temporal theory, empirical modelling, forecasting and policy. Theory formulations need to account for humans inevitably facing disequilibria, so need strategies for correcting errors after unanticipated location shifts. Empirical models must check for genuine long-run connections between variables using cointegration techniques, detect past location shifts, and incorporate feedbacks implementing how agents correct their previous mistakes. Forecasts must allow for the uncertainty arising from cumulating shocks, and could switch to robust devices after systematic failures. Tests have been formulated to check for models not being invariant to location shifts, and for policy changes even causing such shifts, potentially revealing that those models should not be used in future policy decisions.

Policymakers must explicitly recognise the challenges of implementing policy in non-stationary environments. Regulation of integrated processes such as atmospheric

CO_2 concentrations is challenging due to their accumulation: for example, in climate policy, net-zero emissions are required to stabilise outcomes (see Allen, 2015). Invariance of the parameters in policy models to a policy shift is a necessary condition for that policy to be effective and consistent with anticipated outcomes. The possibility of location shifts does not seem to have been included in risk models of financial institutions, even though such shifts will generate many apparently extremely unlikely successive bad draws relative to the prevailing distribution, as seen in Figure 9.

Caution is advisable when acting on forecasts of integrated series or during turbulent times, potentially leading to high forecast uncertainty and systematic forecast failure, as seen in Figures 12 and 13. Conversely, the tools described above for handling shifts in time series enabled Statistics Norway to quickly revise their economic forecasts after Lehman Brothers' bankruptcy. Demographic projections not only face evolving birth and death rates as in Figure 5, but also sudden shifts, as with migration, so like economics, must tackle the two forms of non-stationarity simultaneously.

Location shifts that affect the equilibrium means of cointegrating models initially cause systematic forecast failure, then often lead to incorrectly predicting rapid recovery following a fall, but later under-estimating a subsequent recovery. Flash estimates of GDP frequently exhibit that problem, as many methods for interpolating missing data on disaggregates implicitly treat them as I(0). Using robust forecasting devices like those recorded in Figure 13 after a shift or forecast failure can help alleviate both problems.
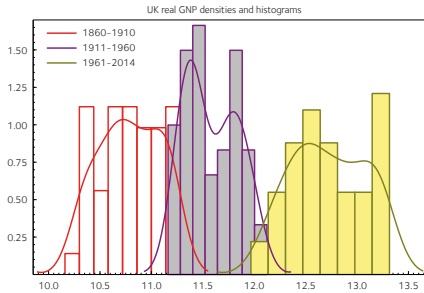
Figure 14: Histograms and densities of logs of UK real GNP in each of three 50-year epochs.

The key feature of every non-stationary process is that the distribution of outcomes shifts over time, illustrated in Figure 14 for histograms and densities of logs of UK real GNP in each of three 50-year epochs. Consequently, probabilities of events calculated in one time period do not apply in another: recent examples include increasing longevity affecting pension costs, and changes in frequencies of flooding vitiating flood-defence systems.

On the other hand, we also noted some benefits of stochastic trends and location shifts revealing genuine links between variables, which is invaluable knowledge in a policy context.

References

Allen, M. (2015). Short-lived Promise? The Science and Policy of Cumulative and Short-lived Climate Pollutants. Oxford Martin Policy Paper, Oxford Martin School, Oxford, UK.

Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de L'École Normale Supérieure 3*, 21–86.

Box, G. E. P. and G. M. Jenkins (1976). *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day. First published, 1970.

Carroll, L. (1865). *Alice's Adventures in Wonderland*. London: Macmillan and Co.

Carroll, L. (1899). *Through the Looking-Glass and What Alice Found There*. London: Macmillan and Co.

Castle, J. L., J. A. Doornik, D. F. Hendry, and F. Pretis (2015). Detecting location shifts during model selection by step-indicator saturation. *Econometrics 3(2)*, 240–264.

