



Nuffield
College
UNIVERSITY OF OXFORD

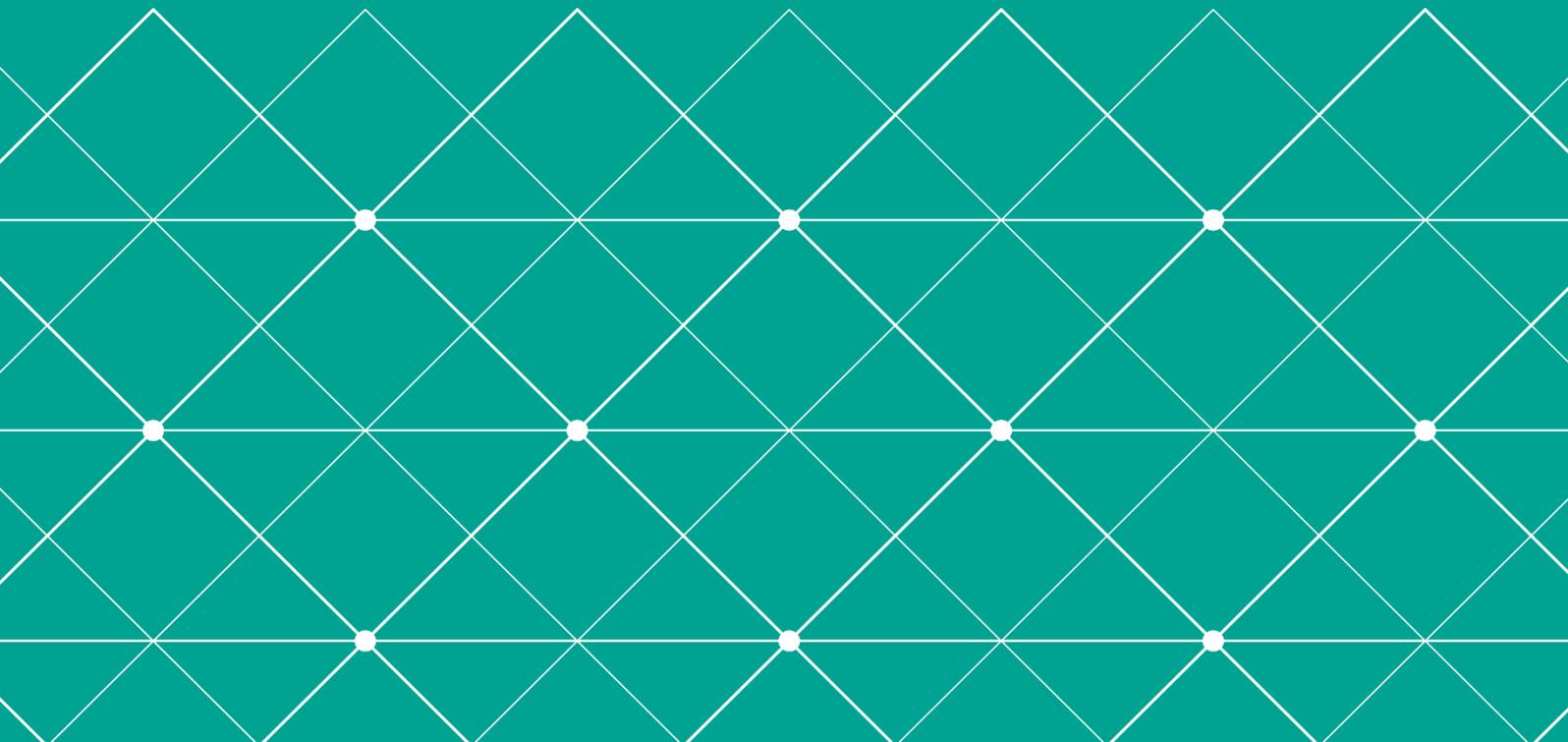
ECONOMICS DISCUSSION PAPERS

Paper No: 2021-W03

Econometrics for Modelling Climate Change

By Jennifer L. Castle and David F. Hendry

May 2021



Econometrics for Modelling Climate Change

Jennifer L. Castle and David F. Hendry*
Climate Econometrics, Nuffield College, University of Oxford, UK

March, 2021

Abstract

Greenhouse gas emissions, such as carbon dioxide, nitrous oxide and methane, are a major cause of climate change as they cumulate in the atmosphere and re-radiate the sun's energy. As such emissions are currently mainly due to economic activity, economic and climate time series share many features such as considerable inertia, stochastic trends and distributional shifts. Consequently, tools for empirically modelling non-stationary economic outcomes are also appropriate for studying many aspects of observational climate-change data. Moreover, both disciplines lack complete knowledge of their respective data generating processes (DGPs), so model search retaining viable theory but allowing for shifting distributions is important. Reliable modelling of both climate and economic-related time series requires finding an unknown DGP (or close approximation thereto) to represent multivariate evolving processes subject to abrupt shifts. Consequently, to ensure that DGP is nested within a much larger set of candidate determinants, model formulations to search over should comprise all potentially relevant variables, their dynamics, indicators for perturbing outliers, shifts, trend breaks and non-linear functions, while retaining well-established theoretical insights.

The model selection approach at Climate Econometrics uses a variant of machine learning with multi-path block searches commencing from very general specifications, usually with more candidate explanatory variables than observations, to discover well-specified and undominated models of the non-stationary processes under analysis. To do so requires applying appropriate indicator saturation estimators (ISEs), a class that includes impulse indicators for outliers, step indicators for location shifts, multiplicative indicators for parameter changes, and trend indicators for trend breaks, all of which are illustrated here in simple settings. All ISEs entail more candidate variables than observations, often by a large margin when implementing combinations, yet can detect the impacts of shifts and policy interventions to avoid non-constant parameters in models, as well as improve forecasts. To characterize non-stationary observational data one must handle all substantively relevant features jointly: a failure to do so leads to non-constant and mis-specified models and hence incorrect theory evaluation and policy analyses. The approach is applied to empirical climate modelling including an application to a model of UK CO₂ emissions.

JEL classifications: C5, C01, C18, C87, Q54.

Keywords: Climate Change; Greenhouse Gas Emissions; Climate Econometrics; Saturation Estimation; Model Selection; *Autometrics*.

*Prepared for the *Oxford Research Encyclopedia of Economics and Finance*: please do not quote without permission from the authors. Financial support from the Robertson Foundation (award 9907422) and Nuffield College is gratefully acknowledged, as are important contributions to the background research from Susana Campos-Martins, Jurgen A. Doornik, Luke P. Jackson, Søren Johansen, Jonas Kurler, Andrew B. Martinez, Bent Nielsen, Felix Pretis and Ryan Rafaty. All calculations and graphs use *PcGive* (Doornik and Hendry, 2018) and *OxMetrics* (Doornik, 2018). email: jennifer.castle@magd.ox.ac.uk and david.hendry@nuffield.ox.ac.uk

Contents

1	Introduction	5
2	Taking Stock of Climate Change: Earth, Air, Fire and Water	8
2.1	Earth	10
2.2	Air	11
2.3	Fire	12
2.4	Water	12
3	Stationary econometric theory and non-stationary time series	14
4	The key implications of shifting distributions	15
5	The importance of handling shifts	18
5.1	Simulating SIS for (15) with an intercept shift	20
5.2	Simulating SIS for (15) with a shift in dynamics	20
5.3	Why is SIS effective for a shift in dynamics?	22
5.4	Tackling a shift in dynamics by multiplicative indicator saturation	23
5.5	Tackling changing trends by trend-indicator saturation	23
6	Modelling and forecasting wide-sense non-stationary processes	27
6.1	Selecting models	28
6.2	Forecasting in non-stationary worlds	29
6.3	Exogeneity in Ice-Ages forecasts	30
6.4	Explaining Cardt	31
6.5	Smooth robust forecasts	32
6.6	The ‘Paradox’ of stagnant real wages yet rising real incomes	35
7	Updating the UK CO₂ model	36
7.1	Cointegration	38
7.2	‘Forecast’ evaluation	40
7.3	Super exogeneity tests	42
7.4	Evaluating the UK’s 2008 Climate Change Act	43
7.5	Implications of the UK’s CO ₂ emissions model	43
8	Conclusions	44
9	Appendix	51

Figures

1	(a) Monthly atmospheric CO ₂ measured at Mauna Loa in parts per million (ppm); (b) annual global ocean heat content to a depth of 700m over 1957–2019 in 10 ²² Joules; (c) annual global mean surface temperature deviations in degrees K since 1880; (d) annual UK coal use in millions of tonnes (Mt) since 1860.	5
2	(a) Thousand-year changes in parts per million (ppm) of CO ₂ in the atmosphere, ending with CO ₂ changes in the last 250 years; (b) reductions in CO ₂ during 2020 from the SARS-Cov-2 pandemic (Mt).	6
3	Ice-age time series: (a) Ice volume ; (b) atmospheric CO ₂ in parts per million (ppm); (c) Antarctic temperature; (d) 3D plot of Ice, Antarctic temperature and atmospheric CO ₂ levels.	9
4	Ice-age orbital drivers: (a) eccentricity (<i>Ec</i>); (b) obliquity (<i>Ob</i>); (c) precession (<i>Pr</i>); (d) Summer-time insolation at 65° south (<i>St</i>).	10
5	Rivers-in-the-sky; dust-storm; wild fire; coastal flooding.	11
6	Frequency of Central England Winter Temperatures below 2° per decade. Data from Met Office Hadley Centre via https://www.metoffice.gov.uk/hadobs/hadcet/	14
7	Trends across seasons. Left: CET over 1659–2019; Right: Boston over 1743–2015.	15
8	Illustrating the impact of a distributional shift on expectations.	16
9	(a) Monthly changes in atmospheric CO ₂ ; (b) annual changes in global ocean heat; (c) annual changes in global mean surface temperature; (d) annual changes in UK coal use.	18
10	(a) A time series from (15); (b), (c) recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs); (d) forecast Chow tests.	19
11	(a) One time series from (15); (b), (c) simulated recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs) when applying SIS; (d) resulting forecast Chow tests.	20
12	(a) A time series from (20); (b), (c) recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs); (d) forecast Chow tests.	21
13	(a) A time series from (20); (b), (c) simulated recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs) when applying SIS; (d) resulting forecast Chow tests.	22
14	(a) A time series from (20); (b), (c) simulated recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs) when applying MIS; (d) resulting forecast Chow tests.	23
15	UK time series of (a) output per worker per year (productivity); (b) births per thousand of the population; (c) employment; (d) cumulative confirmed COVID-19 cases.	24
16	Time series from (27) and (28) of (a) y_t ; (b) z_t ; (c) Δy_t ; (d) Δz_t	25
17	Actual and fitted values and residuals from (a), (b) regressing y_t on z_t without TIS and (c), (d) with.	26
18	Actual and fitted values and residuals from selecting: (a), (b) Δy_t from Δz_t with SIS; (c), (d) y_t with TIS when z_t is forced; (e), (f) Δy_t on SIS with Δz_t forced.	26
19	Actual and forecast values for changes and derived levels from models of Δy_t on Δz_t without SIS (a), (b) denoted $(\hat{\quad})$ and with (c), (d), $(\tilde{\quad})$	27
20	Actual and forecast values for levels from models of (a) y_t on a forced constant and forced z_t with TIS $(\hat{\quad})$ and (b) without TIS $(\tilde{\quad})$, and (c) with a step intercept correction from $t = 84$ $(\bar{\quad})$	28
21	Three scenarios of different atmospheric CO ₂ levels: endogenous, and fixed at 400ppm and 560ppm for (a) Ice volume; (b) temperature.	32
22	Explaining forecasting by Cardt: (a) trend line; (b) Cardt forecast; (c) 3 earlier-based forecasts; (d) average forecast added.	33

23	(a) UK productivity since 1860 with overall and six separate trend lines at roughly 25-year subperiods; (b) UK productivity since 1860 with seven TIS trend lines selected at 0.01%.	33
24	Forecasts of UK productivity: (a) Office for Budget Responsibility; (b) a smooth random walk; (c) smooth double difference device; (d) smooth double robust device.	34
25	Forecasts of UK productivity using Cardt: (a) annual dynamic forecasts estimated from 2000 onwards; (b) the annual dynamic forecasts recorded on a longer time-series; (c) 5 year ahead forecasts for the quarterly data estimated from 1997Q1 onwards with uncertainty bands; (d) the quarterly 5 year ahead forecasts (produced every two quarters) recorded on a longer time-series.	35
26	(a) UK log real wages and log output per worker per year; (b) Log wage share since 1995; (c) Logs of employment and population; (d) Log real GDP per capita and log real GDP per employee.	36
27	(a) UK CO ₂ emissions per capita in tons per annum (p.a.) to 2018; (b) UK coal (millions of tonnes, Mt), oil (Mt), natural gas (millions of tonnes of oil equivalent, Mtoe) and wind+solar (Mtoe), all to 2018; (c) ratio of CO ₂ emissions to the capital stock on a log scale to 2017; (d) US CO ₂ emissions per capita, in tons p.a., 1850–2019	37
28	Shifting sub-period distributions of UK CO ₂ emissions.	37
29	(a) E_t and \tilde{E}_t ; (b) \tilde{Q}_t without impulse indicators, centered on a mean of zero.	39
30	(a) Actual, fitted and ‘forecast’ values for ΔE_t from (39); (b) residuals and ‘forecast’ errors scaled by the equation standard error; (c) residual density and histogram with a Normal density; (d) residual autocorrelation.	40
31	Actual, fitted and ‘forecast’ values for UK CO ₂ emissions.	40
32	Conditional 1-step ‘forecasts’ for ΔE_t (a) from (39) over 2014–2017 (denoted $\hat{\cdot}$) with ± 2 forecast standard error bars and the robust forecasts $\widetilde{\Delta E}_{T+h T+h-1}$; (b) the derived ‘forecasts’ in levels; (c) same as (a) but commencing in 2009.	41
33	(a) Outcomes and 1-step ‘forecasts’ ± 2 forecast standard errors as bars and fans, without and with indicators; (b) Outcomes and h -step ‘forecasts’ and a comparison with Cardt ‘forecasts’.	42
34	(a) UK CO ₂ emissions and CCA2008 CO ₂ targets; (b) deviations from the target values with step indicators.	44

1 Introduction

The four words in our title of ‘econometrics’, ‘modelling’, ‘climate’ and ‘change’ are obviously key to this chapter, both singly and jointly, but closely connected because of **change**: Hepburn and Schwarz (2020) provide clear answers to most common questions about climate change. Greenhouse gas (GHG) emissions, especially carbon dioxide (CO₂), methane and nitrous oxide, have been increasing at a rapid rate, now cumulating in the atmosphere at more than 3 parts per million per annum (about 24 gigatons of CO₂) as shown in Figure 1a from 1958(1) to 2021(2). The single linear trend in Panel (a) is to emphasize that the increases are also increasing. If humanity is to avoid catastrophic climate change, the present upward trends in GHG emissions must be reversed to become rapid downward trends. However, such changes will not be smooth, but rather erratic as different sources of emissions can be more easily reduced and perhaps eliminated, as the United Kingdom has almost done with coal as shown in Panel (d). Figure 1 also illustrates the shifts and changing trends in two other climate time series, namely ocean heat content in Panel (b) and global mean surface temperature in Panel (c).¹

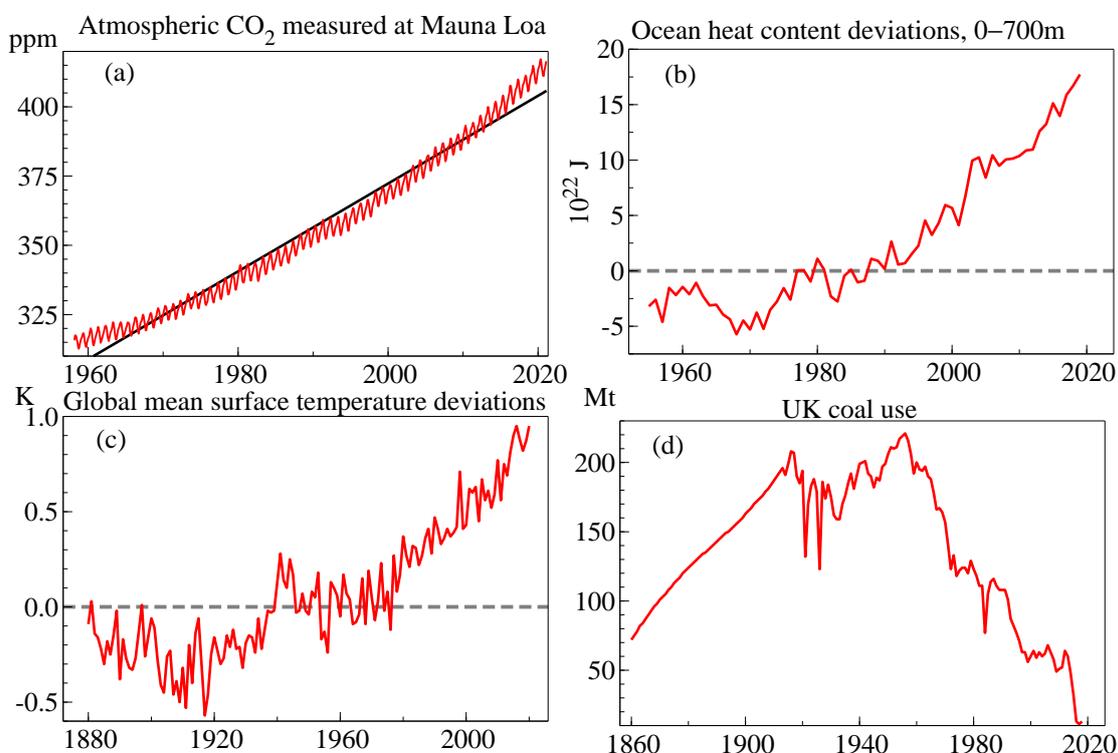


Figure 1: (a) Monthly atmospheric CO₂ measured at Mauna Loa in parts per million (ppm); (b) annual global ocean heat content to a depth of 700m over 1957–2019 in 10²²Joules; (c) annual global mean surface temperature deviations in degrees K since 1880; (d) annual UK coal use in millions of tonnes (Mt) since 1860.

It is not all bad news, as the data on UK coal use (coal is one of the worst pollutants and CO₂ emitters) show its near complete elimination over the same period when the UK economy actually grew considerably in per capita terms. The other aspect of Figure 1 is to emphasize how great and how fast climate change is occurring, especially since 1980, with substantial rises in ocean heat and air temperatures.

Atmospheric CO₂ concentrations lay within the range of roughly 175ppm to 300ppm over 800,000 years of Ice-Ages before the Industrial Revolution. Back then, it took thousands of years to move from

¹Sources: (a) <https://siweb.ucsd.edu/programs/keelingcurve/>; (b) https://www.nodc.noaa.gov/OC5/3M_heat_content/; (c) <https://climate.nasa.gov/vital-signs/global-temperature/>; (d) <http://www.carbonbrief.org/analysis-uk-cuts-carbon-record-coal-drop>.

its lowest level to its highest and down again (see e.g., Castle and Hendry, 2020, §6). Yet as Figure 2(a) shows in dramatic detail, humanity has added almost 100ppm in 60 years.

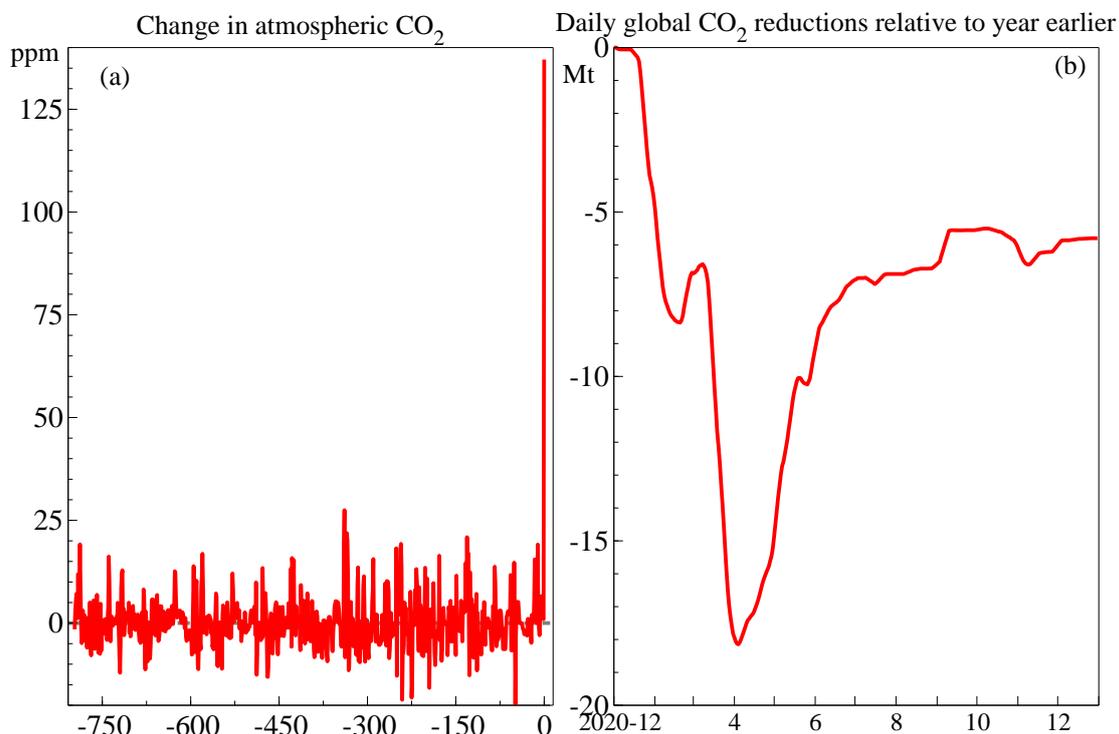


Figure 2: (a) Thousand-year changes in parts per million (ppm) of CO₂ in the atmosphere, ending with CO₂ changes in the last 250 years; (b) reductions in CO₂ during 2020 from the SARS-Cov-2 pandemic (Mt).

The SARS-Cov-2 pandemic with its all too frequent occurrences of intermittent ‘lockdowns’ and associated sudden drops in toxic emissions like nitrogen oxides and in CO₂ as seen in Figure 2(b),² may at last have driven home the need to handle both location shifts and trend breaks. However, it would take an eagle eye to see the small fall during 2020 of less than 20Mt in Figure 2(b) relative to the trend in CO₂ emissions in Figure 1a.

Greenhouse gases (GHGs) in the atmosphere, especially water vapour and carbon dioxide, are crucial in maintaining life. The Earth’s climate may seem stable currently, but it is always in flux, and always has been. When GHGs are too depleted, the planet cools, once being a ‘snowball’ (see e.g., Hoffman and Schrag, 2000) with glaciation in Death Valley, whereas excessive GHGs lead to very warm periods as in the Permian, Cretaceous and most ‘recently’, the Paleocene-Eocene Thermal Maximum (PETM) about 50 million years ago.

Past climate change was from natural forces, including plate tectonics, volcanism and developments like photosynthesis, so before the Anthropocene, planet Earth had experienced a wide range of climates. Given the present proliferation of life, many forms of life had to have survived despite these great changes. Indeed, the fossil record suggests life thrived in global temperatures much higher and lower than today’s. Nevertheless, that same record reveals that large numbers of species have disappeared, becoming extinct in the process of **climate change**, even if much later, new species evolved from survivors. In particular, the ‘mass extinctions’ visible in fossil records seem due to major climate changes, arising from a variety of causes. Global cooling occurred at the end of the Ordovician and Devonian periods, the latter from increased photosynthesis reducing atmospheric CO₂, whereas temperatures were far higher

²<https://public.wmo.int/en/programmes/global-climate-observing-system/global-climate-indicators>

during the worst mass extinction at the Permian–Triassic (P/Tr) boundary from massive volcanism creating the large igneous province (LIP) in Siberia. The mass extinction at the end of the Triassic was probably due to the massive LIP formation of the Central Atlantic Magmatic Province, whereas the well-known extinction of non-avian dinosaurs at the Cretaceous–Tertiary (K/T) boundary is attributed to the impact from a large meteor at Chicxulub near the Yucatan peninsula leading to large changes in climate, perhaps exacerbated by another LIP forming the Deccan Traps in India. Although all of these events occurred many millions of years ago, the central message that climate change was the key determinant in every case is still relevant: large scale warming or cooling both lead to major species extinctions.

Climate science has established a vast body of knowledge about the processes and causal links in the Earth’s climate system. The climate of planet Earth depends on the energy balance between incoming radiation from the Sun and re-radiation from the planet, mediated by the differential absorption and reflection properties of the land, atmosphere, ice and oceans, respectively discussed in Section 2 in terms of Earth, Air, Fire and Water. Climate-change analysis is primarily based on physical process models which embody laws of conservation and energy balance at a global level: see reports by the Intergovernmental Panel on Climate Change (IPCC: <https://www.ipcc.ch/>).

Such well-established climate theories can be retained in econometric models as their core theory (see e.g., Kaufmann, Kauppi, Mann, and Stock, 2013, Pretis, 2019 and Brock and Miller, 2020). For example, Kaufmann *et al.* (2013) link statistical models driven by stochastic trends to physical climate systems; and Pretis (2019) establishes an equivalence between a cointegrated vector autoregressive system (CVAR) and two-component (i.e., atmosphere and oceans) energy-balance models of the climate.

Nevertheless, climate science knowledge is incomplete. Greenhouse gas emissions depend on changeable human behaviour, essentially unpredictable volcanic eruptions that can have global climate impacts, and the rate of loss of sea ice, which both alters the Earth’s albedo and the oceans’ uptake and retention of CO₂. Moreover, which fossil fuels are burnt matters as their CO₂ emissions per million British thermal units (Btu) of energy produced differ substantially as Table 1 shows.³

Table 1: Pounds of CO₂ emitted per million Btu of energy produced.

Coal (anthracite)	228.6
Coal (bituminous)	205.7
Coal (lignite)	215.4
Coal (sub-bituminous)	214.3
Diesel fuel & heating oil	161.3
Gasoline	157.2
Propane	139.0
Natural gas	117.0

Switching from coal to natural gas almost halves the GHG emitted per Btu even if it still produces a considerable amount of CO₂ so is hardly a ‘solution’. Economic, social and behavioural changes all require empirically modelled relationships—hence the role for econometrics. Indeed, most social, economic and environmental observational time series are evolving processes with stochastic trends and sudden shifts, so are *wide-sense non-stationary*, and not just unit-root processes where differencing could create a stationary time series. A wide-sense non-stationary process does not have a constant distribution over time, and viable econometric methods must be able to tackle such changes. Moreover, the DGPs of wide-sense non-stationary time series are almost certainly unknown, so have to be discovered from the available evidence: see Hendry and Doornik (2014). Any approach to doing so will inevitably be heavily data based even if guided by subject-matter theory (which is plentiful in the climate context), although we note that physics itself has recently been approached in a similar vein: see Qin (2020).