Castle, J. L. and N. Shephard (Eds.) (2009). *The Methodology and Practice of Econometrics*. Oxford: Oxford University Press.

Clements, M. P. and D. F. Hendry (1998). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.

Clements, M. P. and D. F. Hendry (1999). *Forecasting Non-stationary Economic Time Series*. Cambridge, Mass.: MIT Press.

Davidson, J. E. H., D. F. Hendry, F. Srba, and J. S. Yeo (1978). Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. *Economic Journal 88*, 661–692.

Doornik, J. A. (2009). *Autometrics*. See Castle and Shephard (2009), pp. 88–121.

Doornik, J. A. and D. F. Hendry (2013). *Empirical Econometric Modelling using Pc-Give: Volume I*. (7th ed.). London: Timberlake Consultants Press.

Duffy, J. and D. F. Hendry (2015). The impact of near-integrated measurement errors on modelling long-run macroeconomic time series. Discussion paper, Economics Department, Oxford University.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflation. *Econometrica 50*, 987–1007.

Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics 16*, 121–130.

Granger, C. W. J. and P. Newbold (1974). Spurious regressions in econometrics. *Journal of Econometrics 2*, 111–120.

Hendry, D. F. (1980). Econometrics: Alchemy or science? *Economica 47*, 387–406.

Hendry, D. F. (1999). An econometric analysis of US food expenditure, 1931–1989. See Magnus and Morgan (1999), pp. 341–361.

Hendry, D. F. (2004). The Nobel Memorial Prize for Clive W.J. Granger. *Scandinavian Journal of Economics 106*, 187–213.

Hendry, D. F. (2009). The methodology of empirical econometric modeling: Applied econometrics through the looking-glass. In T. C. Mills and K. D. Patterson (Eds.), *Palgrave Handbook of Econometrics*, pp. 3–67. Basingstoke: Palgrave Macmillan.

Hendry, D. F. (2015). *Introductory Macro-econometrics: A New Approach*. London: Timberlake Consultants. http://www.timberlake.co.uk/macroeconometrics.html.

Hendry, D. F. and K. Juselius (2000). Explaining cointegration analysis: Part I. *Energy Journal 21*, 1–42.

Hendry, D. F. and G. E. Mizon (2014). Unpredictability in economic analysis, econometric modeling and forecasting. *Journal of Econometrics 182*, 186–195.

Hendry, D. F. and M. S. Morgan (Eds.) (1995). *The Foundations of Econometric Analysis.* Cambridge: Cambridge University Press.

Hendry, D. F. and C. Santos (2010). An automatic test of super exogeneity. In M. W. Watson, T. Bollerslev, and J. Russell (Eds.), *Volatility and Time Series Econometrics*, pp. 164–193. Oxford: Oxford University Press.

Johansen, S. and B. Nielsen (2009). An analysis of the indicator saturation estimator as a robust regression estimator. See Castle and Shephard (2009), pp. 1–36.

Kaufmann, R. K., H. Kauppi, M. L. Mann, and J. H. Stock (2013). Does temperature contain a stochastic trend: linking statistical results to physical mechanisms. *Climatic Change 118*(3-4), 729–743.

Lee, R. D. and L. R. Carter (1992). Modelling and forecasting U.S. mortality. *Journal of the American Statistical Association 87*, 659–671.

Magnus, J. R. and M. S. Morgan (Eds.) (1999). *Methodology and Tacit Knowledge: Two Experiments in Econometrics.* Chichester: John Wiley and Sons.

Morgan, M. S. (1990). *The History of Econometric Ideas.* Cambridge: Cambridge University Press.

Phillips, P. C. B. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics 33*, 311–340.

Pretis, F. (2015). Econometric models of climate systems: The equivalence of two-component energy balance models and cointegrated VARs. Working Paper 750, Economics Department, Oxford University.

Pretis, F., J. Reade, and G. Sucarrat (2016). General-to-specific (gets) modelling and indicator saturation with the *R* package *gets*. Working Paper 794, Economics Department, Oxford University.