³Source: US Department of Energy <https://www.eia.gov/tools/faqs/faq.php?id=73&t=11>.

Our approach at Climate Econometrics (<http://www.climateeconometrics.org/>, capitalized to differentiate it from the general research area) to modelling observational time series is complementary to physical-process climate models. Indicator saturation estimation is used to locate outliers, shifts and breaks, so entails that there are usually more candidate explanatory variables N than observations T . We use *Autometrics*, a variant of machine learning that explores multi-path block searches when $N > T$, to discover a well-specified and undominated model of the processes under analysis (see Doornik, 2009, also available in R by Pretis, Reade, and Sucarrat, 2018, and as the Excel Add-in *XLModeler*, <https://xlmodeler.com/>). While we focus on first moments in this chapter, Engle and Campos-Martins (2020) and Campos-Martins and Hendry (2020) provide studies of risk and volatility.

The structure of the chapter is as follows. Section 2 sets the stage for the empirical approach to modelling climate change by taking stock in the ancient, but still relevant, concepts of Earth, Air, Fire and Water. This discussion highlights the imperative for rapid action to address climate change, and considers where economics and social science enter. Given the context presented in Section 2 revealing the rapidly changing state of nature, Section 3 juxtaposes stationary econometric theory with non-stationary time series. Then Section 4 describes the unfortunate implications for statistical modelling theory and practice of shifting distributions. Section 5 discusses the importance of handling location shifts and parameter changes by saturation estimation in the simplest context of a first-order scalar autoregressive DGP, and how that might be achieved, then considers detecting trend breaks. In both cases, the general approach is aimed to tackle settings where shifts and breaks occur an unknown number of times at unknown dates by unknown magnitudes and directions relative to the model in use.

Section 6 describes our approach to jointly tackling all the main problems facing analyses of empirical evidence on wide-sense non-stationary processes in the framework of model discovery, from model formulation, through selection, to evaluation, before addressing how to forecast in a wide-sense non-stationary setting. The changing status from endogenous to exogenous regressors is observed in the production of forecasts over 110,000 years into the future. We review two forecasting methods designed to handle shifting distributions, Cardt (a univariate statistical predictor designed for non-stationary time-series data) and a smooth robust differenced device (using local estimates of the long-run mean and growth rates). Both are illustrated on UK productivity data. Section 7 updates a UK CO₂ model with two additional years of data. This enables the impact of the UK's Climate Change Act of 2008, captured in the model by a step dummy in 2010, to be evaluated as forecasts show it was key to a level shift down in UK CO₂ emissions. Section 8 concludes with data definitions and sources in the Appendix 9.

2 Taking Stock of Climate Change: Earth, Air, Fire and Water

Over the last 800,000 years, Ice-Ages have induced large switches in climate from cold to cool.⁴ Figure 3 graphs (a) the Ice volume measure; (b) atmospheric CO₂; (c) Antarctic temperature; and (d) 3D plot of Ice, Antarctic temperature and CO₂. The first three are recorded at 1000-year intervals in Figure 3, where the X-axes in such graphs are labelled by the time before the present starting 800,000 years ago. Visually, one can see that ice volumes cumulate more slowly than they melt, suggesting a non-linear relation. The changing albedo of ice coverage and the increasing release (rather than absorption) of CO₂ as oceans warm help explain the relative rapidity with which glacial periods switched. A second notable feature is

⁴Creating these data series has taken a massive international effort, collecting raw observations, where (e.g.) drilling in East Antarctica was completed to just a few meters above bedrock (see Parrenin *et al.*, 2007), then adjusting them to the common time scale and frequency of the European Project for Ice Coring in Antarctica–EPICA–Dome C (EDC3). Synchronization between the EPICA Dome C and Vostok ice core measures over the period –145,000 to the present was based on matching residues from volcanic eruptions (see Parrenin *et al.*, 2012). Ice volume estimates are from Lisiecki and Raymo (2005) (based on $\delta^{18}\text{O}$ as a proxy measure). Antarctic-based land surface temperature proxies were taken from Jouzel *et al.* (2007), and the paleo record from deep ice cores of atmospheric CO₂ from Lüthil *et al.* (2008). Sea level data, based on sediments, can be obtained from Siddall *et al.* (2003).

the considerable increase in the variation of temperature and atmospheric CO₂ after about 440,000 years ago, with much higher peaks, and corresponding deeper troughs, in ice volume. The 3D plot shows how the impact of a given temperature on ice volume changes with the level of atmospheric CO₂.

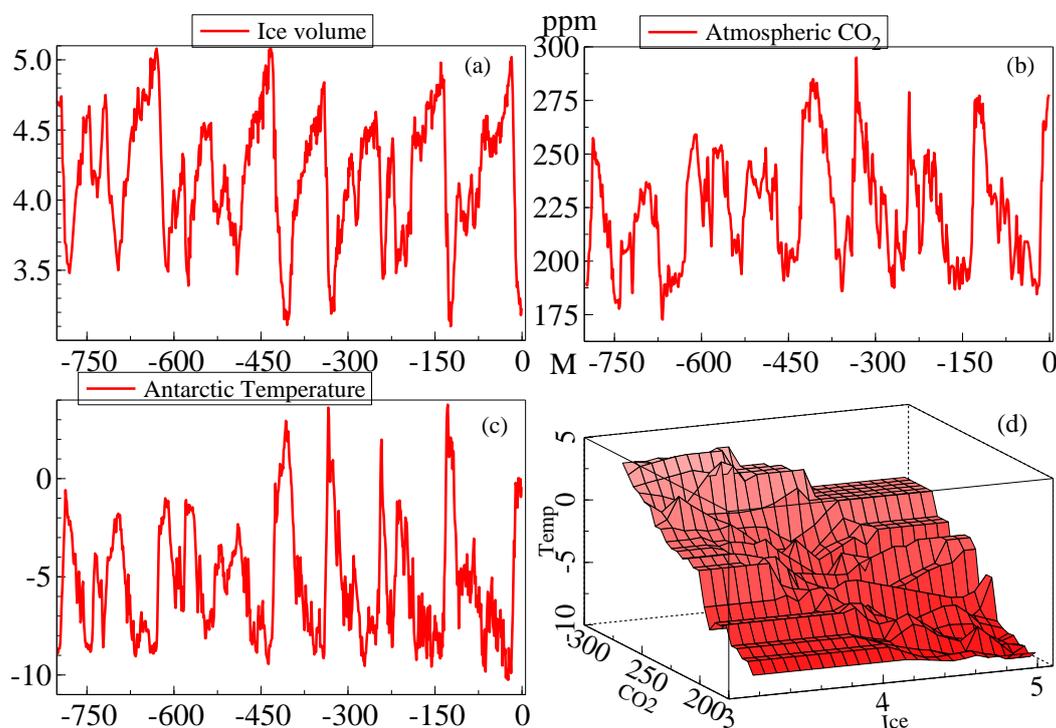


Figure 3: Ice-age time series: (a) Ice volume ; (b) atmospheric CO₂ in parts per million (ppm); (c) Antarctic temperature; (d) 3D plot of Ice, Antarctic temperature and atmospheric CO₂ levels.

Variations in the Earth’s orbital trajectory round the Sun were first hypothesised to drive Ice Ages by Croll (1875), and confirmed by Milankovitch (1969) (after whom the glacial cycles are usually named). Milankovitch calculated the resulting solar radiation at different latitudes and corrected Croll’s assumption that minimum winter temperatures precipitated Ice Ages by showing that cooler summer maxima were key to glaciation.

Figure 4 records (a), the eccentricity of the Earth’s orbit (where zero denotes circularity); (b) the ‘tilt’ of Earth relative to its axis (obliquity, measured in degrees); (c) precession of the equinox (which determines whether the Northern Hemisphere is closest to the Sun during its summer or winter, also measured in degrees) and (d) the resulting summer-time insolation calculated at 65° south (how much solar radiation reaches the planet at that latitude). The first three variables have periodicities of 100,000 years (varying from the gravitational influences of other planets in the solar system); 41,000 years for obliquity; and 2 cycles of 23,000 and 19,000 years for precession (partly due to the Earth not being an exact sphere). While these orbital series are strongly exogenous in any models of Earth’s climate, they seem to be non-stationary from shifting distributions, not unit roots. Interactions between these orbital variations also affect the lengths of glacial and interstadial periods as well as the timing of switches given the existing extent of ice coverage.

As stressed above, **change** is the key word—and humanity is now changing the climate by its vast emissions of greenhouse gases, especially from burning fossil fuels, dramatically highlighted by Figure 2. We now consider the Earth’s limited available land, atmosphere and water resources to show that humanity really can alter the climate, and is doing so in myriad ways. Earth, Air, Fire and Water have been ubiquitous concepts nearly globally from ancient times. Although they are not ‘elements’ as once believed, all four are ‘essential ingredients’ of life. The next four sub-sections discuss their roles in

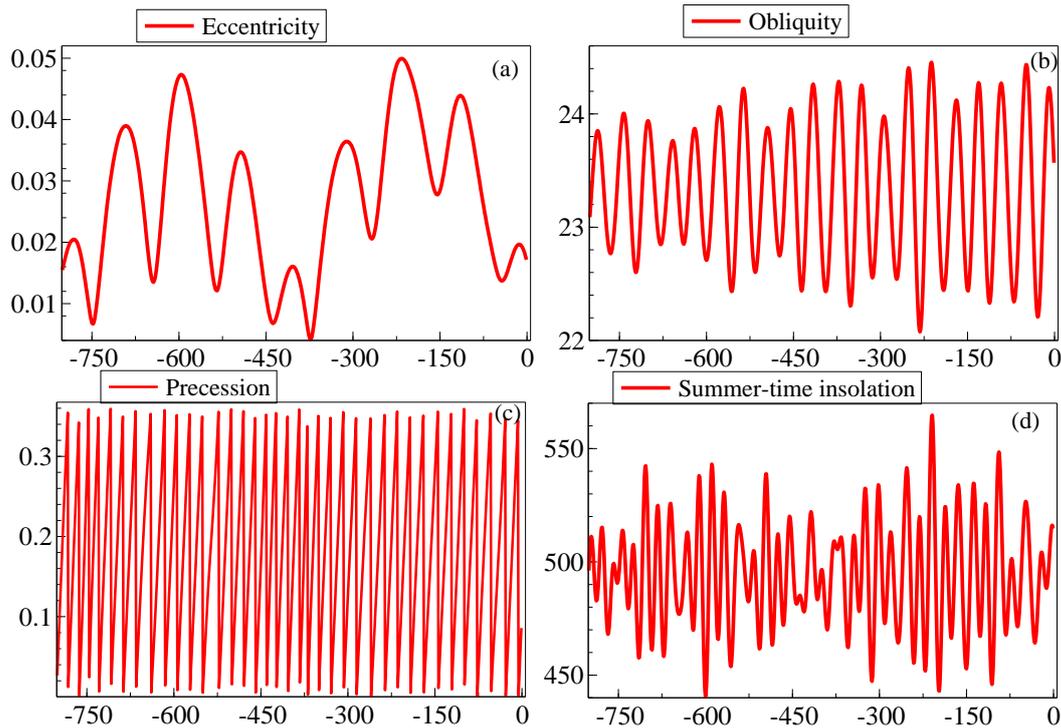


Figure 4: Ice-age orbital drivers: (a) eccentricity (Ec); (b) obliquity (Ob); (c) precession (Pr); (d) Summer-time insolation at 65° south (St).

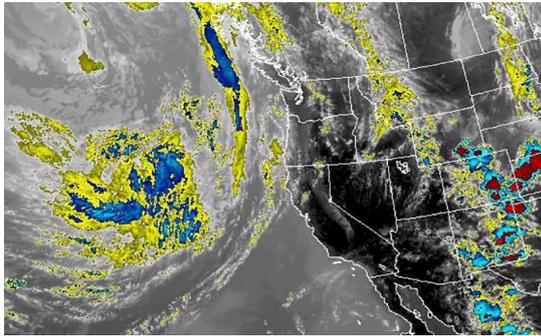
climate change, the dangers of precipitating an anthropogenic mass extinction, and actions humanity could take to avoid that.

2.1 Earth

Continents and their topography are shaped by plate tectonics and the resulting volcanic eruptions, both of which also affect climate and have played a key role in past great extinctions. ‘Earth’ is being used here as a place carrier for **land**, both as available living space and as providing soil for forests, other ‘wild’ areas, and agriculture for food supply—roughly 40% of the planet’s land area, or 50m km². Vegetation and soil together lock in about three times as much CO₂ as the atmosphere currently holds, as well as absorbing around a third of emissions. However, higher temperatures and greater rainfall will release some of that carbon: see Eglinton, Galy and Hemingway *et al.* (2021). Inner-city vertical and underground farms seem viable and may become essential as the climate warms.

Crops are frequently grown using artificial fertilizers (often made from methane, and leaching nitrous oxide in runoff) and based on farmland created by deforestation, wetland draining and mangrove removal, all adding to GHG emissions. Together with the methane given off by animal husbandry, especially by cattle, sheep and goats, these all lead to substantial GHG emissions—humanity’s climate change ‘foodprint’. Reducing such emissions will not be easy, but steps can be taken as discussed in <https://committees.parliament.uk/writtenevidence/21638/html/>. Land around volcanoes is fertile, so basalt dust could be added as a fertilizer and also absorbs atmospheric CO₂: see Nunes, Kautzmann, and Oliveira (2014) and Beerling, Kantzas and Lomas *et al.* (2020). Biochar produced by pyrolysis of biomass and added to soil could also increase crop yields while reducing GHG emissions: see Woolf, Amonette and Street-Perrott *et al.* (2010).

Climate change is increasing extreme land flooding from ‘rivers in the sky’, which can hold 15 times more water than the Mississippi River: witness the massive floods in Australia during mid-March 2021 spanning a large area, with almost a **meter** of rain falling at Nambucca Heads over 6 days. There have



<https://www.psl.noaa.gov/arportal>



<https://public.wmo.int/en/our-mandate/focus-areas/environment/SDS>



<https://earthobservatory.nasa.gov/images/81919/rim-fire-california>



https://en.wikipedia.org/wiki/Coastal_flooding

Figure 5: Rivers-in-the-sky; dust-storm; wild fire; coastal flooding.

also been massive floods in East Africa bringing locust swarms and diseases like cholera.

Climate change is also increasing extreme drought, leading to a loss of crops, with the resultant stress on some plants like sorghums producing toxic hydrogen cyanide (see Shehab, Yao and Wei *et al.*, 2020). Flooding and drought both lead to loss of soil, either from erosion or dust storms. Sea level rises cause coastal flooding, reducing usable land area, as well as forcing migration. Figure 5 illustrates.

2.2 Air

Air is again a place carrier, here denoting the **atmosphere** comprising mainly nitrogen (78%) and oxygen (21%), with greenhouse gases like water vapour (0.4%), carbon dioxide (CO₂), nitrous oxide (N₂O) and methane (CH₄) as well as ozone and some noble gasses. Earth's atmospheric blanket is essential to life, but seen from the space station, the atmosphere is a thin blue line round the planet, not much thicker than a sheet of paper round a soccer ball, so even small additional volumes of GHGs can have a large impact. Earth's gravity and magnetic field are essential to retain our atmosphere against the solar wind and also to protect the ozone layer from damaging radiation.

Atmospheric gases have changed greatly over deep time, especially from volcanism and the exchange of CO₂ for oxygen through photosynthesis, so Earth's range has included Ice Ages and tropical conditions, but Mars and Venus warn that atmospheric protection needs to be 'just right'. Increased greenhouse gases generated by human activity, especially burning fossil fuels, are now the major cause of climate change. GHGs receive then re-radiate energy at different wavelengths between ultraviolet and longwave infrared: this re-radiation is responsible for the atmospheric greenhouse effect. Eunice Foote (1856) showed that a flask of CO₂ heated greatly in the sun, whereas those of water vapour and dry air did not, closely followed by the independent experimental evidence in John Tyndall (1859). Ortiz and Jackson (2020) analyse Foote's contribution and Levendis, Kowalski, Lu, and Baldassarre (2020)

suggest an alternative experiment.

Nitrous oxide emissions from nitrogen and phosphate fertilizers have doubled in the last 50 years so now are about 7% of greenhouse gas emissions (see Tian, Xu and Canadell *et al.*, 2020). Catalytic converters add to this growing problem as N₂O is nearly 300 times more potent per molecule than CO₂ as a greenhouse gas.

Atmospheric methane is now double the highest level over the past 800,000 years. CH₄ is about 20 times as powerful as CO₂ as a GHG with a half-life in the upper atmosphere of around 15 years, gradually getting converted to CO₂. Current estimates of methane in hydrates are over 6 trillion tonnes, roughly twice the GHG equivalent in all fossil fuels: the release of even a small proportion of that would be disastrous.

Chlorofluorocarbons (CFCs) were destroying the ozone layer before the highly successful Montreal Protocol in 1987, a potential role model for a far more ambitious global commitment beyond the Paris Accord at CoP21. However, replacement refrigerants like halons and halocarbons including hydrochlorofluorocarbons (HCFCs), and hydrofluorocarbons (HFCs) remain dangerous greenhouse gases, meriting research for safer alternatives.

2.3 Fire

Fire is the place carrier for **energy**, currently obtained from burning vast volumes of fossil fuels. Humanity cannot continue to consume fossil fuels on the present scale yet stay within the ‘carbon budget’ required to achieve ‘net zero’ GHG emissions—essential to prevent dangerous climate change. The resulting changing global temperatures are leading to increased frequency and severity of wild fires from Australia, Amazon, and California to Siberia, which is a potential tipping point from tundra melting. Wild fires create fire-induced thunderclouds, known as pyrocumulonimbus clouds, which increase aerosol pollutants trapped in the stratosphere and upper atmosphere, can generate fire tornados, and lead to flash floods while also igniting further spot fires and temporarily cooling the planet like a moderate volcanic eruption. There is increased deforestation, especially in the Amazon and other tropical rainforests, with a loss of biodiversity as well as increased emissions from the burning of some of the forests.

The current main hope is to replace fossil fuel based energy by renewable sources from Earth (using thermal energy such as ground-heat and air-source heat pumps), Air (utilizing wind energy from onshore and offshore wind turbines), Fire (from sunlight via solar cells, and nuclear, potentially including small modular reactors) and Water (using hydroelectric energy from dams, diversion facilities or pumped storage facilities). Table 2 records recent estimates of electricity generating costs in £/MWh by different technologies. The costs of all renewable sources of electricity have been falling rapidly and seem likely to continue to do so. The share of UK electricity generated by renewables reached a peak of 60.5 per cent in April 2020, according to National Grid data.

Zero GHG electricity generation from renewables is technically feasible, but requires a huge increase in output and storage capacity (for windless cloudy periods) dependent on a very large investment. Sufficient supply could sustain electric transport, removing emissions from oil, and replace much household use of gas for heating and cooking. However, an electricity grid needs second by second balancing of electricity flows in an otherwise increasingly non-resilient system dependent on highly variable renewable supplies, so both backup and instantaneously accessible storage are essential, suggesting an intelligent system that could exploit electric vehicle-to-grid storage: see Noel *et al.* (2019).

2.4 Water

Water may seem limitless and it surprises many people that in fact there is very little water, especially fresh water. It is common to think that Earth is the ‘Blue Planet’ where we are surrounded by an abundance of water, but we are fooled by widespread shallow oceans: the Atlantic is only about 2.25 miles deep on average. The Pacific is wider and is deeper at about 2.65 miles on average: at its deepest in

Table 2: Power generating technology costs in £/MWh

Source	2015	2025	2040
Solar Large-scale PV (Photovoltaic)	80	44	33
Wind Onshore	62	46	44
Wind Offshore	102	57	40
Biomass	87	87	98
Nuclear Pressurized Water Reactor	93	93	93
Natural Gas Combined Cycle Gas Turbine (CCGT)	66	85	125
CCGT with carbon capture and sequestration (CCS)	110	85	82

Nuclear power guaranteed price of £92.50/MWh for Hinkley Point C in 2023. Lowest cost in **bold**; next lowest in *italic*; sans serif if less than 2015. Assumes increasing carbon taxes and falling CCS costs over time. Source: *Electricity Generation Costs 2020*, UK Department for Business, Energy and Industrial Strategy (BEIS)

the Challenger Deep of the Mariana Trench, it is roughly 6.8 miles down. In total, the Pacific holds 170 million cubic miles of water, just over half the 330 million cubic miles of water on Earth. If all sources of water on the planet were collected in a sphere, it would be just just 860 miles in diameter. Consequently, it is easy to heat the oceans by emitting excessive volumes of CO₂ into the atmosphere (see Figure 1b), pollute them, fill them with plastic waste, and turn sea water to a weak carbonic acid. Ocean acidification impacts on many ocean species, especially calcifying organisms like oysters and corals, but also on fish and seaweeds, affecting the entire ocean foodchain.

The worldwide ocean ‘conveyor belt’ circulates heat and nutrients, and carries oxygen to depths, maintaining the health of the oceans. A key driver is that warm water from the Gulf Stream moves north, evaporates, becomes saltier and denser, cools and so sinks, and flows south. Melting northern hemisphere ice could disrupt this circulation by diluting the denser salty water, as well as increase sea levels by about 18cm by 2100: see Hofer, Lang and Amory *et al.* (2020). Added to increasing volume from thermal expansion, rising sea levels have serious implications for coastal flooding, although the rises are not uniform, and recent coastal elevation measures have tripled estimates of global vulnerability to sea-level rises: see Kulp and Strauss (2019).

Conversely, Southern Ocean sea ice can dramatically lower ocean ventilation by reducing the atmospheric exposure time of surface waters and by decreasing the vertical mixing of deep ocean waters leading to holding 40ppm atmospheric CO₂ less at its maximum: see Stein, Timmermann, Kwon, and Friedrich (2020). Currently, oceans absorb almost a third of anthropogenic CO₂ emissions, but as they warm they will hold less.

Fresh water is a hugely important commodity and needs to be treated carefully. 19% of California’s electricity consumption goes toward water-related applications, such as treating, transporting, pumping and heating. Additionally, about 15% of its in-state electricity generation comes from hydropower, yet the frequency of both high- and low-flow extreme streamflow events have increased significantly across the United States and Canada over the last century (see Dethier, Sartain, Renshaw, and Magilligan, 2020).

Kelp ‘forests’ and seagrass ‘meadows’ absorb and store CO₂, help offset increasing carbonic acidification, provide nurseries for young sea life and help protect coasts against rising sea levels. Improving seaweed farming and raising aquaculture production by marine protection areas seem sensible, noting that off-shore wind farms also act as marine reserves.

While we have discussed each of Earth, Air, Fire and Water separately, the interactions between them are obvious. Given their roles in climate change, they form the basis of many of the time-series that we use to analyse climate change. In the next section we explore the problems of assuming such data are stationary when in fact they are highly non-stationary, in part due to the many changes discussed above.

3 Stationary econometric theory and non-stationary time series

In contrast to shifting observational data, much of econometric theory and its applications to time-series data implicitly or explicitly assumes that the data generating process (DGP) is stationary. Theory derivations all too often rely on proofs about the properties of estimators and tests that are invalidated by failures of stationarity that cannot be resolved by differencing (perhaps after taking logs) to ‘remove’ stochastic trends, because the means and variances of the differenced series under analysis remain non-constant. For example, the widely used theorem that the conditional expectation of a variable is the minimum-variance unbiased predictor is false when distributions shift; and the famous law of iterated expectations also fails: see Hendry and Mizon (2014). Before we consider these fundamental issues in detail in Section 4, as a contrast we will first describe the outcomes under stationarity.

In a stationary setting, key results about asymptotic distributions of estimators rely on the fact that ‘later data’ can improve estimates based on earlier observations as follows. Consider the following simple process over a period $t = 1, \dots, T$:

$$y_t = \beta + \epsilon_t \text{ where } \epsilon_t \sim \text{IN} [0, \sigma_\epsilon^2] \quad (1)$$

where β is a constant and $\text{IN} [\mu, \sigma_\epsilon^2]$ denotes an independent Normal distribution with mean μ and variance σ_ϵ^2 . Based on a sub-sample $t = 1, \dots, T_k < T$ the least-squares estimator of β denoted $\tilde{\beta}_{T_k}$ is:

$$\tilde{\beta}_{T_k} = \frac{1}{T_k} \sum_{t=1}^{T_k} y_t \sim \text{N} [\beta, T_k^{-1} \sigma_\epsilon^2] \quad (2)$$

The precision of estimation will be higher for larger T , so is improved by later data (i.e., future data relative to T_k). Importantly, under the above assumptions, all sub-samples deliver unbiased estimates of β , which is the parameter relevant to all the time periods analyzed.

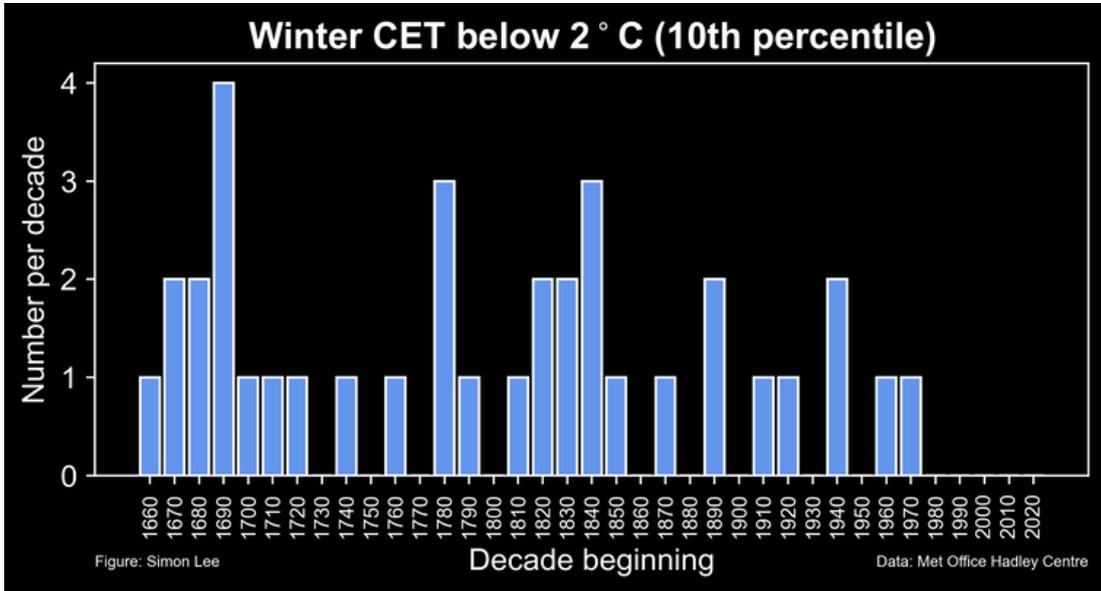


Figure 6: Frequency of Central England Winter Temperatures below 2° per decade. Data from Met Office Hadley Centre via <https://www.metoffice.gov.uk/hadobs/hadcet/>.