Qin, D. (1993). *The Formation of Econometrics: A Historical Perspective.* Oxford: Clarendon Press.

Qin, D. (2013). *A History of Econometrics: The Reformation from the 1970s.* Oxford: Clarendon Press.

Soros, G. (2008). *The New Paradigm for Financial Markets.* London: Perseus Books.

Taleb, N. N. (2007). *The Black Swan.* New York: Random House.

Yule, G. U. (1897). On the theory of correlation. *Journal of the Royal Statistical Society 60*, 812–838.

Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time series (with discussion). *Journal of the Royal Statistical Society 89*, 1–64.

Notes

1   This is called weak stationarity: other definitions exist, but we will mainly use that concept.

2   It is sometimes argued that economic time series could be stationary around a deterministic trend, but it seems unlikely that GNP would continue trending up if nobody worked.

3   Extensive histories of econometrics are provided by Morgan (1990), Qin (1993, 2013), and Hendry and Morgan (1995).

4   Davidson, Hendry, Srba and Yeo (1978), and much of the subsequent literature, call these 'error correction'.

About the authors

Professor Sir David Hendry is Director of the Economic Modelling Programme (EMoD) within the Institute for New Economic Thinking at the Oxford Martin School, Professor of Economics at the University of Oxford, and a Fellow of Nuffield College. He was previously Professor of Econometrics at the London School of Economics and Political Science. He was knighted in 2009 and is an Honorary Vice-President and past President of the Royal Economic Society; Fellow of the British Academy, Royal Society of Edinburgh, Econometric Society, Academy of Social Sciences, and Journal of Econometrics; Foreign Honorary Member, American Academy of Arts and Sciences and American Economic Association; and Honorary Fellow, International Institute of Forecasters. He is a Thomson Reuters Citation Laureate, and has published more than 200 papers and 25 books on econometrics, climate change, and forecasting.

Dr Felix Pretis is a British Academy Post-Doctoral Research Fellow at the Department of Economics at the University of Oxford, co-director of the Climate Econometrics project and a James Martin Fellow affiliated with EMoD within the Institute for New Economic Thinking at the Oxford Martin School. He received his DPhil in Economics from the University of Oxford. He has published papers on econometric methods in climatology and the detection of structural breaks in time series, and has taught statistical model selection and forecasting at the University of Oxford as well as universities in Mexico, Japan, France, and Cuba.

Acknowledgments

Economies, societies, and many natural systems evolve and change, sometimes dramatically, so good models and accurate forecasts are vital for policymakers to prepare for and navigate these changes successfully. Yet history is littered with forecasts that went badly wrong, sharply illustrated during the recent recession. A glance at most economic and related time series, such as greenhouse gases, reveals the invalidity of an assumption of stationarity, whereby the mean and variance are constant over time. Nevertheless, many models used in empirical research, forecasting or for guiding policy have been predicated on treating observed data as stationary, when in fact such analysis must take non-stationarity into account if it is to deliver useful outcomes. The problem for policymakers is not a plethora of excellent models from which to choose, but to find stable relationships that survive long enough to be useful. This paper offers guidance for policymakers and researchers on identifying what forms of non-stationarity are prevalent, what hazards each form implies for empirical modelling and forecasting, and for any resulting policy decisions, and what tools are available to overcome such hazards.

Oxford Martin Policy Papers harness the interdisciplinary research of the Oxford Martin School to address today's critical policy gaps. The development of these policy papers is overseen by the Oxford Martin Policy and Research Steering Group, chaired by Professor Sir John Beddington FRS, Senior Adviser to the Oxford Martin School. This paper reflects the views of the authors, and does not necessarily reflect the position of this Steering Group, the Oxford Martin School or the University of Oxford.

OXFORD
MARTIN
SCHOOL

UNIVERSITY OF
OXFORD