Another often implicit assumption is that the pre-specified model is the DGP, so ‘optimal estimation’ is the main issue, rather than the more realistic need to discover a reasonable approximation to that DGP, in which case modelling is essential. If most time-series are non-stationary from unanticipated shifts, then there is little likelihood of any pre-specified model being complete, correct and immutable, so finding those shifts is important, a topic we return to in Section 5.

The impacts of climate change show up in Winter Central England Temperatures (CET) and in seasonal temperature trends. Building on data by Parker, Legg, and Folland (1992) and Manley (1974), Figure 6 shows the frequency of Central England Winter Temperatures below 2° per decade from 1659–2019, recorded in a figure by Simon Lee.⁵

Based on the same data as above for CET, and comparing to those for Boston over 1743–2015, Figure 7 shows that temperature trends across seasons are very marked: both figures are from Hillebrand and Proietti (2017).⁶ In Central England, winters are warming but summers are not, whereas in Boston, warming is similar across all months. These long-term temperature trends are due to climate change, not weather variation, of which there is plenty. Temperature distributions are shifting on both sides of the Atlantic.

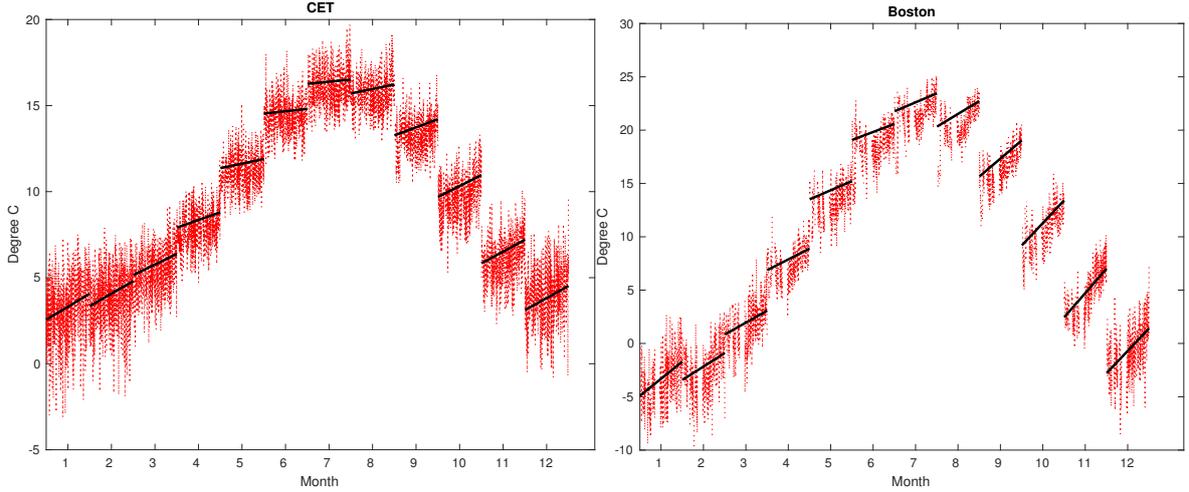


Figure 7: Trends across seasons. Left: CET over 1659–2019; Right: Boston over 1743–2015.

4 The key implications of shifting distributions

Given the manifest evidence of non-stationarity visible in the above graphs, we turn to consider how that modifies the implications of (1) and (2). If instead of a constant mean, the intercept was β_1 for $t = 1, \dots, T_1$ and $\beta_2 \neq \beta_1$ for $t = T_1 + 1, \dots, T$, then for $T_k = T_1$ the estimator $\tilde{\beta}_{T_1}$ in (2) would unbiasedly estimate β_1 , which is the relevant intercept at that time, whereas the full-sample $\hat{\beta}_T$ when $T_k = T$ would be:

$$\hat{\beta}_T = \frac{1}{T} \sum_{t=1}^T y_t \sim \text{N} [r_1 \beta_1 + (1 - r_1) \beta_2, T^{-1} \sigma_\epsilon^2] \quad (3)$$

where $r_1 = T^{-1} T_1$. Thus, $\hat{\beta}_T$ is relevant neither historically nor at the sample end. Moreover, forecasts based on $\hat{\beta}_T$ will be systematically biased:

$$\text{E} [y_{T+1} - \hat{y}_{T+1|T}] = \text{E} [\beta_2 + \epsilon_{T+1} - \hat{\beta}_T] = r_1 (\beta_2 - \beta_1) \quad (4)$$

Thus, any policy using either $\tilde{\beta}_{T_k}$ or $\hat{\beta}_T$ would have incorrect implications at T or later so it is essential to detect such shifts and reformulate empirical models accordingly. Examples highlighting how misleading empirical models can be when large outliers or location shifts are not modelled include Hendry and

⁵<https://theconversation.com/climate-change-is-making-extreme-cold-much-less-likely-despite-the-uk-plummeting-to-23-c-155177>

⁶The Boston time series is from NOAA's National Centers for Environmental Information Global Historical Climatology Network (GHCN) Station Number 42572509000.

Mizon (2011) showing a positive price elasticity for the demand for food in the USA, and Castle and Hendry (2014) revealing the key role of a long-run mean shift in models of UK real wages.

Not only will estimates of coefficients in models be non-constant, so biased for both the before and after shift parameters, as noted above, the basic statistical tools of conditional expectations and the law of iterated expectations also fail.

When the underlying distributions shift, expectations operators need to denote not only the random variables under analysis and their time, but also their distributions at that time and the information set being conditioned on. This requires three-way time dating as in $E_{D_{y_t}}[y_{t+1} | \mathcal{I}_{t-1}]$, which denotes the conditional expectation using the information set \mathcal{I}_{t-1} formed at time t of the vector random variable y_{t+1} integrated over $D_{y_t}(\cdot)$:

$$E_{D_{y_t}}[y_{t+1} | \mathcal{I}_{t-1}] = \int y_{t+1} D_{y_t}(y_{t+1} | \mathcal{I}_{t-1}) dy_{t+1} \quad (5)$$

As knowledge of $D_{y_{t+1}}(\cdot)$ is unavailable at t , the fundamental problem is obvious from (5): the expectation is not over $D_{y_{t+1}}(\cdot)$, so if $D_{y_t}(\cdot) \neq D_{y_{t+1}}(\cdot)$ because the distribution has shifted, there is no reason why $E_{D_{y_t}}[y_{t+1} | \mathcal{I}_{t-1}]$ should be informative about $E_{D_{y_{t+1}}}[y_{t+1} | \mathcal{I}_{t-1}]$, the correct conditional mean as Figure 8 illustrates.

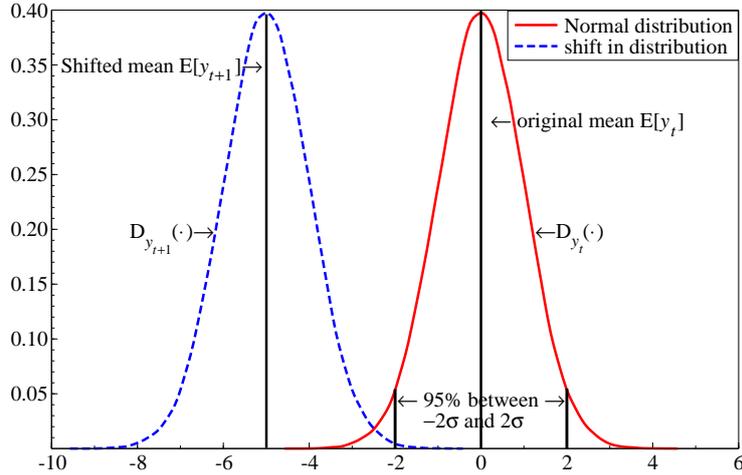


Figure 8: Illustrating the impact of a distributional shift on expectations.

Simply using E to denote the expectation can potentially mislead as in letting:

$$y_{t+1} = E[y_{t+1} | \mathcal{I}_t] + v_{t+1} \quad (6)$$

so that by taking conditional expectations of both sides:

$$E[v_{t+1} | \mathcal{I}_t] = 0 \quad (7)$$

This merely establishes that at time t , it is expected that the next error will have a mean of zero, but does not prove that the model used for $E[y_{t+1} | \mathcal{I}_t]$ will produce an unbiased prediction of y_{t+1} , as it is sometimes misinterpreted as doing.

Instead, when $y_t \sim \text{IN}[\mu_t, \sigma_y^2]$ is unpredictable because future changes in μ_t are unknowable (like SARS-Cov-2 virus driving the pandemic outbreak), the expectation $E_{D_{y_t}}[y_{t+1}]$ need not be unbiased for the mean outcome at $t + 1$. From (5):

$$E_{D_{y_t}}[y_{t+1}] = \int y_{t+1} D_{y_t}(y_{t+1}) dy_{t+1} = \mu_t \text{ (say)} \quad (8)$$

whereas:

$$\mathbb{E}_{D_{y_{t+1}}}[y_{t+1}] = \int y_{t+1} D_{y_{t+1}}(y_{t+1}) dy_{t+1} = \mu_{t+1} \quad (9)$$

so that $\mathbb{E}_{D_{y_t}}[y_{t+1}]$ does not correctly predict $\mu_{t+1} \neq \mu_t$. Thus, the expectation $\mathbb{E}_{D_{y_t}}[y_{t+1}]$ formed at t is not an unbiased predictor of the outcome μ_{t+1} at $t + 1$, although the ‘crystal-ball’ predictor $\mathbb{E}_{D_{y_{t+1}}}[y_{t+1}]$ based on knowing $D_{y_{t+1}}$ would be unbiased.

Returning to (6) at time t and subscripting the expectations operator as in (5):

$$y_{t+1} = \mathbb{E}_{D_{y_t}}[y_{t+1} | \mathcal{I}_t] + v_{t+1} \quad (10)$$

so (7) becomes:

$$\mathbb{E}_{D_{y_t}}[v_{t+1} | \mathcal{I}_t] = 0 \quad (11)$$

which does not entail that:

$$\mathbb{E}_{D_{y_{t+1}}}[v_{t+1} | \mathcal{I}_t] = 0 \quad (12)$$

whereas (12) is required for an unbiased prediction. The conditional expectation is the minimum mean-square error predictor only when the distribution remains constant, and fails under distributional shifts.

Moreover, the law of iterated expectations only holds inter-temporally when the distributions involved remain the same. When the variables correspond to drawings at different dates drawn from the same distribution so $D_{y_t} = D_{y_{t+1}}$:

$$\mathbb{E}_{D_{y_t}}[\mathbb{E}_{D_{y_{t+1}}}[y_{t+1} | y_t]] = \mathbb{E}_{D_{y_{t+1}}}[y_{t+1}]. \quad (13)$$

Thus, if the distributions remain constant, the law of iterated expectations holds, but it need not hold when distributions shift:

$$\mathbb{E}_{D_{y_t}}[\mathbb{E}_{D_{y_{t+1}}}[y_{t+1} | y_t]] \neq \mathbb{E}_{D_{y_{t+1}}}[y_{t+1}] \quad (14)$$

as $D_{y_{t+1}}(y_{t+1}|y_t) D_{y_t}(y_t) \neq D_{y_{t+1}}(y_{t+1}|y_t) D_{y_{t+1}}(y_t)$ unlike the situation in (13) where there is no shift in distribution.

Changes that alter the means of the data distributions are location shifts, so DGPs with such shifts are obviously non-stationary. Unfortunately there is a widespread use of ‘non-stationary’ to refer just to DGPs with stochastic trends, leading to the non-sequitur that ‘differencing induces stationarity’, as well as our need to call time series with shifts and possibly also stochastic trends ‘wide-sense non-stationary’. Further, stochastic trends are usually assumed to apply constantly to an entire sample, but unit roots in DGPs are not an intrinsic property and can also change. Not only are there strong trends in Figure 1, these vary considerably over time as shown by their changing rates of growth. Thus, several of the differenced time series are not stationary, as shown by Figure 9. In particular, changes in atmospheric CO₂ have continued to trend up despite the Paris Accord, and there have been large variance changes in the differenced series for UK coal use. As documented in Castle and Hendry (2020), the distributions of the UK’s CO₂ emissions have shifted considerably over time.

Both the ‘Financial crisis’ of 2008 and the Sars-Cov-2 pandemic in 2020 were essentially unanticipated, but have had major impacts on economic activity as well as on health, lives and livelihoods, so the importance of non-stationarity replacing stationarity as the standard assumption for time-series analysis may at last be realised. However, once the comfort blanket of stationarity is discarded, a fundamental problem is that there may be multiple unknown shifts in many different facets of DGPs at different points in time, with differing magnitudes and signs. Many ‘proofs’ and statistical derivations of time-series modelling become otiose, and some like the so-called ‘Oracle principle’ for model-selection algorithms are irrelevant when parameters change over relatively short periods of time.

Since climate change is driven by economic activity—which is wide-sense non-stationary and riddled with abrupt, usually unanticipated, changes—these difficulties confront empirical modelling of many observational climate time series. The next section addresses why undertaking econometric modelling of changing climate time series requires handling various forms of shift.

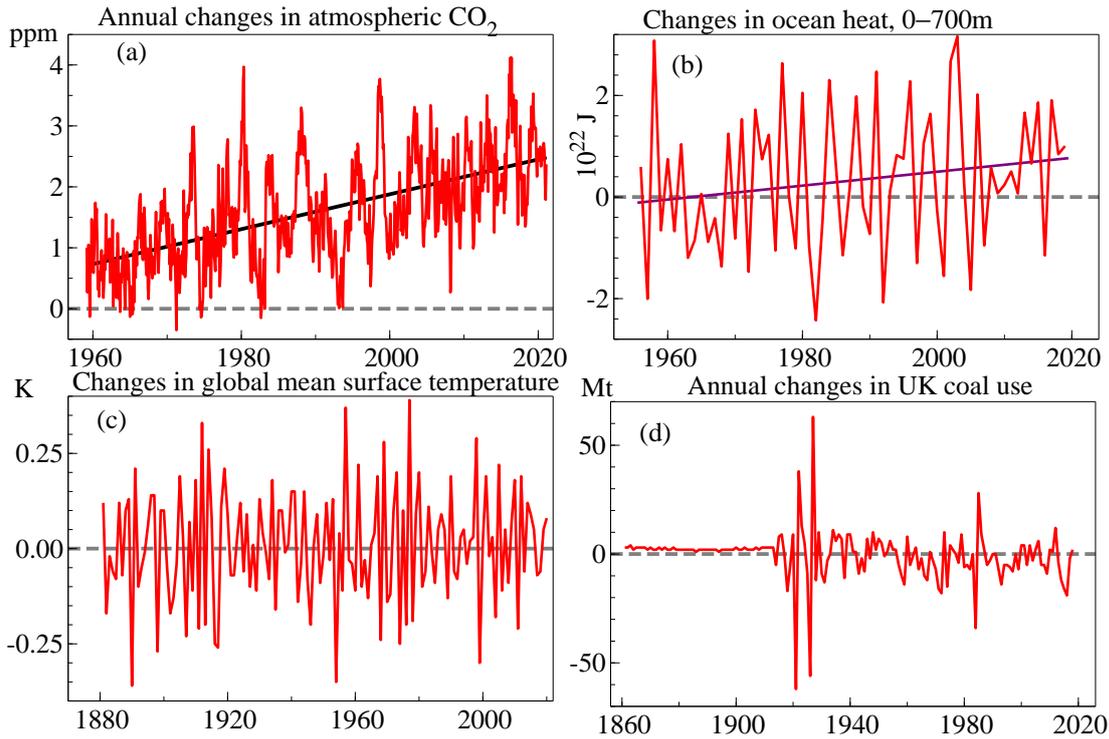


Figure 9: (a) Monthly changes in atmospheric CO₂; (b) annual changes in global ocean heat; (c) annual changes in global mean surface temperature; (d) annual changes in UK coal use.

5 The importance of handling shifts

There are many methods designed either to reveal parameter non-constancy by estimation or testing, or tackle the problem by formulating changing-parameter models. The first group uses recursive data samples or moving windows thereof. In the recursive setting, T_k in (2) defines an initial sub-sample, and observations are added sequentially to provide $\tilde{\beta}_{T_k+j}$ for $j = 1, \dots, T - T_k$. We reinterpret this procedure as initially including impulse indicators 1_t that are zero at all observations except unity at t for $t = T_k + 1, \dots, T$, then sequentially removing these as the estimation sub-sample size increases. The values of these impulse indicators can be very informative both about outliers and location shifts in the remaining data but are usually ignored. Moreover, there may have been shifts in the initial sample $1, \dots, T_k$, which will not be detected, so the initial estimate is already biased. Similar comments apply to using moving windows.

A different issue confronts testing for non-constancy in that the model must already be available, either specified a priori or more likely selected from some set of trial runs. But these estimates ignored the changes that are then tested for, and if constancy is rejected, the model must be reformulated—vitiating the earlier specifications. A similar difficulty faces methods for estimating changing parameters after fitting a pre-specified model: not only are there all the other unknown aspects of non-stationarity noted above, observational-data model specification is also uncertain in terms of the relevant variables, their lags and non-linearities.

To illustrate the issues raised by location shifts, we generalise (1) to the dynamic model:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \epsilon_t, \quad t = 1, \dots, T \quad (15)$$

where $T = 100$, $\beta_1 = 0.5$, with $\sigma_\epsilon^2 = 0.1$, but $\beta_0 = 1$ till $T = 80$ then $\beta_0 = 0$ for the remainder of the sample. A draw from (15) is shown in Figure 10, Panel (a). The average simulation full-sample

estimates from $M = 1000$ replications were:

$$\hat{y}_t = \begin{matrix} 0.914 & y_{t-1} & + & 0.119 & & \hat{\sigma}_\epsilon^2 = 0.14 \\ (0.045) & & & (0.082) & & \end{matrix} \quad (16)$$

The estimates in (16) are far from the DGP parameter values. The estimate of β_1 being close to unity is a standard outcome from a failure to model a step shift, and warns that near estimated ‘unit roots’ need not signal a stochastic trend rather than a location shift. Also, $\hat{\sigma}_\epsilon^2$ is 40% larger than σ_ϵ^2 on average.

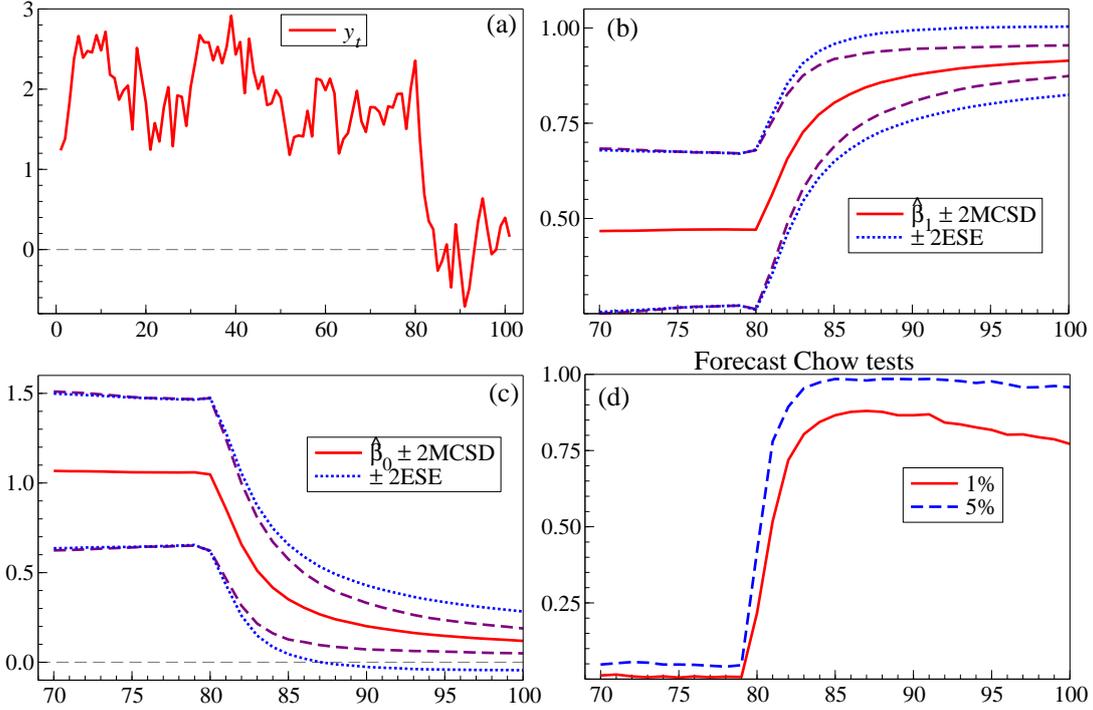


Figure 10: (a) A time series from (15); (b), (c) recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs); (d) forecast Chow tests.

Providing the DGP is known, recursive estimation reveals the non-constant parameter estimates as Figure 10, Panels (b) & (c) show, confirmed by high levels of rejection of the null of constancy on forecast Chow (1960) tests (see Hendry, 1984, for an analysis of Monte Carlo methods). By themselves the recursive plots show that there was change, but do not isolate the source. One ‘identification’ route is to fit the model separately to the data before and after $T = 80$. For a single data draw, this delivers:

$$\hat{y}_t = \begin{matrix} 0.661 & y_{t-1} & + & 0.68 & & \hat{\sigma}_\epsilon^2 = 0.10 \\ (0.084) & & & (0.168) & & \end{matrix} \quad t = 2, \dots, 80 \quad (17)$$

$$\hat{y}_t = \begin{matrix} 0.521 & y_{t-1} & + & 0.009 & & \hat{\sigma}_\epsilon^2 = 0.10 \\ (0.105) & & & (0.073) & & \end{matrix} \quad t = 81, \dots, 100 \quad (18)$$

revealing the large shift in the intercept with little change in the other estimates. Our methods, here step-indicator saturation (SIS) explained below, are designed to find location shifts and outliers while selecting relevant variables. Applying SIS to the full sample, where $S_{\{j\}} = 1, t \leq j, j = 1, \dots, T - 1$, and zero otherwise, but always retaining the intercept and y_{t-1} selecting which step indicators to retain at 0.5% yields:

$$\hat{y}_t = \begin{matrix} 0.608 & y_{t-1} & - & 0.011 & + & 0.795 & S_{\{80\}} & & \hat{\sigma}_\epsilon^2 = 0.10 \\ (0.066) & & & (0.072) & & (0.138) & & & \end{matrix} \quad t = 2, \dots, 100 \quad (19)$$

Thus, the intercept is estimated as 0.784 up to $t = 80$ and essentially zero thereafter. Despite having started with a candidate variable set of 98 step indicators, all the data (less two degrees of freedom) can be used in (19) to estimate β_1 , thereby delivering a more precise, as well as a constant, outcome.

5.1 Simulating SIS for (15) with an intercept shift

The results in (19) are for a single draw of the time series, but Autometrics can simulate SIS. Using the same settings as for Figure 10 and selecting step indicators at 0.5% with no backtesting until the final stage, yields the estimates reported in Figure 11.

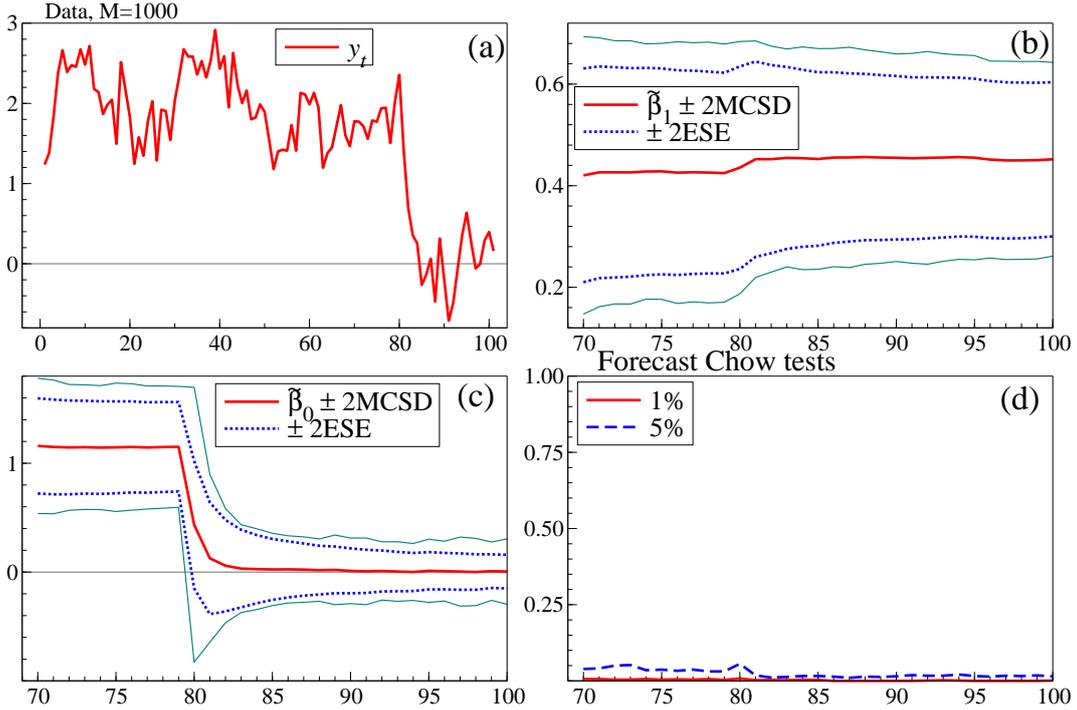


Figure 11: (a) One time series from (15); (b), (c) simulated recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs) when applying SIS; (d) resulting forecast Chow tests.

The outcome is dramatically different from that in Figure 10. The estimate of β_1 is now relatively constant at between 0.4 and 0.5; the estimate of β_0 switches quickly from around unity to zero after $t = 80$; and the forecast tests reject at about their nominal significance levels. The step indicator at $t = 80$ was selected with a probability of 0.84 rising to 0.97 within ± 1 of $t = 80$. Irrelevant indicators were selected with probability 0.008 (the empirical gauge), close to the nominal significance level of 0.005, so there is almost no over-fitting. The MCSDs are somewhat wider than the ESEs, so the latter slightly underestimate the uncertainty.

5.2 Simulating SIS for (15) with a shift in dynamics

A legitimate question is what if, instead of β_0 shifting, the intercept remained constant and the dynamic parameter β_1 changed? For example:

$$\begin{aligned} y_t &= \beta_0 + \beta_1 y_{t-1} + \epsilon_t, \quad t = 1, \dots, T_1 \\ y_t &= \beta_0 + \beta_1^* y_{t-1} + \epsilon_t, \quad t = T_1 + 1, \dots, T \end{aligned} \quad (20)$$

Setting the same parameter values for β_0 , β_1 , σ_ϵ and $T_1 = 80$ as before with $\beta_1^* = 0.75$, we repeated the above Monte Carlo leading to Figure 12.

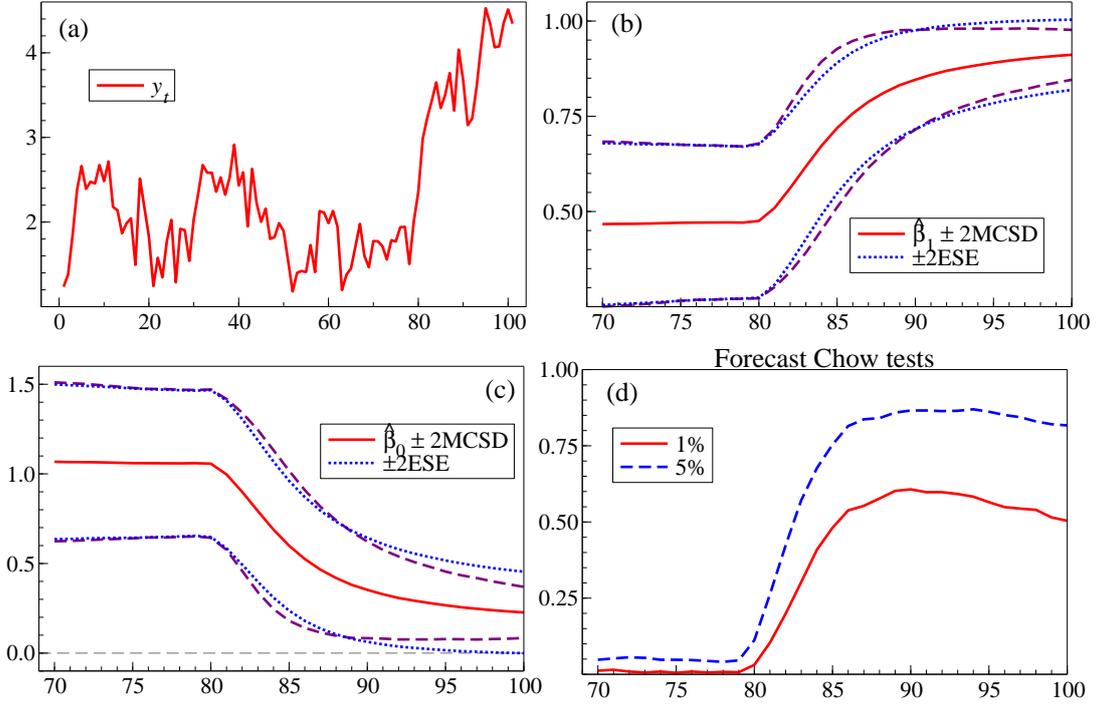


Figure 12: (a) A time series from (20); (b), (c) recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs); (d) forecast Chow tests.

Apart from Panel (a), the other panels are similar to those in Figure 10. The average simulation full-sample estimates from $M = 1000$ replications are now:

$$\hat{y}_t = \begin{matrix} 0.911 & 0.227 \\ (0.046) & (0.113) \end{matrix} y_{t-1} + \quad (21)$$

so are also closely similar to (16). Indeed, for the single data draw, the sub-sample fits are identical to (17) for the first period but for the second sub-period:

$$\hat{y}_t = \begin{matrix} 0.678 & 1.28; \\ (0.125) & (0.466) \end{matrix} y_{t-1} + \quad t = 81, \dots, 100, \quad \hat{\sigma}_\epsilon^2 = 0.10 \quad (22)$$

so the estimate of β_1 is barely altered but $\hat{\beta}_0$ is greatly shifted. Consequently, it may not surprise that SIS can again capture much of the shift in the intercept:

$$\hat{y}_t = \begin{matrix} 0.71 & 1.14 & - & 0.568 \\ (0.06) & (0.22) & & (0.119) \end{matrix} y_{t-1} + S_{80}; \quad t = 2, \dots, 100, \quad \hat{\sigma}_\epsilon^2 = 0.10 \quad (23)$$

The simulation outcomes when applying SIS reinforce these conclusions. First Figure 13 shows that the estimates of β_1 increase after $t = 80$ although they remain close to 0.5 rather 0.75. As before, the estimates of β_0 change after $t = 80$ gradually rising to 2, and the forecast Chow tests reject close to their nominal significance levels. The step indicator at $t = 80$ was selected with a probability of 0.46, rising to 0.73 within ± 1 of $t = 80$, 0.89 within ± 2 and 1.0 within ± 3 of $t = 80$, so detection is more spread out. Irrelevant indicators were selected with probability 0.010, somewhat above the nominal significance level of 0.005 reflecting the issue being a shift in dynamics of which only the induced location shift is detected. The MCSDs are wider than the ESEs, so the latter again underestimate the uncertainty.

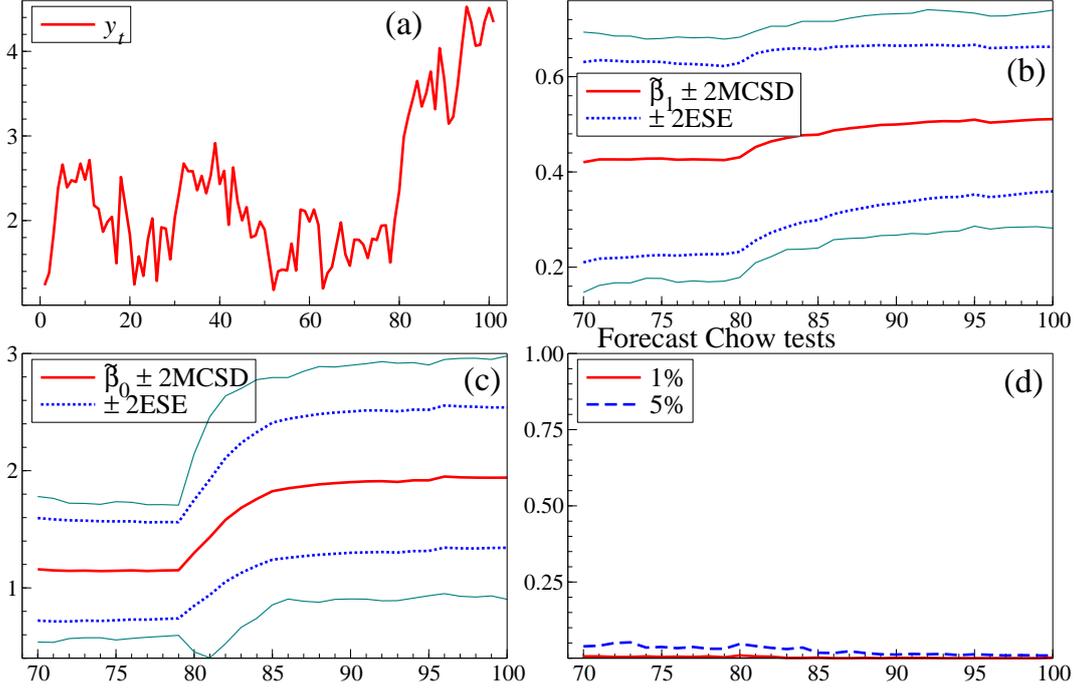


Figure 13: (a) A time series from (20); (b), (c) simulated recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs) when applying SIS; (d) resulting forecast Chow tests.

5.3 Why is SIS effective for a shift in dynamics?

It may seem puzzling that SIS can help facing a shift in the dynamics, but the crucial modelling problems are those that arise from shifts in the long-run, or equilibrium, mean rather than in other parameters, as emphasized in the forecasting context by Clements and Hendry (1998,1999). Indeed, notice that the dominant visual feature of both data graphs above are the sudden departures from their previous average locations. Let $\mu = \beta_0/(1 - \beta_1)$ which shifts to $\mu_{\beta_0^*}^* = \beta_0^*/(1 - \beta_1)$ in (15), so goes from 2 to 0 as seen in Figure 10 Panel (a). Then $\mu_{\beta_1^*}^* = \beta_0/(1 - \beta_1^*)$ in (20) changes from 2 to 4 as in Figure (12) Panel (a). Defining $\nabla\mu^* = \mu^* - \mu$ and $\nabla\beta_1^* = \beta_1^* - \beta_1$, the two cases can be written as:

$$y_t - \mu = \beta_1 (y_{t-1} - \mu) + (1 - \beta_1) \nabla\mu_{\beta_0^*}^* \mathbf{1}_{\{t>T_1\}} + \epsilon_t \quad (24)$$

$$y_t - \mu = \beta_1 (y_{t-1} - \mu) + (1 - \beta_1^*) \nabla\mu_{\beta_0^*}^* \mathbf{1}_{\{t>T_1\}} + \nabla\beta_1^* (y_{t-1} - \mu) \mathbf{1}_{\{t>T_1\}} + \epsilon_t \quad (25)$$

Expressed as mean-zero deviations about their equilibrium means, when either DGP shifts, the problem is the sudden appearance of a non-zero intercept where previously it was zero. That is what SIS can correct. For the β_0 shift, removing the new intercept recreates the previous DGP, whereas the β_1 shift also induces a third term that has a mean of zero, so SIS cannot correct that. From (24) and (25), predicted values for the step-indicator magnitudes can be calculated, leading to -1 and 0.5 respectively. For example, from (19), $\mu = 2$, so S_{80} should have a coefficient of unity, not significantly different from the 0.8 found, whereas in (23), it should be -0.5 , close to the outcome.

While this illustration is for a simple setting, the principles generalise to more general models including selecting across variables as well as indicators.

5.4 Tackling a shift in dynamics by multiplicative indicator saturation

Although SIS offsets much of the effect of the changing dynamics, the more appropriate approach is to model the change in β_1 . Our tool for doing so is multiplicative indicator saturation, denoted MIS (see e.g., Kitov and Tabor, 2015, and Castle, Doornik, and Hendry, 2020). In MIS, y_{t-1} is multiplied by almost every step indicator creating the $T - 2$ additional candidate variables $S_j \times y_{t-1}$ for $j = 2, \dots, T - 1$. Thus, SIS is in fact MIS for the constant term, but is a special case of importance given the pernicious effect unmodelled location shifts have on forecasts.

For one draw applying selection at 0.1% over the MIS regressors we obtain:

$$\hat{y}_t = \begin{matrix} 0.74 & 1.02 & - & 0.25 & S_{80} \times y_{t-1} & t = 2, \dots, 100, & \hat{\sigma}_\varepsilon^2 = 0.08 \\ (0.05) & (0.16) & & (0.04) & & & \end{matrix} \quad (26)$$

which almost exactly replicates the DGP. The outcome shows a shift in $\hat{\beta}_1$ from 0.49 before $t = 80$ to 0.74 after, with a constant intercept. The simulation outcomes are reported in figure 14. The interaction indicator at $t = 80$ was selected with a probability of 0.42, rising to 0.62 within ± 1 of $t = 80$, 0.81 within ± 2 and 0.89 within ± 3 of $t = 80$. Irrelevant indicators were selected with probability 0.007, which falls to 0.003 excluding the indicators within ± 3 of $t = 80$. The MCSDs are again wider than the ESEs, so the latter underestimate the uncertainty.

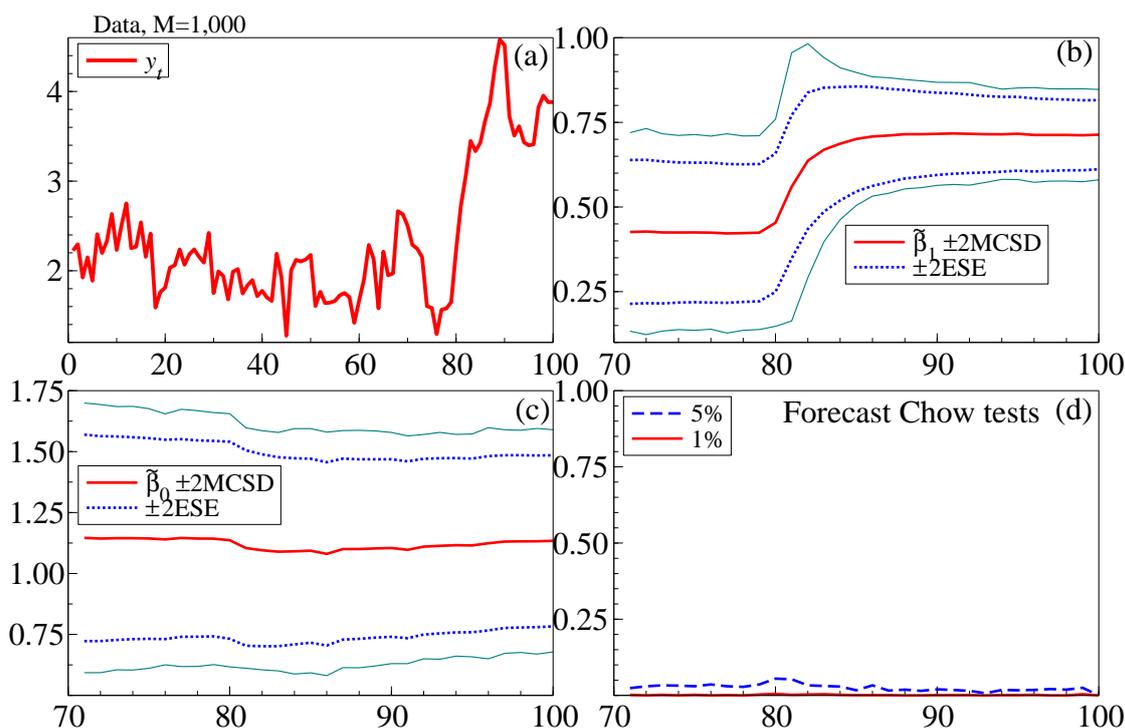


Figure 14: (a) A time series from (20); (b), (c) simulated recursive parameter estimates with estimated standard errors (ESEs) and Monte Carlo standard deviations (MCSDs) when applying MIS; (d) resulting forecast Chow tests.

5.5 Tackling changing trends by trend-indicator saturation

As noted in the introduction, if humanity is to avoid catastrophic climate change, the present upward trends in GHG emissions must be reversed to become rapid downward trends. Thus, graphs of fossil fuel use and GHG emissions shaped like the \cap in Figure 1(d) may become common, as may less extreme

but still marked trend changes like those in Figure 1(a)–(c). Economics, demography and epidemiology have all experienced sudden large changes in trend as Figure 15 illustrates.

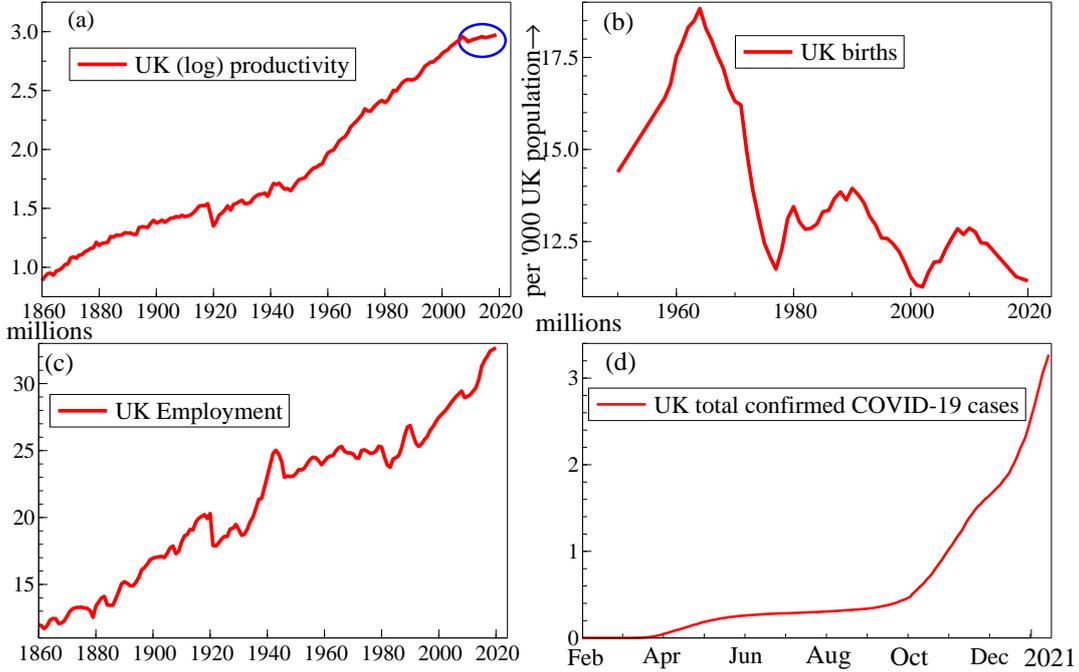


Figure 15: UK time series of (a) output per worker per year (productivity); (b) births per thousand of the population; (c) employment; (d) cumulative confirmed COVID-19 cases.

Productivity has increased at varying rates since 1860, but has stagnated since 2008 as highlighted by the ellipse, badly over-forecast by the UK Office of Budget Responsibility not adjusting to the ‘flat-lining’, and much improved by the robust predictor proposed by Martinez, Castle, and Hendry (2021). The birth rate was steadily increasing until the introduction of oral contraception in the mid-1960s, then fell sharply till the late 1970s and has fluctuated since. Employment has expanded greatly since 1860 with major fluctuations during world wars and severe depressions, but at a much more rapid rate since 2000 (which helps explain the productivity slowdown). Finally, panel (d) shows dramatic changes in the trend of confirmed COVID-19 cases. To empirically model trends which change by unknown magnitudes at unknown points in time an unknown number of times requires a general tool like trend indicator saturation, denoted TIS (see Castle, Doornik, Hendry, and Pretis, 2019, and Walker, Pretis, Powell-Smith, and Goldacre, 2019, for an application). In effect, TIS is MIS for the trend term but merits a separate analysis because like the constant, a trend is a deterministic term and also that $\sum_{t=1}^T t^2 = \frac{1}{6}T(T+1)(2T+1)$ so grows at $O(T^3)$ as against $O(T^2)$ for sums of squares of stationary variables.

Much macroeconomic modelling is in differences of variables or their logs to eliminate trends and possibly unit roots. However, the potency of detecting breaks is much higher when working in levels than in changes. To illustrate, we create an artificial trend-break DGP and apply TIS to the levels of the time series and compare break detection to that by SIS applied to the changes.

Figure 16 records a single representative draw of the resulting time series denoted y_t , z_t , $\Delta y_t = y_t - y_{t-1}$ and Δz_t where:

$$y_t = \beta_{0,t} + \beta_{1,t}z_t + \epsilon_t \text{ where } \epsilon_t \sim \text{IN}[0, \sigma_\epsilon^2] \quad (27)$$

and:

$$z_t = \gamma_0 + \gamma_1 t + \nu_t \text{ where } \nu_t \sim \text{IN}[0, \sigma_\nu^2] \quad (28)$$

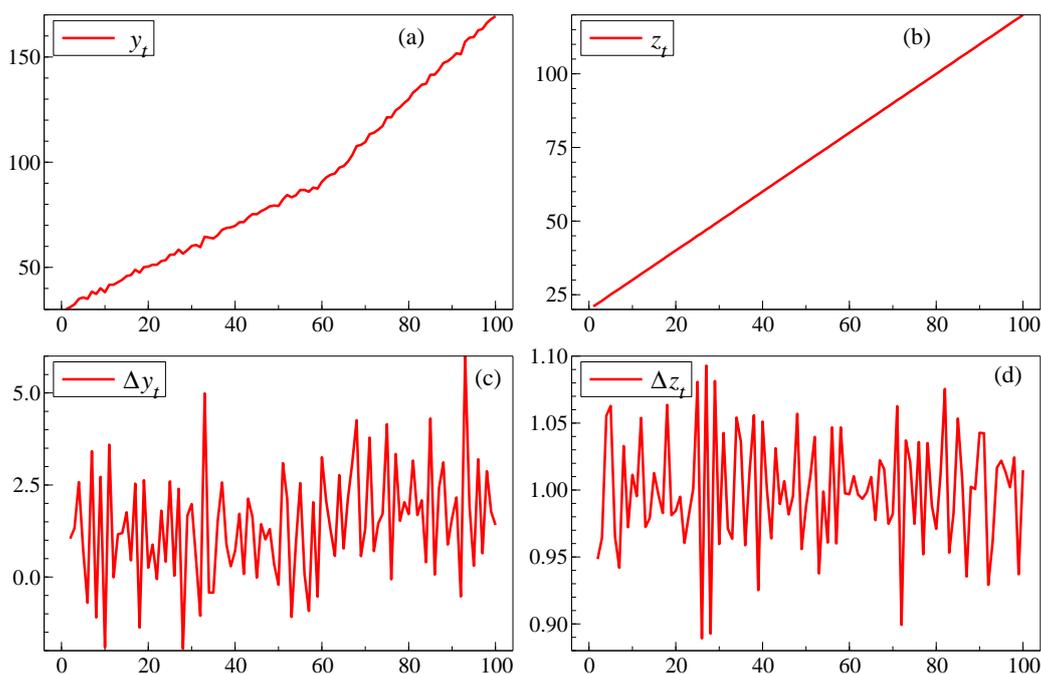


Figure 16: Time series from (27) and (28) of (a) y_t ; (b) z_t ; (c) Δy_t ; (d) Δz_t .

so z_t acts as the trend. We set $\beta_{0,t} = 10$ and $\beta_{1,t} = 1$ for $t = 1, \dots, 60$ then $\beta_{0,t} = -70$ and $\beta_{1,t} = 2$ for $t = 61, \dots, 100$ with $\gamma_0 = 0$, $\gamma_1 = 1$, $\sigma_\epsilon^2 = 1$ and $\sigma_\nu^2 = 0.001$ deliberately set to a tiny value to mimic a trend as Figure 16 confirms.

‘Ocular econometrics’ on Figure 16 shows an obvious trend break, but not an obvious shift in the changes Δy_t . Applying TIS to the levels with a target nominal significance value of 0.001 picks up the shift, but SIS applied to the changes needs to be at 0.005 to do so, as we now show. Although doubling of the growth rate is a very large change, smaller changes, such as a 50% increase, barely registered in Δy_t . We consider 4 cases: (i) estimating the relation between y_t and z_t without allowing for the trend shift; (ii) estimating that relation with TIS at 0.001; (iii) regressing $\Delta y_t = \psi_0 + \psi_1 \Delta z_t$; and (iv) estimating that last equation with SIS. Figure 17 shows the outcomes for (i) & (ii), and Figure 18 for the differenced case (iii) & (iv).

Not handling the break in the trend as in Figure 17(a), (b) is disastrous—the deviations from trend would be unrelated to ‘excess demand’ if y_t was GDP—whereas in (c), (d) the shift is detected by TIS.⁷

Figure 18 records two cases with SIS in the top row (a), (b) when Δz_t is not forced, and when it is (i.e., not selected over) in (e), (f), whereas (c), (d) in the middle row are with TIS when z_t is forced. SIS at 0.005 detects the shift, dropping Δz_t (when only the constant is forced), whereas Δz_t is retained with an unrestricted constant. An intercept is equally good here at representing the mean change, but would not work if z_t was exactly the trend, so Δz_t was constant whereas Δz_t would matter more than the intercept if σ_ν was larger. There is a small loss of fit if SIS is not used, and as the residual autocorrelation test is significant in all three cases, there is little sign of the step shift. Differencing doubles the error variance so detecting the step shift rather than the trend break is harder, reflected in the need to use a less stringent nominal significance level.

Although the fit is similar between (c) and (d) as measured by $\hat{\sigma}_\epsilon$, there is nevertheless a substantial impact on forecasts of y_t derived from Δy_t as Figure 19 shows for multi-step forecasts. There is little difference between (iii) & (iv) in the root mean square forecast errors (RMSFEs) for forecasts of the

⁷The shift did not quite match the node at the break point so is picked up by the implicit difference of the indicators at $t = 58$: the timing in PcNaive starts at 0, so dates are shifted back, and the trend indicators end at the date shown.

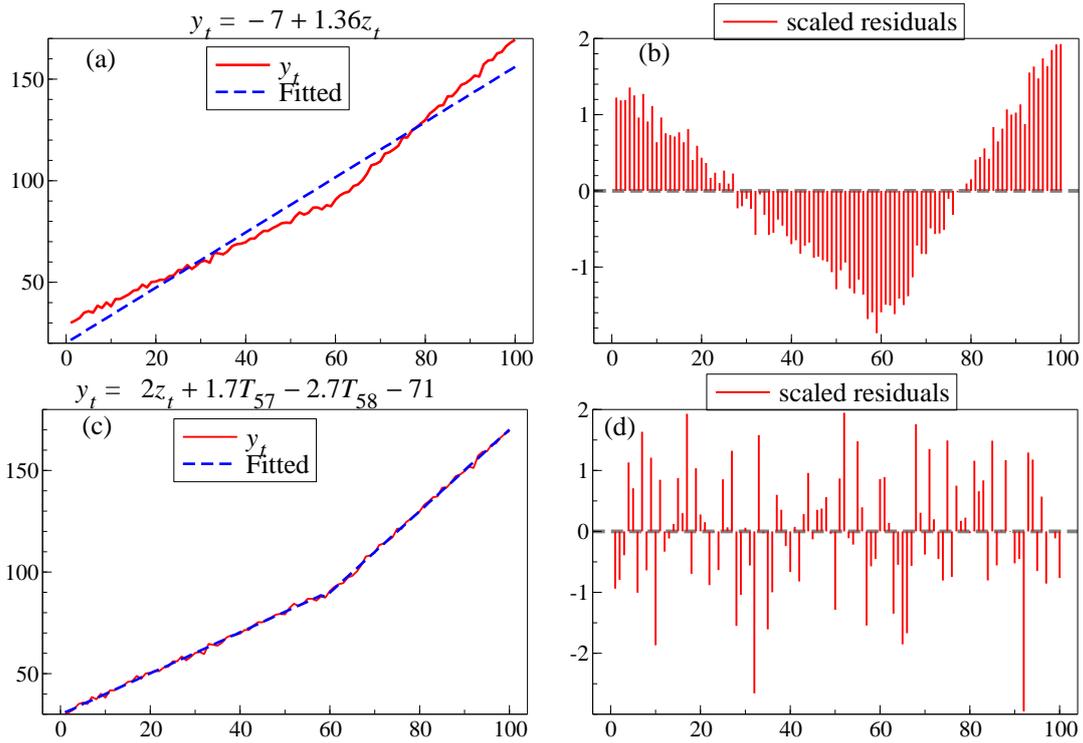


Figure 17: Actual and fitted values and residuals from (a), (b) regressing y_t on z_t without TIS and (c), (d) with.

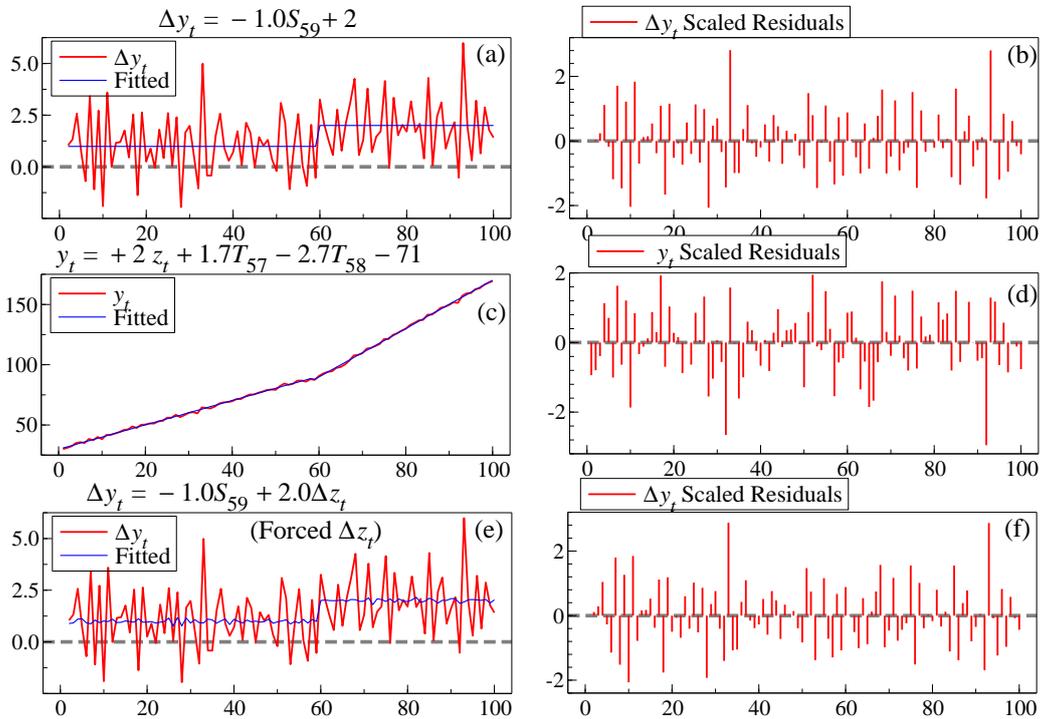


Figure 18: Actual and fitted values and residuals from selecting: (a), (b) Δy_t from Δz_t with SIS; (c), (d) y_t with TIS when z_t is forced; (e), (f) Δy_t on SIS with Δz_t forced.

changes Δy_t , but cumulating these forecasts to derive the levels' forecasts leads to the RMSFE for the model without SIS being more than four times larger than that with, and those forecasts from (iii) being

systematically too low.

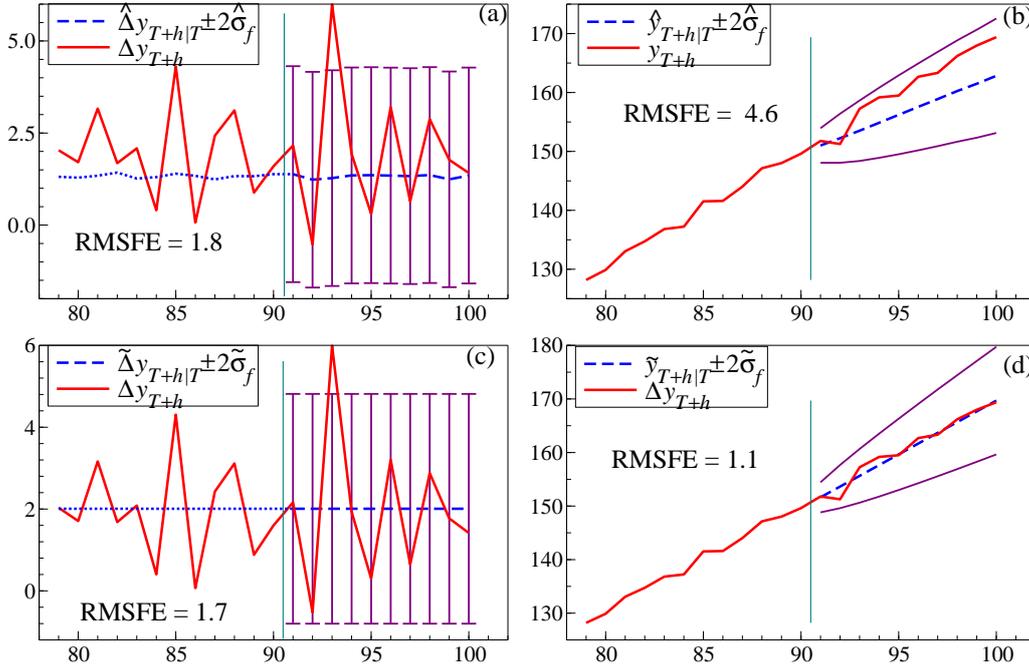


Figure 19: Actual and forecast values for changes and derived levels from models of Δy_t on Δz_t without SIS (a), (b) denoted $(\hat{\cdot})$ and with (c), (d), $(\tilde{\cdot})$.

Figure 20 records multi-step forecast values for the levels from models of y_t on a forced constant and forced z_t with TIS (b) and without (a), as well as with a step intercept correction from $t = 84$ on. All three graphs also show robustified forecasts (see Hendry, 2006). The forecasts with TIS have dramatically smaller RMSFEs than those without. While the robust device is a great improvement for (a) when TIS was not used, being based on differencing the data, its RMSFEs are similar to those in Figure 19 for the cumulated differenced forecasts without SIS.

Models in differenced data face this risk of appearing to have few or no step shifts when estimated and produce respectable forecasts of future changes, but suffer systematic forecast failure for the entailed levels of the data.

6 Modelling and forecasting wide-sense non-stationary processes

In this section, we describe our approach to jointly tackling all the main problems facing analyses of empirical evidence on wide-sense non-stationary processes. The framework is one of model discovery, from model formulation, through selection, to evaluation. Formulation entails commencing from a very large initial specification intended to nest the DGP as closely as possible while retaining available theory information. Selection requires searching over all non-retained potential determinants jointly with indicator saturation estimation, to find which variables, lags, and functional forms are relevant and which observations need separate handling. Evaluation involves testing for a range of possible mis-specifications. Forecasting hinges on the wide-sense non-stationarity of the data, with forecasting methods derived for stationary settings otiose in the real world.

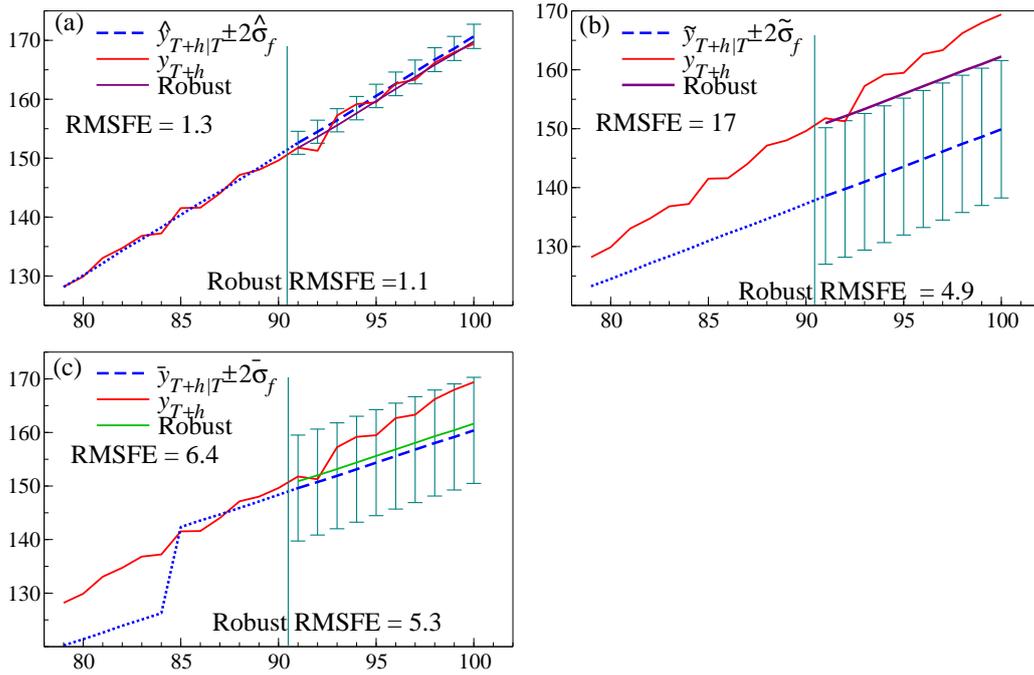


Figure 20: Actual and forecast values for levels from models of (a) y_t on a forced constant and forced z_t with TIS ($\hat{\cdot}$) and (b) without TIS ($\tilde{\cdot}$), and (c) with a step intercept correction from $t = 84$ ($\bar{\cdot}$).

6.1 Selecting models

The proposed framework begins by embedding the climate theory within a much more general model specification that allows for influences that are not explicitly included in the theory models. This could be due to the theory model itself being incorrect or incomplete, or it could be due to external effects that lie outside of the climate theory, for example volcanic eruptions or economic responses to a pandemic. The theory is the object of study, but the target for model selection is the data generating process for the set of variables under analysis. In order to evaluate the objective or the postulated theory, the model must be robust to outliers, shifts, omitted variables, non-linearity, mis-specified dynamics, incorrect distributions, non-stationarity, and invalid conditioning.

The approach can orthogonalize all additional variables with respect to the theory variables so the distributions of the estimators of parameters of the object are unaffected by selection, see Hendry and Johansen (2015). The selection algorithm retains without selection all the variables in theory model when selecting other features. This enables tighter than conventional significance levels, reducing the retention of irrelevant candidates without jeopardizing the retention of theory relevant variables. This is a win-win situation. When the theory model is complete and correct, the resulting model will deliver precisely the same estimates as directly fitting it to data, even if selecting from a vastly larger specification, but if the theory is not complete and correct the modeller would discover a better formulation.

The implication of commencing with much larger models than the theory would suggest is that the starting point for model selection would include more candidate variables, N , than the number of observations, T . A search algorithm is needed that can handle such a situation and we use *Autometrics* which implements a block search algorithm to handle more variables than observations, see Doornik (2009) and Hendry and Doornik (2014). The regressors are divided into sub-blocks but the theory variables are retained at every stage, selecting only over the putative irrelevant variables at a stringent significance level. The sub-blocks include expanding and contracting searches to handle correlated regressors that may need to enter jointly to be significant. It is almost costless to check large numbers of rival variables,

so there are huge benefits if the initial specification is incorrect and the enlarged general unrestricted model nests the local data generating process (LDGP).

Search is unavoidable as climate and economic variables are all interrelated with high correlations. There are many selection algorithms in the literature that assume a lack of correlation and therefore pursue single path searches. These include forward search and stepwise regression, 1-cut selection and backward elimination, and Lasso, see Tibshirani (1996), and its variants. The benefits of *Autometrics* include formal tests for congruence to ensure all tests of reduction are valid, multipath search to avoid path dependence and ensure that the initial ordering of regressors does not matter, increased efficiency relative to estimating all 2^N possible models, which is infeasible in most cases as N becomes even moderately large, and a well-defined stopping point at terminal models using encompassing tests against the general unrestricted model to ensure there isn't a substantial loss of information. Although such a multipath search is not as fast and simple as single path procedures, and there is some dependence on how the blocking is implemented, increased computing power mean that these costs are not large.

Three important automatic generalisations for specifying the general unrestricted model include (i) adding in many lags of the regressors to allow for a sequential factorization and the selection algorithm will then apply a lag length reduction stage ensuring there are no unmodelled dynamics; (ii) including non-linear transformations of the regressors to allow for general unspecified forms of non-linearity using polynomial and exponential expansions; and (iii) a variety of indicator saturation estimators (ISEs) to model many different aspects of wide-sense non-stationarity. As illustrated above, each ISE is designed to match a specific problem: IIS to tackle outliers, SIS for location shifts, MIS for parameter changes, TIS for trend breaks, as well as designed-indicator saturation for modelling phenomena with a regular pattern, such as detecting the impacts on temperature of volcanic eruptions (DIS: see Pretis, Schneider, Smerdon, and Hendry, 2016). Importantly, saturation estimators can be used in combination as we do below, where IIS and SIS combined is called super saturation (see Ericsson and Reisman, 2012, and Kurle, 2019). All saturation estimators can be applied when retaining without selection a theory-model that is the objective of a study, while selecting from other potentially substantive variables. Saturation estimators, and indeed our general approaches, have seen applications across a range of disciplines including dendrochronology, volcanology, geophysics, climatology, and health management, as well as economics, other social sciences and forecasting. Although theory models are much better in many of these areas than in economics and other social sciences, modelling observational data faces most of the same problems, which is why an econometric toolkit can help.

6.2 Forecasting in non-stationary worlds

'Conventional' economic forecasting uses a theory-based system that models the main variables of interest. Examples include a dynamic-stochastic general equilibrium (DSGE) model, a variant of a vector autoregression (VAR), or a simultaneous-equations model. Some systems are closed, in that all variables are modelled, but most are open with 'offline' assumptions made about future values of unmodelled 'exogenous' variables. Almost all economic systems are equilibrium-correction models (EqCMs) or differenced variants thereof. Clements and Hendry (1998, 1999) develop taxonomies of all forecast errors in closed systems to show that the key determinant of forecast failure is an unmodelled shift in the equilibrium mean. Hendry and Mizon (2012) extend the taxonomy to open systems and show that result still holds, but there are additional potential sources of forecast failure deriving from changes in the 'exogenous' variables. Introductions to forecasting facing breaks are provided by Castle, Clements, and Hendry (2016, 2019).

Nothing can solve the problem of an unanticipated shift in the equilibrium mean over the forecast horizon, of which this century has already witnessed several including the dot-com crash, the financial crisis and the Covid-Sars-2 pandemic. Forecasting such shifts before they occur has not yet proved feasible even if ex post claims that they were predicted abound. However, a number of approaches have been proposed to counter the problem of forecast failure due to in-sample shifts in EqCMs including differ-

encing, and developing predictors that are robust after breaks: see e.g., Hendry (2006), Castle, Clements, and Hendry (2015) and Martinez, Castle, and Hendry (2021) (described in Section 6.5). Nevertheless, avoiding systematic mis-forecasting has a cost in larger root-mean square forecast errors (RMSFEs), especially when there is no shift to offset.

The system we developed for forecasting Ice Ages data is a simultaneous-equations model where the measures of Earth’s orbital trajectory are entered as unmodelled ‘exogenous’ drivers, so in Section 6.3 we address the extent to which it circumvents the problems raised by Hendry and Mizon (2012) after discussing the role of the exogeneity assumptions in the Ice-Ages forecasts. We then describe our forecasting device Cardt, a modification of the predictor we used in the M4 competition, see Doornik, Castle, and Hendry (2020a). This will be applied to forecasts of UK productivity in Section 6.5 and forecasting UK CO₂ emissions in Section 7.2. Section 6.5 describes smooth robust forecasting methods and applies them to 5-year ahead predictions of UK productivity, along with the Cardt forecasts, before addressing a productivity puzzle.

6.3 Exogeneity in Ice-Ages forecasts

As Pretis (2019) remarks ‘Econometric studies beyond IAMs (integrated assessment models) are split into two strands: one side empirically models the impact of climate on the economy, taking climate variation as given ... the other side models the impact of anthropogenic (e.g., economic) activity onto the climate by taking radiative forcing—the incoming energy from emitted radiatively active gases such as CO₂—as given... This split in the literature is a concern as each strand considers conditional models, while feedback between the economy and climate likely runs in both directions.’ Pretis (2021) addresses the exogeneity issue in more detail. Examples of approaches conditioning on climate variables such as temperature include Burke, Hsiang, and Miguel (2015), Pretis, Schwarz, Tang, Haustein, and Allen (2018), Burke, Davis, and Diffenbaugh (2018), and Davis (2019). Hsiang (2016) reviews such approaches to climate econometrics. Examples from many studies modelling climate time series include Estrada, Perron, and Martínez-López (2013), Kaufmann, Kauppi, Mann, and Stock (2011), Kaufmann, Kauppi, Mann, and Stock (2013) and Pretis and Hendry (2013).⁸

The dynamic simultaneous system in Castle and Hendry (2020) for modelling and forecasting Antarctic Ice volume and Temperature, and atmospheric CO₂ over the last 800,000 years of Ice Ages (see Figure 3) conditioned on the contemporaneous and lagged values of the Earth’s orbital variables of eccentricity, obliquity and precession shown in Figure 4. The open model taxonomy in Hendry and Mizon (2012) added nine potential sources of forecast error to the ten that occur in closed models, so we now address their possible impacts on the forecast accuracy of the system 100,000 years into the future. They show that despite the in-sample forecasting model being correctly specified and all unmodelled variables (denoted by the vector z_t) are strongly exogenous with known future values, changes in dynamics can lead to forecast failure when the z_t have non-zero means.

To illustrate, let the DGP of a vector y_t conditional on known z_t be:

$$y_t = \Psi_1 y_{t-1} + \Psi_2 z_t + \epsilon_t \text{ where } \epsilon_t \sim \text{IN}[\mathbf{0}, \Omega_\epsilon] \quad (29)$$

which has a zero intercept. If z_t has a non-zero mean, $E[z_t] = \mu$, then when Ψ_1 has all its eigenvalues inside the unit circle:

$$E[y_t] = \psi = (\mathbf{I} - \Psi_1)^{-1} \Psi_2 \mu \quad (30)$$

which is the equilibrium mean. In terms of deviations from equilibrium means:

$$y_t = \psi + \Psi_1 (y_{t-1} - \psi) + \Psi_2 (z_t - \mu) + \epsilon_t. \quad (31)$$

⁸See <https://us5.campaign-archive.com/?u=d1bdd6126f95e7ead3788a350&id=1cd67a8763> for summaries of many contributions to climate econometrics from members of our network <https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=climateeconometrics>.

If any of the parameters in (30) change, then the equilibrium mean shifts leading to forecast failure until the model is updated despite the future z_t being super-strongly exogenous and known.

Forecasting after the other parameters shift at $T + 1$, so that the DGP becomes:

$$\mathbf{y}_{T+1} = \Psi_1^* \mathbf{y}_T + \Psi_2^* z_{T+1} + \epsilon_{T+1} \quad (32)$$

from a forecast origin at T where \mathbf{y}_T is known using:

$$\hat{\mathbf{y}}_{T+1|T} = \hat{\Psi}_1 \mathbf{y}_T + \hat{\Psi}_2 z_{T+1} \quad (33)$$

leads to the forecast error $e_{T+1|T} = \mathbf{y}_{T+1} - \hat{\mathbf{y}}_{T+1|T}$:

$$e_{T+1|T} = (\Psi_1^* - \Psi_1) \mathbf{y}_T + (\Psi_1 - \hat{\Psi}_1) \mathbf{y}_T + (\Psi_2^* - \hat{\Psi}_2) z_{T+1} + \epsilon_{T+1}. \quad (34)$$

When the in-sample parameter estimates are unbiased:

$$E[e_{T+1|T}] = (\Psi_1^* - \Psi_1) \psi = (\Psi_1^* - \Psi_1) (\mathbf{I} - \Psi_1)^{-1} \Psi_2 \mu. \quad (35)$$

Thus the equilibrium mean shifts when $\mu \neq \mathbf{0}$ and would also do so if there was an intercept $\phi \neq \mathbf{0}$ in the DGP. If neither μ nor ϕ shift, then the key problem appears to be shifts in Ψ_1 , although if $\hat{\mathbf{y}}_T$ has to be estimated, then $\mathbf{y}_T - \hat{\mathbf{y}}_T$ could also cause a systematic forecast error.

The Earth's orbital drivers of eccentricity, obliquity and precession are super-strongly exogenous and known into the distant future, albeit that a sufficiently large rogue object intruding into our solar system could perturb our orbit in unanticipated ways. Excluding that last possibility for the next 100,000 years, so that μ does not change, then the forecasting system being open does not by itself create additional problems. Instead, the change to the system from CO₂ being endogenously determined to being created anthropogenically is a fundamental shift in the DGP seen in Figure 2. Castle and Hendry (2020) tackled this by computing two scenarios where the model for Ice volume and Temperature remains constant, but CO₂ is determined exogenously. The impact on the forecasts is dramatic as Figure 21 shows. The top plot is for Ice under the three scenarios and the lower plot for temperature.

Forecasting 110,000 years ahead under *ceteris paribus*, so there is no human intervention shown as long-dashed with $\pm 2\bar{\sigma}_f$ bars, mimics the Ice-Age data. Forecasting conditional on atmospheric CO₂ remaining at 400ppm (roughly the present level) shown as dotted with solid $\pm 2\tilde{\sigma}_f$ error bands leads to less Ice than the Ice-Age minimum over the last 800,000 years and higher temperatures than the peak of the Ice Age. Finally, at 560ppm (roughly Representative Concentration Pathway, RCP8.5), shown as thick dashed with $\pm 2\hat{\sigma}_f$ fans, temperatures are far higher than during the Ice Ages and Antarctica is almost ice free.

6.4 Explaining Cardt

For a non-stationary time series, we decompose the data into a trend, a seasonal and an irregular component, then forecast the components separately before aggregating their results. The forecasts for the trend and the irregular are computed using Cardt (see Castle, Doornik, and Hendry, 2021), which is a modified version of our M4 competition submission predictor (see Makridakis, Spiliotis, and Assimakopoulos, 2020) described in Doornik, Castle, and Hendry, 2020a. Any seasonal component is extrapolated from most recent estimates of the seasonal pattern. Cardt denotes a Calibrated Average of Rho (ρ), Delta (δ) and THIMA where ρ estimates a simple autoregressive model with seasonality, forcing a unit root if the estimates are close to unity, so then switches to a first-difference model with dampened mean; δ estimates the growth rate based on first differences, but dampened by removing large values and allowing for seasonality; and THIMA is a **t**rend-**h**alved **i**ntegrated **m**oving **a**verage model, namely a dampened trend arbitrarily halved, with an intercept correction estimated by a moving-average model.

Cardt computes the arithmetic average of these three forecasts, which are then calibrated by treating the forecasts as if they were observed: a richer autoregressive model is estimated from this extended data

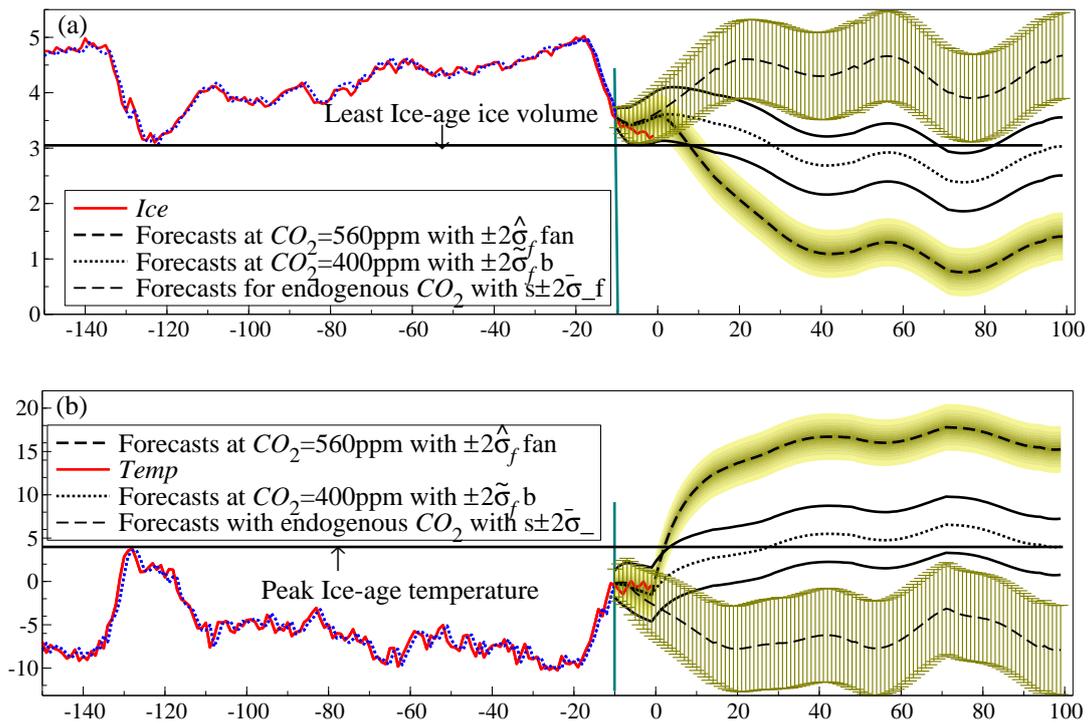


Figure 21: Three scenarios of different atmospheric CO₂ levels: endogenous, and fixed at 400ppm and 560ppm for (a) Ice volume; (b) temperature.

series. The fitted values from this calibrated model are the final forecasts, undoing any transformations such as logs and differencing to obtain the levels. Higher orders of integration (e.g. $I(2)$ and damped $I(2)$) can be handled: see Doornik, Castle, and Hendry (2020b). A second ‘average forecast’ is also calculated by computing forecasts commencing from one, two and three observations earlier, adjusted to match the last known observation.

Figure 6.4 provides a graphical explanation using our COVID-19 forecasts (see Doornik, Castle, and Hendry, 2020c): dates like 2020-10-25 denotes 25th October, 2020. Panel (a) shows a very short section of the data and the current estimated trend (dotted). This is calculated by saturating moving windows of data by linear trends (as with TIS) so initially there are as many trends as observations, from which significant ones are selected using Autometrics. These selected linear trends are averaged to yield an overall flexible trend: see Doornik (2019). Panel (b) shows the forecast computed from separately forecasting the trend and irregular using Cardt, dampening the trend, where the seasonal is extrapolated and added to the forecasts for trend and irregular. The 80% uncertainty bands are shown by the thin dotted lines. Panel (c) adds the forecasts made commencing from up to three days earlier shown by thin lines with circles. Finally, the thick dashed line in Panel (d) shows their average forecast. Differences between the thick dashed line added in (d) and the first Cardt forecasts in (b) indicate shifts: when the average lies above the Cardt forecast, the most recent data have led to lower values, which in the pandemic raises hopes it may be being brought under control.

6.5 Smooth robust forecasts

UK productivity measured as GDP per employee per annum has stagnated since 2008 as shown in Figure 23 (also see Figure 15). The principle underlying robust forecasting devices is to use local estimates of the equilibrium mean rather than full sample ones. Many moving windows approaches implement a similar idea but applied to all the parameters rather than just the equilibrium mean which is the key source

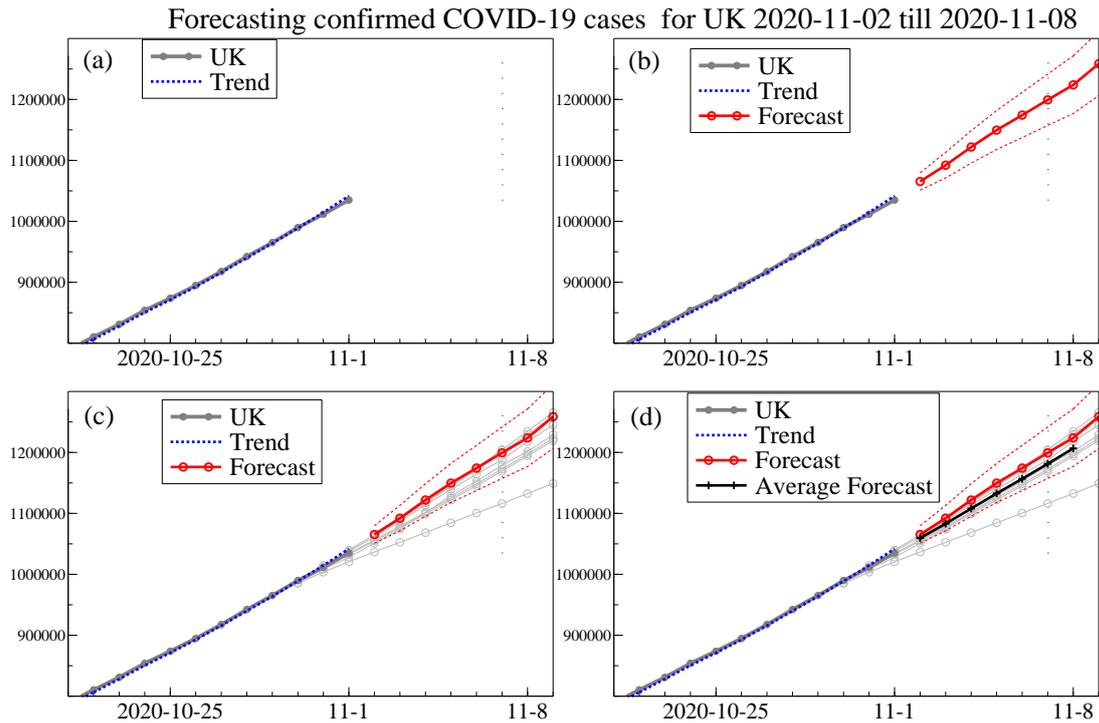


Figure 22: Explaining forecasting by Cardt: (a) trend line; (b) Cardt forecast; (c) 3 earlier-based forecasts; (d) average forecast added.

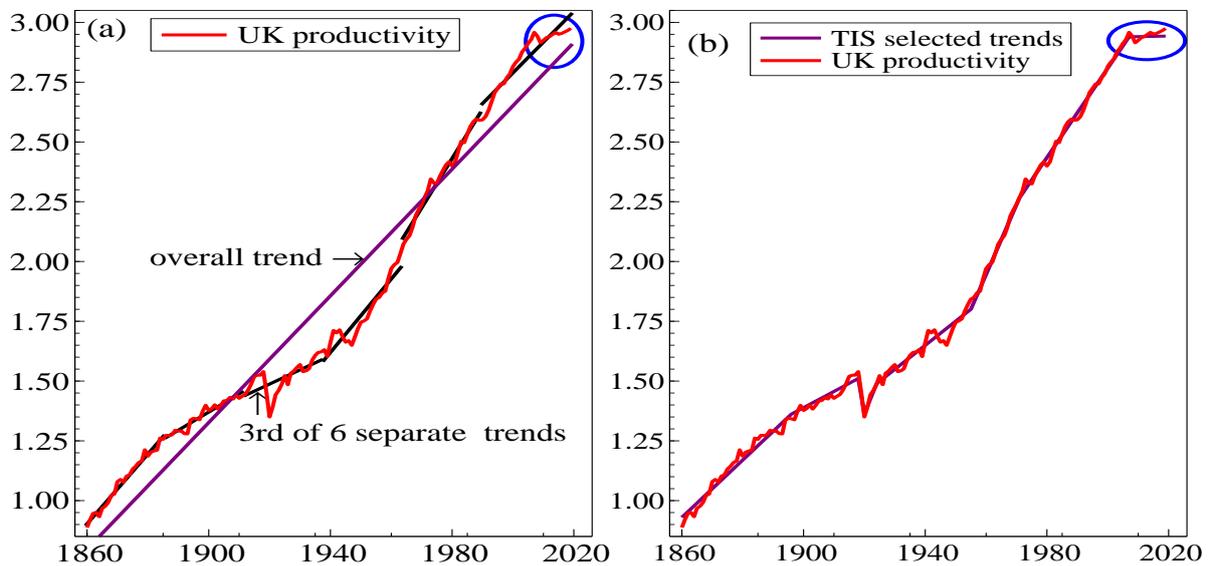


Figure 23: (a) UK productivity since 1860 with overall and six separate trend lines at roughly 25-year subperiods; (b) UK productivity since 1860 with seven TIS trend lines selected at 0.01%.

of systematic mis-forecasting. Simply differencing all variables eliminates the equilibrium-correction mechanism (EqCM), often of importance to policy. The initial approach in Hendry (2006) sought to retain the EqCM but just used the previous observation so although it avoided systematic forecast errors, could be very volatile. Consequently, Castle, Clements, and Hendry (2015) proposed a class of methods

that used moving averages of recent data to smooth the forecasts. Martinez, Castle, and Hendry (2021) show that robust forecasts can be interpreted as alternative local estimators of the long-run mean so develop smoothed estimates thereof.

Figure 24 Panel (a) records the five-year-ahead forecasts made by the Office for Budget Responsibility (OBR) forecasting UK productivity since 2008. As can be seen, these systematically greatly over-predict future productivity year after year. This is a typical example of the problem with EqCMs noted above where the model returns to the in-built equilibrium (here an underlying trend) irrespective of the data having a different trajectory. Panel (b) from Martinez, Castle, and Hendry (2021) reports forecasts from a robust predictor, namely a smooth random walk which differnces the data when forecasting then the forecasts are re-integrated to the level. Panel (c) shows the forecasts from the smooth double difference device. While these are far better than OBR, they highlight the erratic nature of some forecasts as does Panel (d) for the smooth double robust device. In the last two cases, despite the smoothing, small dips in the data are extrapolated into the distant future highlighting the trade-off between the delay in responding and the length of the smoothing period.

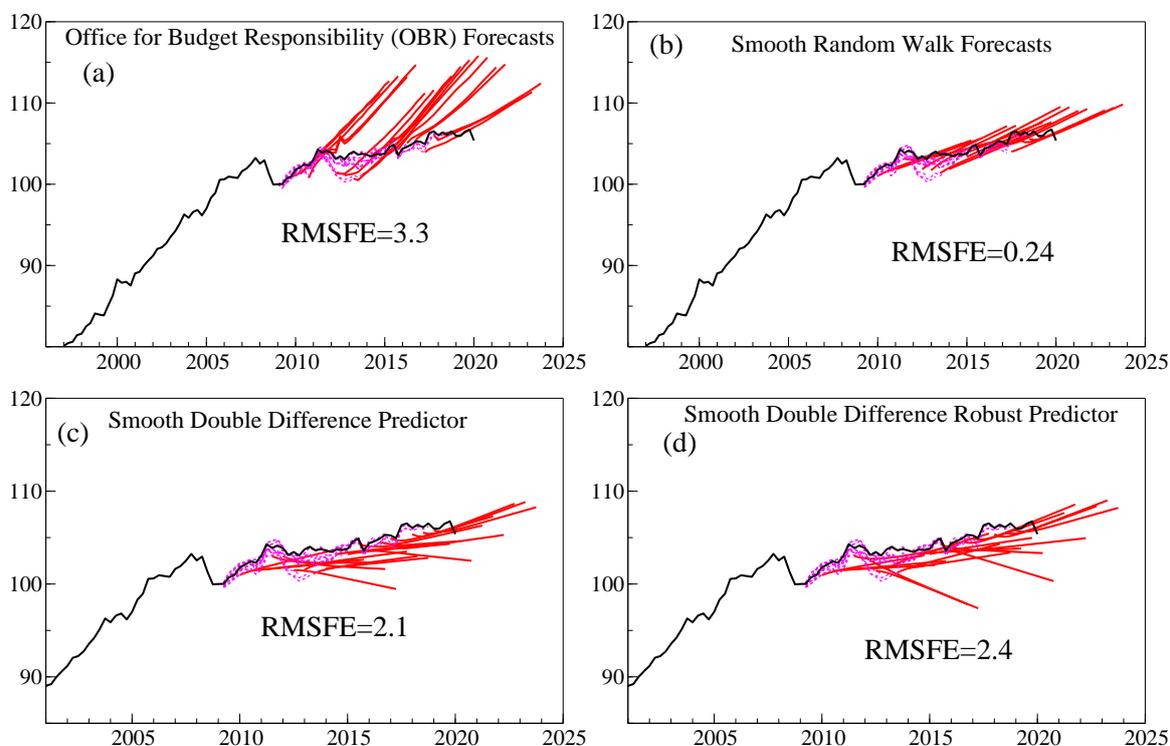


Figure 24: Forecasts of UK productivity: (a) Office for Budget Responsibility; (b) a smooth random walk; (c) smooth double difference device; (d) smooth double robust device.

Figure 25 records the Cardt forecasts for UK productivity, both for the annual time series in figure 23, plotted in the top panels and the quarterly series in Figure 24 plotted in the bottom panels. For both data frequencies the Cardt forecasts adjust too slowly to the break in trend in 2008, predicting higher productivity over the next few years than actually materialised, but it only takes a couple of observations for the forecasts to adapt to the changed trend. The RMSFE is 1.7 for the quarterly 5 year ahead Cardt forecasts, produced every 6 months and commencing evaluation in 2009 to match those of the OBR, delivering substantial gains over the OBR forecasts. Only the smoothed random walk forecasts were more accurate over 5 years ahead. The uncertainty bands on the annual forecasts are extremely wide (from 1.3–7.7) because of the very small sample size used for estimation (just from 2000 onwards), so are not shown, but those for the quarterly data (estimating from 1997Q1 onwards) illustrate their range

for the initial forecasts but rapidly narrowing. Again there is a trade-off between a shorter estimation sample which avoids non-modelled in-sample breaks that would bias the estimates of the forecasting models in Cardt, versus a longer sample needed to narrow down the forecast uncertainty bands. When multiple breaks occur, the RMSFEs are smaller using shorter estimation samples.

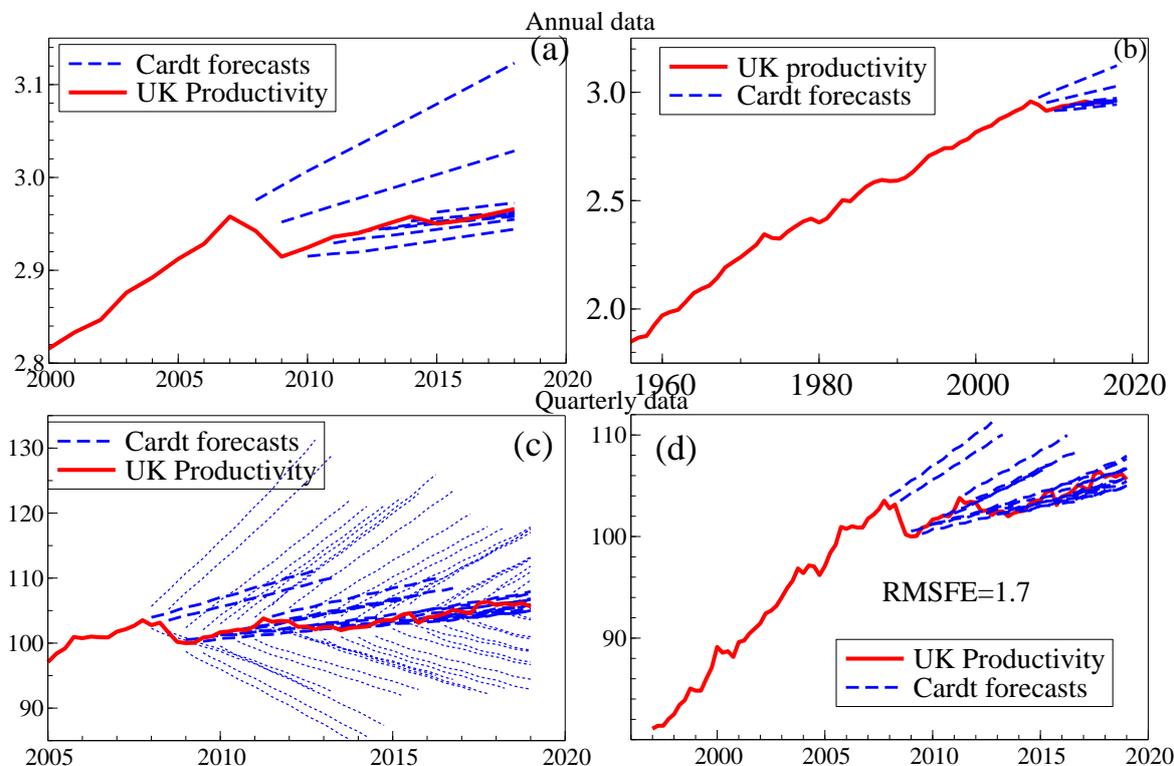


Figure 25: Forecasts of UK productivity using Cardt: (a) annual dynamic forecasts estimated from 2000 onwards; (b) the annual dynamic forecasts recorded on a longer time-series; (c) 5 year ahead forecasts for the quarterly data estimated from 1997Q1 onwards with uncertainty bands; (d) the quarterly 5 year ahead forecasts (produced every two quarters) recorded on a longer time-series.

6.6 The ‘Paradox’ of stagnant real wages yet rising real incomes

Real wages and productivity in the UK have stagnated since 2007. The concern of a failure of ‘living standards’ to rise at the same rate as post-WWII rates obscures the fact that real GDP has grown by more than 20% over the period despite the ‘Great Recession’, so aggregate living standards have in fact risen. The issue matters for climate analyses since it might be falsely thought that because the productivity stagnation followed the Climate Change Act of 2008 (CCA2008) to reduce the UK’s CO₂ emissions, it was due to the changes implemented therein. Rather, the bigger events of 2008—the Financial Crisis and ‘Great Recession’—are more likely culprits for the stagnation. Although the apparent paradox remains, it has an easy resolution: before the pandemic, employment rose considerably faster than population.

Figure 26 records UK log real wages and log output per worker per year (Panel a), along with the wage share (Panel b), employment and population (Panel c) and log real GDP per capita and log real GDP per employee, aka productivity (Panel d). Employment has risen by about 4.6 million since 2000, an increase of around 18% of the previous labour force. Pre-pandemic, the level of employment was the highest ever for the UK and has grown much faster than population over the last quarter century. With so many more employees and real wages relatively constant, total earnings must have increased

considerably and indeed have risen by almost 35% this millennium and 15% since 2008. In fact, real GDP per capita of the population has increased by almost 22% since 2000. The explanation for the dramatic difference from stagnant productivity growth lies in this remarkably different behaviour of aggregate employment and population. The ratio of employment to population has been mostly rising since the 1980s and is currently at the highest ever peacetime level, leading to the divergence between real GDP per capita and real GDP per employee.

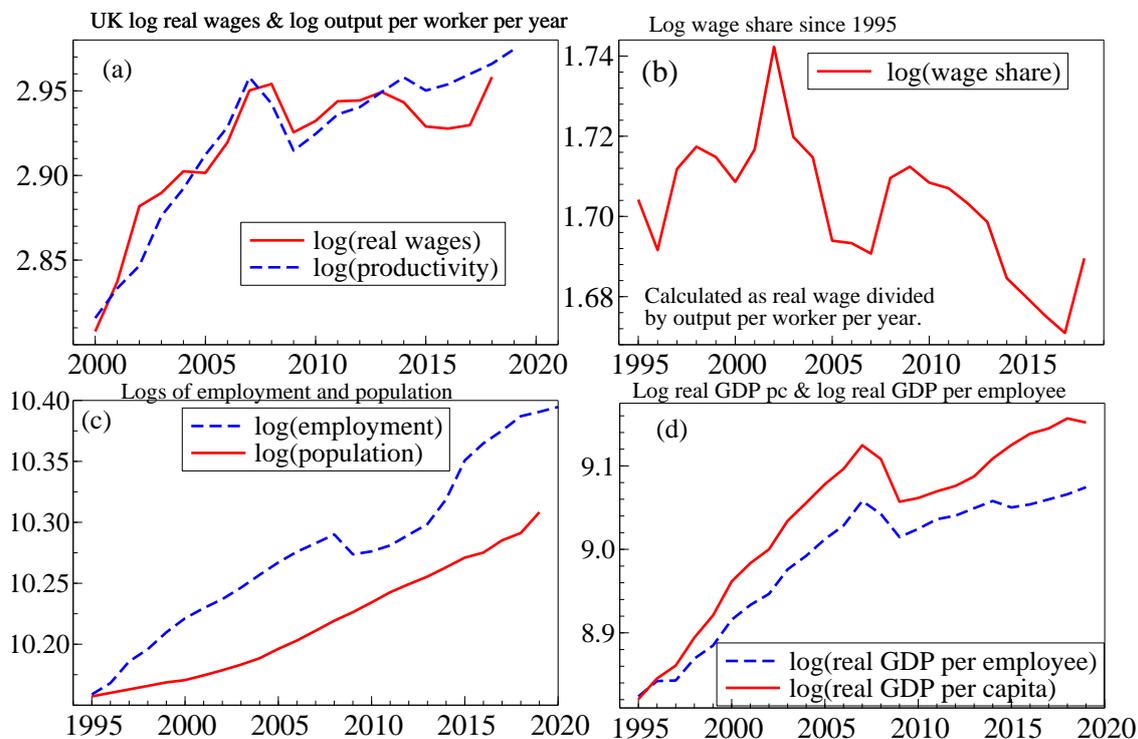


Figure 26: (a) UK log real wages and log output per worker per year; (b) Log wage share since 1995; (c) Logs of employment and population; (d) Log real GDP per capita and log real GDP per employee.

7 Updating the UK CO₂ model

Since modelling UK CO₂ emissions reported Castle and Hendry (2020), two more annual observations have been produced, allowing an important update of their equation. This enables us both to check the role of the step indicator S_{2010} for the UK's CCA2008, and to test the constancy of the relationship. At the time of their model, only two data points were available for estimating the coefficient of S_{2010} when using the final four observations for forecasting; again keeping the last four data points for forecasts doubles to four that number of observations. In turn, the updated estimate is sufficiently precise that S_{2010} can be included in the cointegrating vector.

Figure 27 records the extended UK data series together with US CO₂ emissions per capita, in tons p.a., 1850–2019. UK per capita CO₂ emissions have continued to fall, as has fossil fuel usage, whereas wind+solar has risen rapidly, albeit still less than both oil and gas. However, the US per capita CO₂ emissions remain higher than the highest UK values.

The distributional shifts of total UK CO₂ emissions in Mt p.a. shown in Figure 28 continue to emphasise the need to handle them in modelling. Hendry and Mizon (2011) highlighted how failing to handle shifting and evolving relationships can lead to rejecting a sound theory. Consequently, denoting impulse (IIS) and step (SIS) indicators by $\mathbf{1}_{\{abcd\}}$ and $S_{\{abcd\}}$ respectively where observation $abcd$ is an

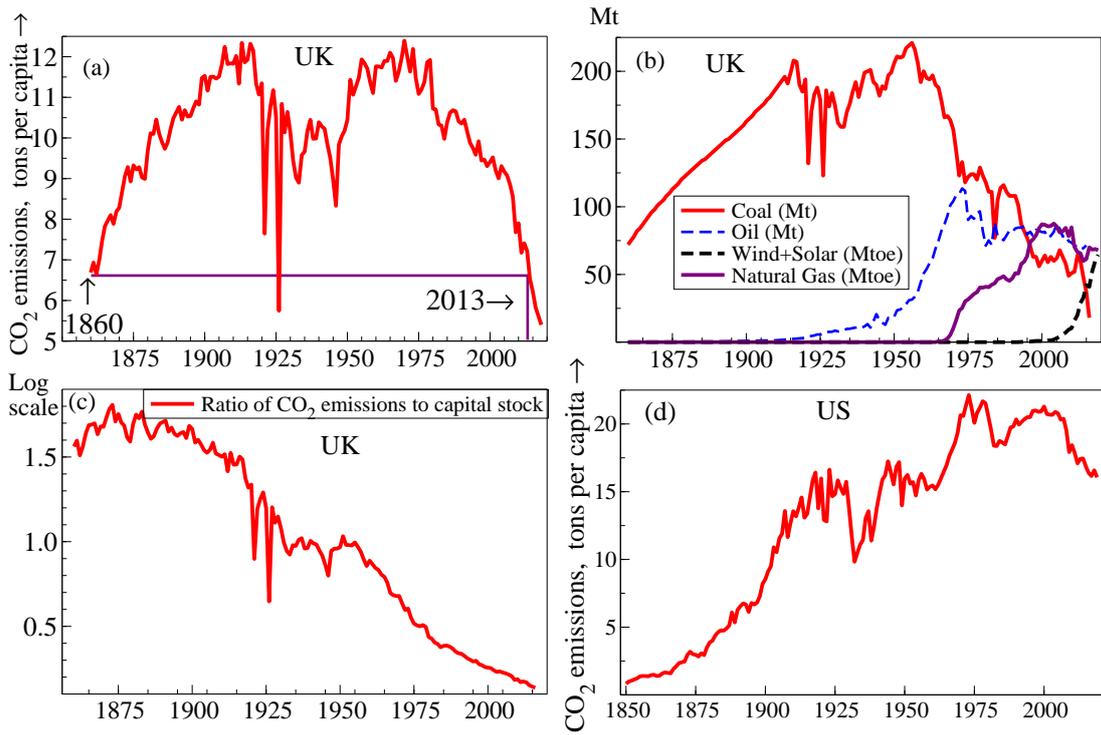


Figure 27: (a) UK CO₂ emissions per capita in tons per annum (p.a.) to 2018; (b) UK coal (millions of tonnes, Mt), oil (Mt), natural gas (millions of tonnes of oil equivalent, Mtoe) and wind+solar (Mtoe), all to 2018; (c) ratio of CO₂ emissions to the capital stock on a log scale to 2017; (d) US CO₂ emissions per capita, in tons p.a., 1850–2019

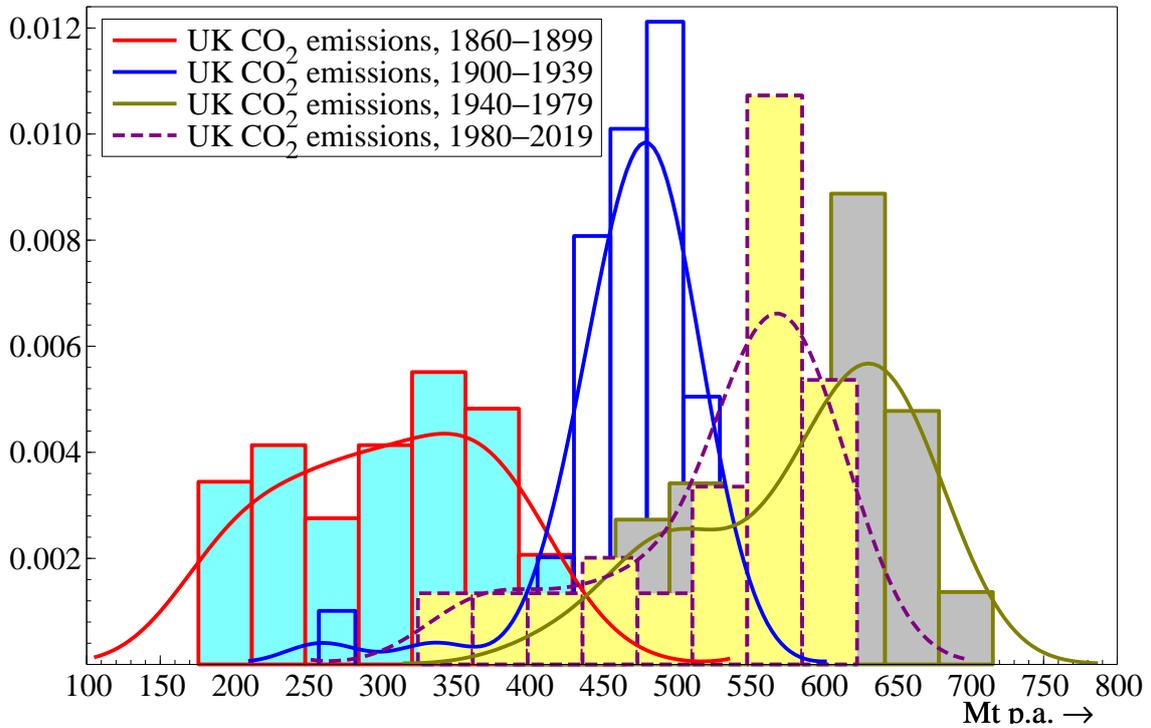


Figure 28: Shifting sub-period distributions of UK CO₂ emissions.

outlier or the end of a step shift, the general unrestricted model (GUM) was formulated as:

$$E_t = \beta_0 + \beta_1 E_{t-1} + \beta_2 C_t + \beta_3 C_{t-1} + \beta_4 O_t + \beta_5 O_{t-1} + \beta_6 k_t + \beta_7 k_{t-1} + \beta_8 g_t + \beta_9 g_{t-1} + \text{indicators} + \epsilon_t. \quad (36)$$

The model was re-selected from data over 1861–2013, testing constancy for 2014–2017, first selecting by super-saturation combining IIS and SIS at $\alpha_1 = 0.001$, with all other explanatory variables retained. Equation (37) records the outcome at this stage.

$$\begin{aligned} \widehat{E}_t = & \quad 0.49 E_{t-1} - \quad 47 \mathbf{1}_{\{1921\}} - \quad 161 \mathbf{1}_{\{1926\}} - \quad 44 \mathbf{1}_{\{1946\}} + \quad 54 \mathbf{1}_{\{1947\}} \\ & \quad (0.06) \quad (13) \quad (20) \quad (10) \quad (11) \\ & + \quad 28 \mathbf{1}_{\{1996\}} - \quad 39 S_{\{1925\}} + \quad 71 S_{\{1927\}} - \quad 32 S_{\{1969\}} + \quad 35 S_{\{2010\}} \\ & \quad (9.8) \quad (14) \quad (13) \quad (7.5) \quad (7.0) \\ & - 189 + 1.86 C_t - 0.84 C_{t-1} + 1.70 O_t - 1.01 O_{t-1} + 0.94 g_t \\ & \quad (91) \quad (0.12) \quad (0.18) \quad (0.26) \quad (0.28) \quad (0.33) \\ & - 1.12 g_{t-1} + 7.97 k_t - 7.32 k_{t-1} \end{aligned} \quad (37)$$

$$\begin{aligned} \widehat{\sigma} &= 9.62 \quad R^2 = 0.995 \quad F_{\text{ar}}(2, 131) = 1.59 \quad \chi_{\text{nd}}^2(2) = 4.98 \quad F_{\text{arch}}(1, 150) = 2.73 \\ F_{\text{Het}}(21, 124) &= 0.79 \quad F_{\text{Reset}}(2, 131) = 2.29 \quad F_{\text{Chow}}(4, 133) = 1.48 \quad F_{\text{nl}}(27, 109) = 1.08 \end{aligned}$$

Coefficient standard errors are shown in parentheses, $\widehat{\sigma}$ is the residual standard deviation, R^2 is coefficient of multiple correlation, F_{ar} tests residual autocorrelation (see Godfrey, 1978), F_{arch} tests autoregressive conditional heteroscedasticity (see Engle, 1982), F_{Het} tests residual heteroskedasticity (see White, 1980), $\chi_{\text{nd}}^2(2)$ tests non-Normality (see Doornik and Hansen, 2008), F_{Chow} is a parameter constancy forecast test over 2012–2016 (see Chow, 1960), F_{Reset} tests non-linearity (see Ramsey, 1969), as does F_{nl} (see Castle and Hendry, 2010). Also, the *PcGive* unit-root t-test value of $t_{ur} = 9.33^{**}$ strongly rejects no cointegration: see Doornik and Hendry (2018), and Ericsson and MacKinnon (2002), for critical values.

No mis-specification tests were significant and five impulse and four step indicators were selected from the 307 candidate variables despite $\alpha_1 = 0.001$. Of these, $\mathbf{1}_{\{1926\}}$ and $S_{\{1927\}}$ can be combined to $\Delta \mathbf{1}_{\{1926\}}$, and $\mathbf{1}_{\{1947\}} - \mathbf{1}_{\{1946\}} = \Delta \mathbf{1}_{\{1947\}}$ leaving three step indicators: $\widehat{\sigma}$ was unaffected by these. SIS indicators in *Autometrics* terminate at the dates shown, so reflect what happened **earlier**. Thus, a positive coefficient for S_{1925} entails a higher level prior to 1926, the date of the Act of Parliament creating UK's first nationwide electricity grid, enhancing its efficiency, but also the General Strike. However, as coal is a regressor, indicators for miners' strikes should only be needed to capture large changes in inventories, perhaps as in 1926. The Clean Air Act of 1956 did not need a step indicator as the drop in CO_2 should again be captured by fall in coal use. Next, 1969 was the start of converting burner equipment from coal gas (about 50% hydrogen) to natural gas (mainly methane) with a considerable expansion in the use of gas over the following decades. The shift in 2010 seems to be a response to the CCA2008 and the European Union Renewables Directive of 2009, rather than the Great Recession, as the larger GDP fall in 1921–1922 did not need a step although there was an impulse indicator for the large outlier in 1921. We did not impose that any policies had an effect—the data tell us it did. The coefficients of all these location shifts have the appropriate signs of reducing and increasing emissions respectively. Selecting the fuel and economic regressors at $\alpha_2 = 0.01$, including the indicators in (37) retained all those variables.

7.1 Cointegration

The cointegrating, or long-run, relation was derived from that equation after transforming the indicators as noted. When mapping to a non-integrated specification that reparametrizes levels variables to first differences, step indicators should be included in the equilibrium correction mechanism (EqCM) to avoid

cumulating to trends. However, they need to be *led* one period as the EqCM will be lagged in the $l(0)$ formulation. Impulse indicators and differenced step indicators are left unrestricted: see the survey articles by Hendry and Juselius (2000, 2001) and Hendry and Pretis (2016). Applications of cointegration analysis in climate econometrics include Kaufmann and Juselius (2010), Kaufmann, Kauppi, Mann, and Stock (2013) and Pretis (2019). This yielded:

$$\begin{aligned} \tilde{E}_{LR} = & \quad 2.0 C + \quad 1.4 O + \quad 1.25 k - \quad 0.35 g + 61 S_{\{1924\}} - 62.0 S_{\{1968\}} \\ & \quad (0.06) \quad (0.19) \quad (0.28) \quad (0.29) \quad (7) \quad (14) \\ & + 70.0 S_{\{2009\}} - 234 . \\ & \quad (13) \quad (170) \end{aligned} \quad (38)$$

All variables are significant at 1% other than g which enters negatively. The coefficients of coal and oil are close to their values in Table 1. With 4 observations on $1 - S_{\{2010\}}$, S_{2010} is now precisely estimated so could be included in \tilde{E}_{LR} .

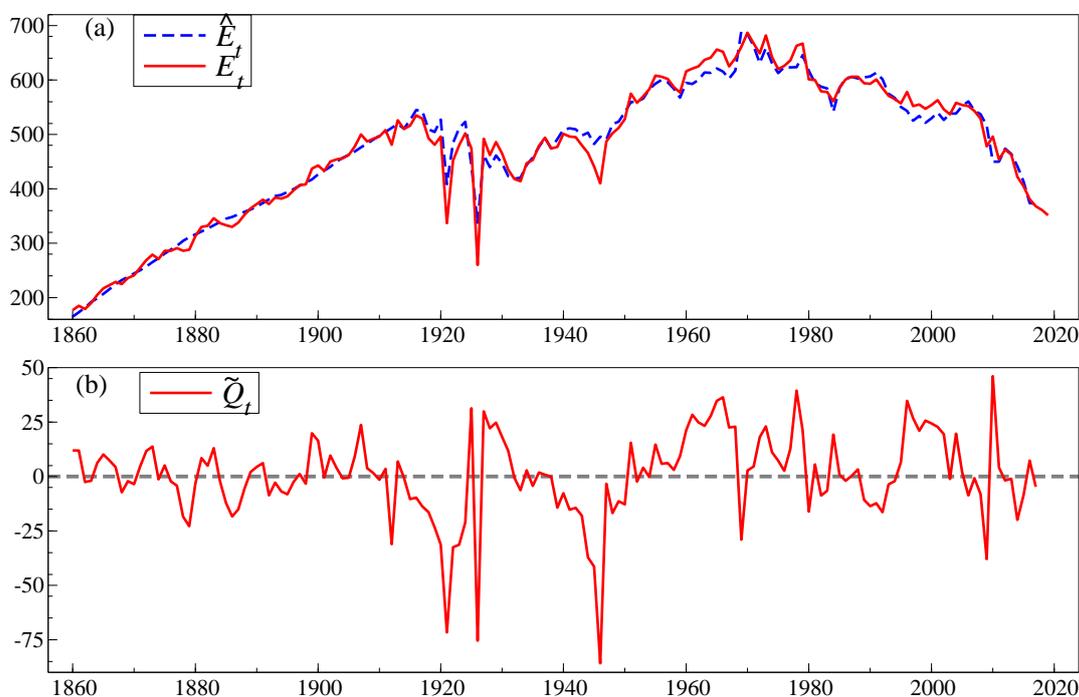


Figure 29: (a) E_t and \hat{E}_t ; (b) \tilde{Q}_t without impulse indicators, centered on a mean of zero.

Transforming to a model in first differences and the lagged EqCM from (38), then re-estimating, revealed a significant non-normality test, so IIS was re-applied at 0.5%, which yielded (for 1861–2013, testing constancy over 2014–2017):

$$\begin{aligned} \widehat{\Delta E}_t = & \quad 1.88 \Delta C_t + \quad 1.72 \Delta O_t + \quad 7.16 \Delta k_t + \quad 0.88 \Delta g_t - \quad 0.50 \tilde{Q}_{t-1} - 59.2 \\ & \quad (0.10) \quad (0.21) \quad (1.10) \quad (0.28) \quad (0.05) \quad (9.2) \\ & - \quad 79.4 \Delta 1_{\{1926\}} + \quad 50.3 \Delta 1_{\{1947\}} - \quad 45.9 1_{\{1921\}} - \quad 28.3 1_{\{1912\}} \\ & \quad (8.77) \quad (6.42) \quad (11.1) \quad (8.93) \\ & + \quad 26.8 1_{\{1978\}} + \quad 27.5 1_{\{1996\}} \end{aligned} \quad (39)$$

$$\begin{aligned} \hat{\sigma} = 8.88 \quad R^2 = 0.94 \quad F_{ar}(2, 139) = 0.49 \quad \chi_{nd}^2(2) = 1.68 \quad F_{Het}(14, 134) = 1.02 \\ F_{arch}(1, 151) = 0.54 \quad F_{Reset}(2, 139) = 1.5 \quad F_{ni}(15, 126) = 1.34 \quad F_{Chow}(4, 141) = 1.76. \end{aligned}$$

Changes in coal, oil, k and g all lead to changes in the same direction in emissions, which then equilibrates back to the long-run relation in (38). Figure 30 provides a graphical description of the selected model. We have put ‘forecasts’ in inverted commas since they are for the past, although the data points were outside the sample period used for selection and estimation. As the top-left graph is dominated by the fluctuations in the 1920s, Figure 31 plots the levels outcome from modelling, with the impulse and step indicator dates.

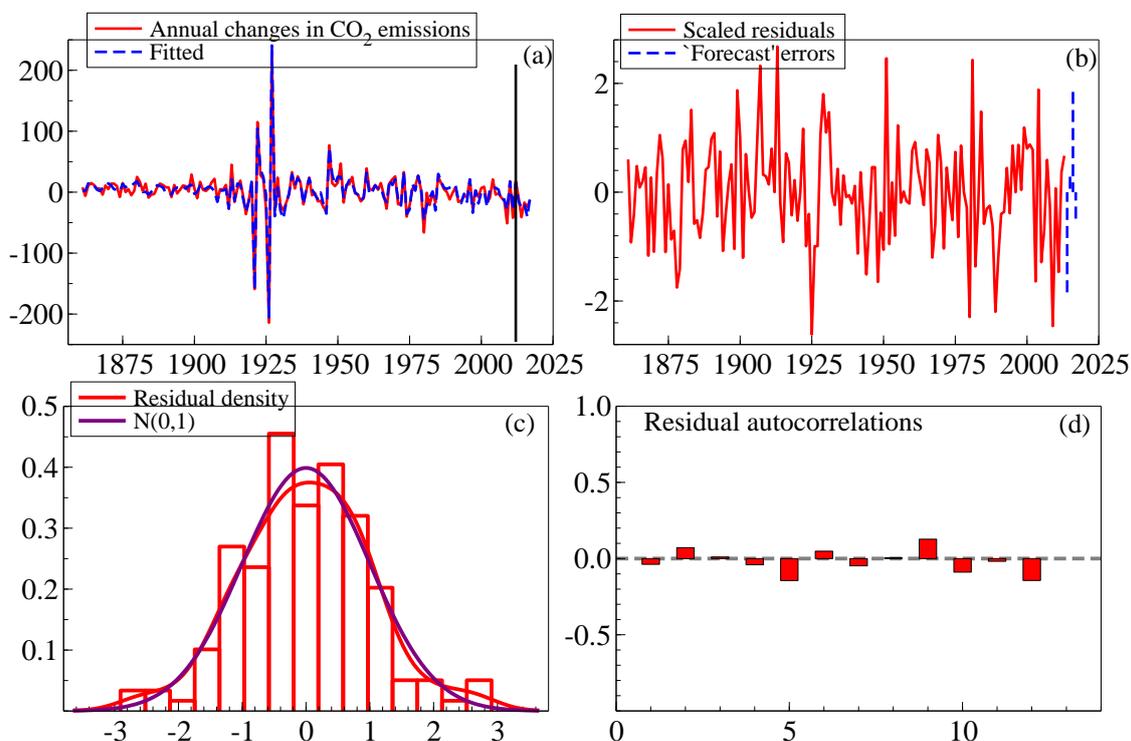


Figure 30: (a) Actual, fitted and ‘forecast’ values for ΔE_t from (39); (b) residuals and ‘forecast’ errors scaled by the equation standard error; (c) residual density and histogram with a Normal density; (d) residual autocorrelation.

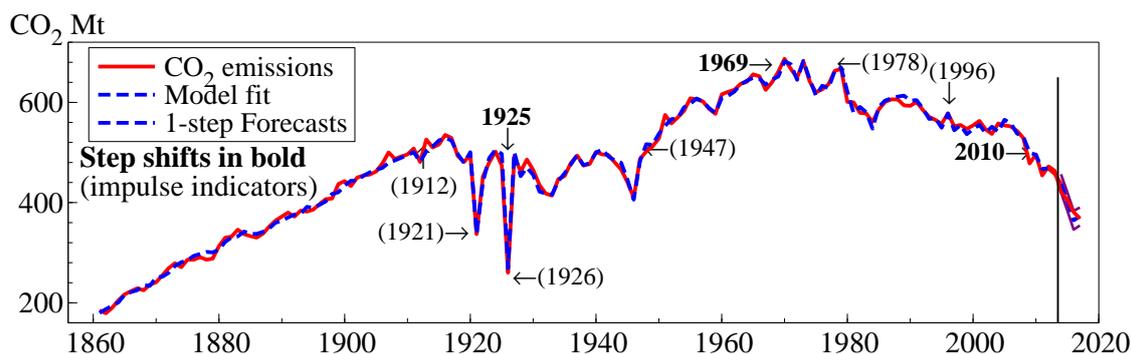


Figure 31: Actual, fitted and ‘forecast’ values for UK CO₂ emissions.

7.2 ‘Forecast’ evaluation

Figure 32 Panel (a) records the change in CO₂ emissions and the fitted values from (39), along with the 1-step conditional ‘forecasts’, denoted $\widehat{\Delta E}_{T+h|T+h-1}$, with $\pm 2\hat{\sigma}_f$ shown as bars, estimating the model

up to 2014. In the same panel, the robust forecasts described in Section 6.5 (differencing the EqCM but without smoothing) are also reported. The forecasts are similar. Panel (b) integrates the forecasts to obtain 1-step ahead forecasts of the level of CO₂ emissions from both the model in (39) (with $\pm 2\hat{\sigma}_f$) and the robust forecasts (denoted with $\pm 2\hat{\sigma}_f$). Again, the forecasts are similar, with no clear advantage to using either the econometric model or its robustified form. Panel (c) shifts the forecast origin forward to 2009, so data up to 2008 is available to estimate the model. Again, the 1-step conditional ‘forecasts’ from the econometric model, denoted $\widehat{\Delta E}_{T+h|T+h-1}$ with $\pm 2\hat{\sigma}_f$ bars, are recorded along with the robust forecasts $\widetilde{\Delta E}_{T+h|T+h-1}$. There is a clear benefit to using the robust forecasts over the longer forecast period. The conditional ‘forecasts’ come before the implementation of the CCA2008, leading to forecasts that are too high from 2012 onwards. The robust device remains accurate even when the CCA2008 is not explicitly modelled, as it differences out the previous in-sample mean for the change in CO₂ emissions, which was higher prior to the CCA2008, leading to biased 1-step ahead forecasts if not handled.

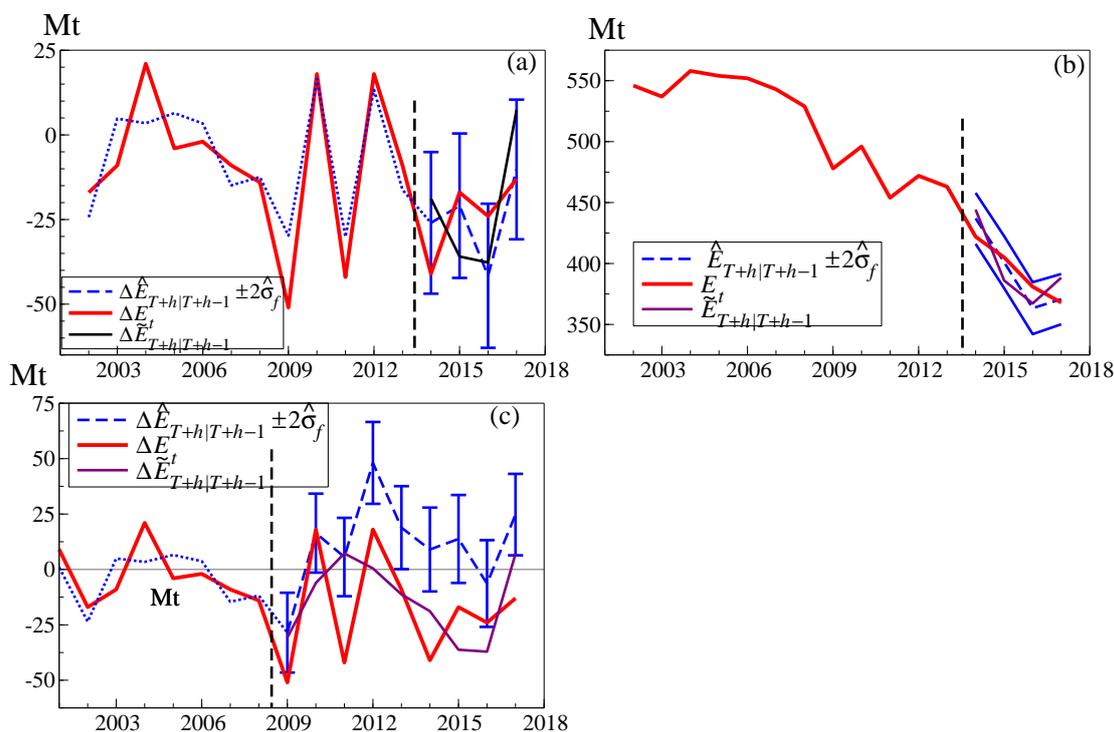


Figure 32: Conditional 1-step ‘forecasts’ for ΔE_t (a) from (39) over 2014–2017 (denoted $\hat{\cdot}$) with ± 2 forecast standard error bars and the robust forecasts $\Delta E_{T+h|T+h-1}$; (b) the derived ‘forecasts’ in levels; (c) same as (a) but commencing in 2009.

The 1-step ahead ‘forecasts’ highlight the impact of the CCA2008 which needs to be either modelled or accounted for (via robustification) for accurate forecasts. Modelling CO₂ emissions within a system allows for unconditional forecasts to be produced, so does not rely on data over the forecast period and hence is feasible *ex ante*. A vector autoregression (VAR) with 2 lags is sufficient to model the dynamics. Figure 33 plots the unconditional system 1-step and dynamic ‘forecasts’ from a VAR in all five variables, either including the indicators found in (39) or without the indicators. In the former case, all outliers and shifts will be captured in the VAR, whereas in the latter they will be ignored, which is a common approach in the economics literature that uses VARs as forecasting benchmarks. Panel (a) plots the 1-step ahead ‘forecasts’ from the VAR without and with indicators, and panel (b) records the dynamic ‘forecasts’. The importance of the step indicators is readily apparent, yielding huge reductions in RMSFEs and correcting the bias evident in the VAR without the steps. The role of the CCA2008 in producing a level shift reduction in CO₂ emissions is clear so needs to be modelled for accurate forecasting.

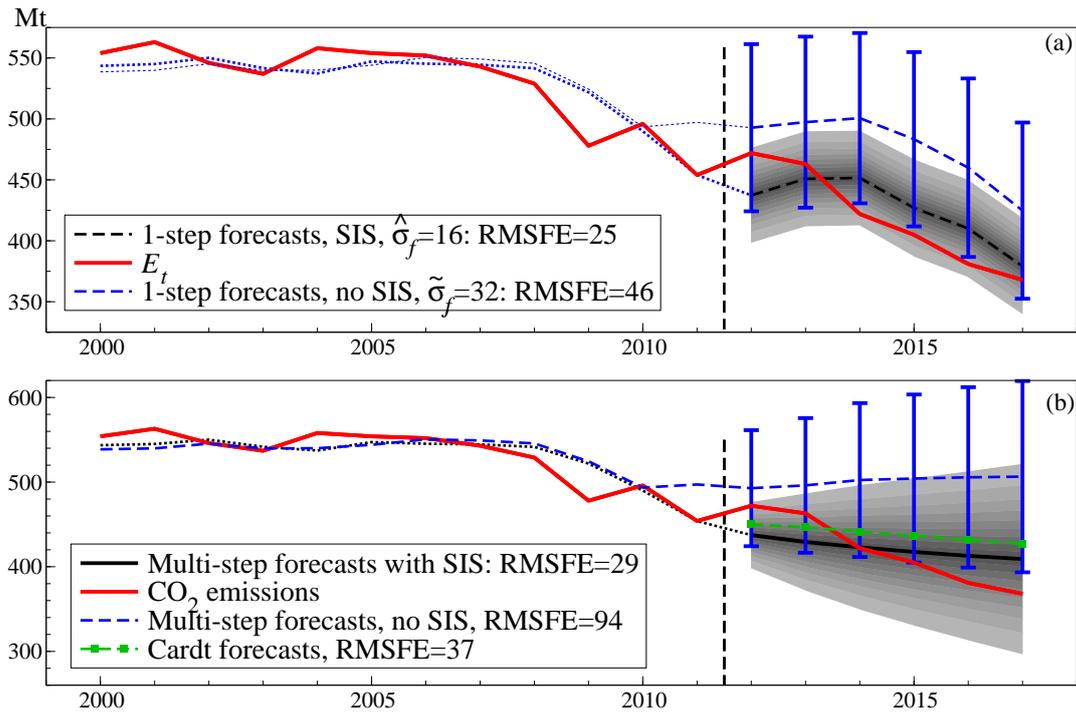


Figure 33: (a) Outcomes and 1-step ‘forecasts’ ± 2 forecast standard errors as bars and fans, without and with indicators; (b) Outcomes and h -step ‘forecasts’ and a comparison with Cardt ‘forecasts’.

Figure 33 also records the ‘forecasts’ for UK CO₂ emissions, comparing Cardt ‘forecasts’ to the VAR ‘forecasts’. The sample estimation period for the Cardt ‘forecasts’ is kept short (2005–2011) to avoid contaminating the estimates of ρ , δ and THIMA with earlier in-sample breaks. As a consequence, the uncertainty bands around the Cardt ‘forecasts’ are very wide, so are not shown. Extending the in-sample estimation period reduces the uncertainty bands but results in a larger bias due to unmodelled shifts in-sample, especially 2010. The Cardt ‘forecasts’ are close to the VAR ‘forecasts’ using SIS, which has a mean ‘forecast’ error of -3 , compared to a mean error for the Cardt ‘forecasts’ of -20 . In contrast, the VAR model without SIS has a mean error of -83 . The key characteristics of the Cardt forecasting device is that it dampens trends and growth rates, it averages across predictors and it robustifies to breaks by over-differencing. As such, it is comparable to the ‘forecasts’ from the VAR with SIS which models step shifts explicitly with step indicators, and is superior to the VAR ‘forecasts’ that do not model location shifts explicitly.

7.3 Super exogeneity tests

Parameter invariance is essential in policy models because without it, a model will mis-predict under regime shifts. The concept of super exogeneity combines parameter invariance with valid conditioning so is crucial for policy, see Engle and Hendry (1993), Krolzig and Toro (2002), Hendry and Massmann (2007), and Hendry and Santos (2010). To test the joint super exogeneity of the regressors, a natural procedure here is to check if any of the indicators in the conditional model enter the equations for the marginal variables. This requires no *ex ante* knowledge of the timing or magnitudes of breaks, or the data generating process of the marginal variables. The test has the correct size under null of super exogeneity for a range of sizes of marginal-model saturation tests, and it has power to detect failures of super exogeneity when location shifts occur in the marginal models.

Having created a VAR, we can test for super exogeneity by a likelihood ratio test of the VAR with the indicators only entering the equation for E_t against entering every equation. This yields $\chi^2(37) = 161^{**}$

which strongly rejects. However, the rejection is due to the indicators for the 1920s also occurring in the equations for GDP and coal, which is not surprising as the post-war crash and the 1926 general strike affected both. Retaining 1_{1921} , 1_{1926} , S_{1925} and S_{1927} in the coal equation and 1_{1921} and 1_{1926} in GDP delivers $\chi^2(31) = 25$ which is insignificant, implying the regressors in the CO₂ emissions model are super exogenous, so they are weakly exogenous and the parameters of the conditional model (39) are invariant to structural breaks in the marginal models.

7.4 Evaluating the UK's 2008 Climate Change Act

The most important implication of the above evidence is that substantial CO₂ reductions have been feasible, so far with little impact on GDP. The UK's CCA2008 established the world's first legally-binding climate-change target to reduce the UK's GHG emissions (6 gases including CO₂, which is approximately 80% of the total) by at least 80% by 2050 from a 1990 baseline. The policy produced a series of five-year carbon budgets, which we mapped to annual targets by starting 20Mt above and ending 20Mt below them in each period. Allowing 20% for other GHG emissions, we call these the Targets for CO₂. Figure 34 Panel (a) plots the Targets and CO₂. TargDiff denotes the difference between these targets and CO₂ emissions, and (40) records the result from selecting step indicators by SIS to describe it over 2008–2020.

$$\widehat{\text{TargDiff}}_t = \underset{(9.1)}{52.3} S_{2013} + \underset{(17)}{49.9} S_{2019} - \underset{(16)}{101} \quad (40)$$

$$\hat{\sigma} = 15.7 \quad R^2 = 0.85 \quad F_{\text{ar}}(1, 9) = 0.38 \quad \chi_{\text{nd}}^2(2) = 0.48$$

$$F_{\text{arch}}(1, 11) = 0.02 \quad F_{\text{Reset}}(2, 8) = 0.00 \quad T = 2008 - 2020$$

Thus, emissions were approximately 52Mt below target after 2013, and fell another 50Mt further below after 2019, part of which is undoubtedly due to the impacts from pandemic lockdowns. Nevertheless, these are large reductions. Farmer *et al.* (2019), suggest exploiting 'sensitive intervention points' to accelerate the post-carbon transition, and include the UK's CCA2008 as a timely SIP with a large effect, corroborated here.

A similar approach could be used to evaluate the extent to which countries met their Paris Accord Nationally Determined Contributions (NDCs), given the appropriate data. The NDCs agreed at COP21 in Paris are insufficient to keep temperatures below 2°C so must be enhanced, and common time frames must be adopted to avoid a lack of transparency in existing NDCs: see Rowan (2019). Since the baseline dates from which NDCs are calculated is crucial, 5-year NDC reviews and evaluation intervals are needed.

In 2019, the UK Government revised its target to one of net zero GHG emissions, entailing no use of coal, oil and natural gas, with no emissions coming from agriculture, construction and waste (currently about 100 Mt p.a.) beyond what can be captured or extracted from the atmosphere. Increases in the capital stock could make the target harder to achieve unless they were carbon neutral. As capital embodies the technology at the time of its construction and is long lived, transition to zero carbon has to be gradual, and necessitates that new capital, and indeed new infrastructure in general, must be zero carbon producing. 'Stranded assets' in carbon producing industries face a problematic future as legislation imposes ever lower CO₂ emissions targets to achieve zero net emissions (see Pfeiffer *et al.*, 2016)).

7.5 Implications of the UK's CO₂ emissions model

Despite more candidate variables than observations, the econometric approach presented in this chapter developed a model to explain the UK's extremely non-stationary CO₂ emissions time-series data over 1860–2017 in terms of coal and oil usage, capital stock and GDP. It was essential to take account of both stochastic trends and distributional shifts. Detection of major policy interventions by indicator saturation estimators yielded a congruent model of CO₂ emissions and accurate forecasts since the CCA2008 came

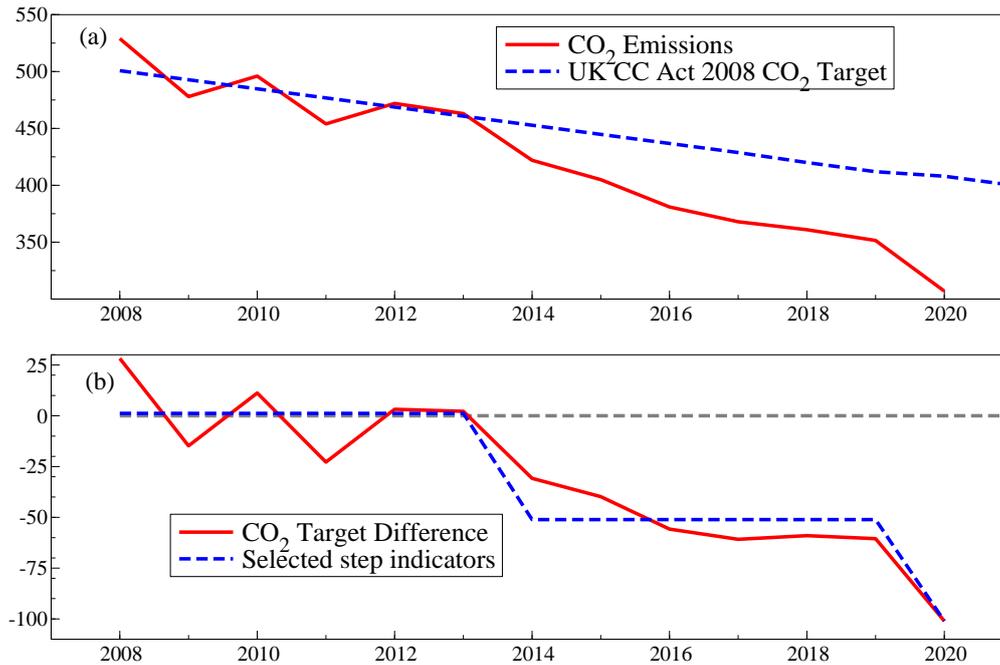


Figure 34: (a) UK CO₂ emissions and CCA2008 CO₂ targets; (b) deviations from the target values with step indicators.

into force. The policy implications highlight that CO₂ emissions are reducing rapidly, but far greater reductions are needed if the UK is to achieve its net-zero emissions target requiring all coal, oil and natural gas to be eliminated or their GHG emissions sequestered. The UK was initially a net CO₂ exporter through embodied CO₂, but is now a net importer, although this component will decrease with falls in the GHG emissions of exporting countries.⁹

The aggregate data provided little evidence of high costs to the large domestic reductions in CO₂ emissions—dropping by 202Mt from 554Mt in 2000 to 352Mt (39%) by end 2019, before the pandemic—whereas real GDP rose by 39% over that period despite the ‘Great Recession’. As new ‘green’ technologies are implemented, careful attention must be paid to local costs of lost jobs: mitigating the inequality impacts of climate induced changes has to be achieved to retain public support.

8 Conclusions

The role of human behaviour in climate change entails that methods developed to model human behaviour in the economic sphere are applicable to modelling climate phenomena as well. There is a two way interaction between the climate and human actions, characterised by wide-sense non-stationary data interacting in complex and non-constant ways. As such, the tools needed to understand these interactions must be able to handle the complex, evolving and shifting interactions over time due to changing human behaviour. In this chapter we outline an approach to modelling such phenomena using time-series econometric tools designed to handle wide-sense non-stationary data from unknown generating processes. The key role that anthropogenic forces play in determining climate can be drawn out by careful modelling of the relationships, embedding our understanding of the climate with economic theory but allowing for data-based search to handle the non-constant distributions. Such an approach allows for testing climate and economic theories, forecasting and policy studies, without contaminating the analyses by unmodelled phenomena. This is essential to provide reliable guidance on how countries can achieve net zero

⁹See <http://www.emissions.leeds.ac.uk/chart1.html> and <https://www.biogeosciences.net/9/3247/2012/bg-9-3247-2012.html>.

emissions to maintain stable global land surface temperatures.

Climate science provides the background to climate econometrics and Section 2 noted the Earth's limited atmosphere and its water resources, establishing that humanity really can alter the climate, and is doing so in many ways. Past climate changes can be related to the 'great extinctions' seen in the geological record, emphasising that it is climate **change** that matters, and the rapidity of change over the most recent past is dramatically faster than any previous changes experienced except perhaps after the meteor impact 65 million years ago. The mid-18th Century Industrial Revolution brought huge benefits, but led to a global explosion in anthropogenic GHG emissions. Emissions are subject to shifts from wars, crises, resource discoveries, technological innovations, pandemics and policy interventions, and the resulting stochastic trends, large shifts and numerous outliers must be handled for viable empirical models of climate phenomena.

Section 3 outlined time-series econometric theory under the assumption of stationarity. Such an assumption is violated by every climate and economics data series imaginable, but is often the premise for climate econometric modelling. Discarding this infeasible setting, Section 4 outlined the implications of shifting distributions for econometric theory. Indicator saturation estimators (ISEs) were described in Section 5, emphasising their key role in robust model discovery. ISEs work well despite creating more candidate variables to select over than observations. A series of simulation experiments highlighted the importance of modelling shifts in both the intercept and the slope parameter. Under the null of no outliers or shifts, there is little loss of efficiency in selecting over T indicators using a tight nominal significance level, even in dynamic models, but the gains under the alternative can be large.

This leads to a more general framework aimed at model **discovery** rather than model building or theory testing, discussed in Section 6. The theory of reduction underpins economic modelling, where the data generating process for the variables under analysis is the target while the theory is retained as the object. As the DGP is unknown, it must be discovered and so automatic model selection is essential. Such discovery needs to be able to handle all the non-stationarities due to outliers, shifting distributions, changing trends, possible non-linearities and omitted variables. Commencing from very general models, models with no losses on reduction are congruent and those that explain rival models are encompassing. Model selection theory poses great difficulties as all statistics for selecting models and evaluating them have distributions, usually interdependent, and are generally altered by every modelling decision. Since congruent and encompassing models are needed we address how they should be selected by discussing *Autometrics*, a multi-path search machine learning algorithm.

The theory for forecasting in a wide-sense non-stationary world is sketched, along with robust forecasting devices including Cardt and a smoothed robust forecasting device which uses localised estimates of the long-run mean and growth rates. One aspect of wide-sense non-stationarity is that the status of regressors can switch from being endogenous to exogenous or vice versa. This was illustrated in a model of past climate variability over the Ice Ages, where a simultaneous-equations system was developed to characterise Antarctic land ice volume, temperature and atmospheric CO₂ levels as non-linear functions of Earth's orbital path round the Sun. The resulting approach clarified the role of CO₂ as a result, rather than a cause of changes in ice volume, with scenario simulations warning of a near ice-free planet with temperatures far above those experienced over the last 800,000 years.

Section 7 updated a model of the UK's CO₂ emissions allowing for two additional years of data. This enabled the UK's Climate Change Act, captured by a step indicator in 2010, to be explicitly modelled, and was shown to matter hugely for the forecast performance of the model. A single equation analysis was developed in four steps to produce a conditional representation. First, impulse and step indicators were selected at very tight significance levels retaining all other regressor variables. Then regressors were selected at looser significance levels. The selected model was solved for the cointegrating (long-run) relation which included step indicators and the non-deterministic terms were reparametrized to differences. Finally, the non-integrated formulation was estimated and used to produce conditional forecasts. A VAR of emissions, coal, oil, GDP and capital stock was constructed to obtain unconditional system forecasts, which were compared with the forecasts from the statistical forecasting device, Cardt.

The importance of modelling in-sample breaks and shifts for forecast performance was emphasised.

The econometric techniques allow for direct linking of climate models with empirical data to further improve econometric research on human responses to climate variability. The approach to jointly addressing all aspects of wide-sense non-stationarity for an unknown data generating process seems most appropriate in climate modelling, where the theory is incomplete, the data are evolving and subject to sudden shifts, there are huge measurement issues, and feedbacks generate non-constant and non-linear relationships. Such an approach should improve our ability to test climate-change mitigation proposals and the role of human behaviour within the climate system, produce more accurate forecasts and scenarios based on differing emissions paths, and provide useful policy analysis to guide the policy response to this most imperative of issues.

References

- Beerling, D. J., E. P. Kantzas and M. R. Lomas *et al.* (2020). Potential for large-scale CO₂ removal via enhanced rock weathering with croplands. *Nature* 583, 242–248. <https://doi.org/10.1038/s41586-020-2448-9>.
- Brock, W. A. and J. I. Miller (2020). Beyond RCP8.5: Marginal mitigation using quasi-representative concentration pathways. WP 1904: <https://econpapers.repec.org/paper/umcwpaper/1904.htm>, Economics Department, University of Missouri.
- Burke, M., W. M. Davis, and N. S. Diffenbaugh (2018). Large potential reduction in economic damages under UN mitigation targets. *Nature* 557, 549–553.
- Burke, M., S. M. Hsiang, and E. Miguel (2015). Global non-linear effect of temperature on economic production. *Nature* 527, 235–239.
- Campos-Martins, S. and D. F. Hendry (2020). Geo-climate, geopolitics and the geo-volatility of carbon-intensive asset returns. Working Paper, Nuffield College, Oxford University.
- Castle, J. L., M. P. Clements, and D. F. Hendry (2015). Robust approaches to forecasting. *International Journal of Forecasting* 31, 99–112.
- (2016). An overview of forecasting facing breaks. *Journal of Business Cycle Research* 12, 3–23.
- (2019). *Forecasting: An Essential Introduction*. New Haven, Ct.: Yale University Press.
- Castle, J. L., J. A. Doornik, and D. F. Hendry (2020). Multiplicative-indicator saturation. Working paper, Nuffield College, Oxford University.
- (2021). Forecasting principles from experience with forecasting competitions. *Forecasting* 3(1), 138–165.
- Castle, J. L., J. A. Doornik, D. F. Hendry, and F. Pretis (2019). Trend-indicator saturation. Working paper, Nuffield College, Oxford University.
- Castle, J. L. and D. F. Hendry (2010). A low-dimension portmanteau test for non-linearity. *Journal of Econometrics* 158, 231–245.
- (2014). Semi-automatic non-linear model selection. In N. Haldrup, M. Meitz, and P. Saikkonen (Eds.), *Essays in Nonlinear Time Series Econometrics*, pp. 163–197. Oxford: Oxford University Press.
- (2020). Climate Econometrics: An Overview. *Foundations and Trends in Econometrics* 10, 145–322.
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28, 591–605.

- Clements, M. P. and D. F. Hendry (1998). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.
- (1999). *Forecasting Non-stationary Economic Time Series*. Cambridge, Mass.: MIT Press.
- Croll, J. (1875). *Climate and Time in Their Geological Relations, A Theory of Secular Changes of the Earth's Climate*. New York: D. Appleton.
- Davis, W. M. (2019). Dispersion of the temperature exposure and economic growth: panel evidence with implications for global inequality. Thesis submitted in partial fulfilment for the MPhil degree, Economics Department, Oxford.
- Dethier, E. N., S. L. Sartain, C. E. Renshaw, and F. J. Magilligan (2020). Spatially coherent regional changes in seasonal extreme streamflow events in the United States and Canada since 1950. *Science Advances* 6 (49), DOI: 10.1126/sciadv.aba5939.
- Doornik, J. A. (2009). Autometrics. In J. L. Castle and N. Shephard (Eds.), *The Methodology and Practice of Econometrics*, pp. 88–121. Oxford: Oxford University Press.
- (2018). *OxMetrics: An Interface to Empirical Modelling* (8th ed.). London: Timberlake Consultants Press.
- (2019). Locally averaged time trend estimation and forecasting. Mimeo, Nuffield College, University of Oxford, Oxford.
- Doornik, J. A., J. L. Castle, and D. F. Hendry (2020a). Card forecasts for M4. *International Journal of Forecasting* 36, 129–134. <https://doi.org/10.1016/j.ijforecast.2019.03.012>.
- (2020b). Short-term forecasting of the coronavirus pandemic. *International Journal of Forecasting*, <https://doi.org/10.1016/j.ijforecast.2020.09.003>.
- (2020c). Statistical short-term forecasting of the COVID-19 pandemic. *Journal of Clinical Immunology and Immunotherapy* 6, DOI: 10.24966/CIIT-8844/1000046.
- Doornik, J. A. and H. Hansen (2008). An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics* 70, 927–939.
- Doornik, J. A. and D. F. Hendry (2018). *Empirical Econometric Modelling using PcGive: Volume I* (8th ed.). London: Timberlake Consultants Press.
- Eglinton, T. I., V. V. Galy and J. D. Hemingway *et al.* (2021). Climate control on terrestrial biospheric carbon turnover. *Proceedings of the National Academy of Sciences* 118(8), <https://doi.org/10.1073/pnas.2011585118>.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987–1007.
- Engle, R. F. and S. Campos-Martins (2020). Measuring and hedging geopolitical risk. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3685213, NYU Stern School of Business, New York.
- Engle, R. F. and D. F. Hendry (1993). Testing super exogeneity and invariance in regression models. *Journal of Econometrics* 56, 119–139.
- Ericsson, N. R. and J. G. MacKinnon (2002). Distributions of error correction tests for cointegration. *Econometrics Journal* 5, 285–318.
- Ericsson, N. R. and E. L. Reisman (2012). Evaluating a global vector autoregression for forecasting. *International Advances in Economic Research* 18, 247–258.
- Estrada, F., P. Perron, and B. Martínez-López (2013). Statistically derived contributions of diverse human influences to twentieth-century temperature changes. *Nature Geoscience* 6, 1050–1055.

- Farmer, J. D., C. Hepburn, M. C. Ives, T. Hale, T. Wetzler, P. Mealy, R. Rafaty, S. Srivastav, and R. Way (2019). Sensitive intervention points in the post-carbon transition. *Science* 364(6436), 132–134.
- Feinstein, C. H. (1972). *National Income, Expenditure and Output of the United Kingdom, 1855–1965*. Cambridge: Cambridge University Press.
- Foote, E. (1856). Circumstances affecting the heat of the sun’s rays. *The American Journal of Science and Arts* 22, 382–383.
- Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica* 46, 1303–1313.
- Hendry, D. F. (1984). Monte Carlo experimentation in econometrics. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics*, Volume 2, Chapter 16, pp. 937–976. Amsterdam: North-Holland.
- (2006). Robustifying forecasts from equilibrium-correction models. *Journal of Econometrics* 135, 399–426.
- Hendry, D. F. and J. A. Doornik (2014). *Empirical Model Discovery and Theory Evaluation*. Cambridge, Mass.: MIT Press.
- Hendry, D. F. and S. Johansen (2015). Model discovery and Trygve Haavelmo’s legacy. *Econometric Theory* 31, 93–114.
- Hendry, D. F. and K. Juselius (2000). Explaining cointegration analysis: Part I. *Energy Journal* 21, 1–42.
- (2001). Explaining cointegration analysis: Part II. *Energy Journal* 22, 75–120.
- Hendry, D. F. and M. Massmann (2007). Co-breaking: Recent advances and a synopsis of the literature. *Journal of Business and Economic Statistics* 25, 33–51.
- Hendry, D. F. and G. E. Mizon (2011). Econometric modelling of time series with outlying observations. *Journal of Time Series Econometrics* 3 (1), DOI: 10.2202/1941–1928.1100.
- (2012). Open-model forecast-error taxonomies. In X. Chen and N. R. Swanson (Eds.), *Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis*, pp. 219–240. New York: Springer.
- (2014). Unpredictability in economic analysis, econometric modeling and forecasting. *Journal of Econometrics* 182, 186–195.
- Hendry, D. F. and F. Pretis (2016). *All Change! The Implications of Non-stationarity for Empirical Modelling, Forecasting and Policy*. Oxford University: Oxford Martin School Policy Paper.
- 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.
- Hepburn, C. and M. Schwarz (2020). Climate change: Answers to common questions. Pictet report, Smith School of Enterprise and the Environment, University of Oxford.
- Hillebrand, E. and T. Proietti (2017). Phase changes and seasonal warming in early instrumental temperature records. *Journal of Climate* 30, 6795–6821.
- Hofer, S., C. Lang and C. Amory *et al.* (2020). Greater Greenland Ice Sheet contribution to global sea level rise in CMIP6. *Nature Communications* 11, <https://doi.org/10.1038/s41467-020-20011-8>.
- Hoffman, P. F. and D. P. Schrag (2000). Snowball Earth. *Scientific American* 282, 68–75.
- Hsiang, S. (2016). Climate econometrics. *Annual Review of Resource Economics* 8(1), 43–75.
- Jouzel, J., V. Masson-Delmotte and O. Cattani *et al.* (2007). Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science* 317, 793–797.

- Kaufmann, R. K. and K. Juselius (2010). Glacial cycles and solar insolation: The role of orbital, seasonal, and spatial variations. *Climate of the Past Discussions* 6, 2557–2591. doi:10.5194/cpd-6-2557-2010.
- Kaufmann, R. K., H. Kauppi, M. L. Mann, and J. H. Stock (2011). Reconciling anthropogenic climate change with observed temperature 1998–2008. *Proceedings of the National Academy of Science* 108, 11790–11793.
- (2013). Does temperature contain a stochastic trend: linking statistical results to physical mechanisms. *Climatic Change* 118, 729–743.
- Kitov, O. I. and M. N. Tabor (2015). Detecting structural changes in linear models: A variable selection approach using multiplicative indicator saturation. Unpublished paper, University of Oxford.
- Krolzig, H.-H. and J. Toro (2002). Testing for super-exogeneity in the presence of common deterministic shifts. *Annales d’Economie et de Statistique* 67/68, 41–71.
- Kulp, S. A. and B. H. Strauss (2019). New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. *Nature Communications* 10, <https://doi.org/10.1038/s41467-019-12808-z>.
- Kurle, J. K. (2019). Essays in climate econometrics. Unpublished M. Phil. Thesis, Economics Department, University of Oxford.
- Levendis, Y. A., G. Kowalski, Y. Lu, and G. Baldassarre (2020). A simple experiment on global warming. *Royal Society Open Science*, <https://doi.org/10.1098/rsos.192075>.
- Lisiecki, L. E. and M. E. Raymo (2005). A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records. *Paleoceanography* 20, doi:10.1029/2004PA001071.
- Lüthil, D., M. Le Floch and B. Bereiter *et al.* (2008). High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature* 453. doi:10.1038/nature06949.
- Makridakis, S., E. Spiliotis, and V. Assimakopoulos (2020). The M4 competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting* 36(1), 54–74.
- Manley, G. (1974). Central England Temperatures: Monthly means 1659 to 1973. *Quarterly Journal of the Royal Meteorological Society* 100, 389–405.
- Martinez, A. B., J. L. Castle, and D. F. Hendry (2021). Smooth robust multi-step forecasting methods. *Advances in Econometrics*, Forthcoming.
- Milankovitch, M. (1969). *Canon of insolation and the ice-age problem*. Washington, D.C: National Science Foundation. English translation by the Israel Program for Scientific Translations of *Kanon der Erdbestrahlung und seine Anwendung auf das Eiszeitenproblem*, Textbook Publishing Company, Belgrade, 1941.
- Mitchell, B. R. (1988). *British Historical Statistics*. Cambridge: Cambridge University Press.
- Noel, L., G. Zarazua de Rubens, J. Kester, and B. Sovacool (Eds.) (2019). *Vehicle-to-Grid: A Sociotechnical Transition Beyond Electric Mobility*. Basingstoke: Palgrave MacMillan.
- Nunes, J., R. Kautzmann, and C. Oliveira (2014). Evaluation of the natural fertilizing potential of basalt dust wastes from the mining district of Nova Prata (Brazil). *Journal of Cleaner Production* 84, 649–656. <https://doi.org/10.1016/j.jclepro.2014.04.032>.
- Ortiz, J. D. and R. Jackson (2020). Understanding Eunice Foote’s 1856 experiments: heat absorption by atmospheric gases. *Royal Society Notes and Records*, <https://doi.org/10.1098/rsnr.2020.0031>.
- Parker, D., T. Legg, and C. Folland (1992). A new daily Central England Temperature series, 1772–1991. *International Journal of Climatology* 12, 317–342.

- Parrenin, F., J.-M. Barnola and J. Beer *et al.* (2007). The EDC3 chronology for the EPICA Dome C ice core. *Climate of the Past* 3, 485–497.
- Parrenin, F., J.-R. Petit and V. Masson-Delmotte *et al.* (2012). Volcanic synchronisation between the EPICA Dome C and Vostok ice cores (Antarctica) 0-145 kyr BP. *Climate of the Past* 8, 1031–1045.
- Pfeiffer, A., R. Millar, C. Hepburn, and E. Beinhocker (2016). The ‘2°C capital stock’ for electricity generation: Committed cumulative carbon emissions from the electricity generation sector and the transition to a green economy. *Applied Energy* 179, 1395–1408.
- Pretis, F. (2019). Econometric models of climate systems: The equivalence of two-component energy balance models and cointegrated VARs. *Journal of Econometrics*, <https://doi.org/10.1016/j.jeconom.2019.05.013>.
- (2021). Exogeneity in climate econometrics. *Energy Economics* 96, <https://doi.org/10.1016/j.eneco.2021.105122>.
- Pretis, F. and D. F. Hendry (2013). Comment on “Polynomial cointegration tests of anthropogenic impact on global warming” by Beenstock *et al.* (2012) – some hazards in econometric modelling of climate change. *Earth System Dynamics* 4, 375–384.
- Pretis, F., J. J. Reade, and G. Sucarrat (2018). Automated general-to-specific (GETS) regression modeling and indicator saturation for outliers and structural breaks. *Journal of Statistical Software* 68, 4, <https://www.jstatsoft.org/article/view/v086i03>.
- Pretis, F., L. Schneider, J. E. Smerdon, and D. F. Hendry (2016). Detecting volcanic eruptions in temperature reconstructions by designed break-indicator saturation. *Journal of Economic Surveys* 30, 403–429.
- Pretis, F., M. Schwarz, K. Tang, K. Haustein, and M. R. Allen (2018). Uncertain impacts on economic growth when stabilizing global temperatures at 1.5°C or 2°C warming. *Philosophical Transactions of the Royal Society A376: 20160460*. <https://doi.org/10.1098/rsta.2016.0460>.
- Qin, H. (2020). Machine learning and serving of discrete field theories. *Science Reports* 10, <https://doi.org/10.1038/s41598-020-76301-0>.
- Ramsey, J. B. (1969). Tests for specification errors in classical linear least squares regression analysis. *Journal of the Royal Statistical Society B*, 31, 350–371.
- Rowan, S. S. (2019). Pitfalls in comparing Paris pledges. *Climatic Change*, <https://link.springer.com/article/10.1007/s10584-019-02494-7>.
- Shehab, A. A., L. Yao and L. Wei *et al.* (2020). The increased hydrocyanic acid in drought-stressed sorghums could be alleviated by plant growth regulators. *Crop and Pasture Science* 71(5), 459–468.
- Siddall, M., E. J. Rohling and A. Almogi-Labin *et al.* (2003). Sea-level fluctuations during the last glacial cycle. *Nature* 423, 853–858.
- Stein, K., A. Timmermann, E. Y. Kwon, and T. Friedrich (2020). Timing and magnitude of Southern Ocean sea ice/carbon cycle feedbacks. *PNAS* 117 (9), 4498–4504. <https://www.pnas.org/content/117/9/4498>.
- Tian, H., R. Xu and J. Canadell *et al.* (2020). A comprehensive quantification of global nitrous oxide sources and sinks. *Nature* 586, 248–256. <https://doi.org/10.1038/s41586-020-2780-0>.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society B*, 58, 267–288.
- Tyndall, J. (1859). Note on the transmission of radiant heat through gaseous bodies. *Proceedings of the Royal Society of London* 10, 37–39.

- Walker, A., F. Pretis, A. Powell-Smith, and B. Goldacre (2019). Variation in responsiveness to warranted behaviour change among NHS clinicians: a novel implementation of change-detection methods in longitudinal prescribing data. *British Medical Journal* 367, 15205.
- White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48, 817–838.
- Woolf, D., J. E. Amonette and F. A. Street-Perrott *et al.* (2010). Sustainable biochar to mitigate global climate change. *Nature Communications* 1(5), <https://doi.org/10.1038/ncomms1053>.

9 Appendix

Data definitions and sources

E_t	=	CO ₂ emissions in millions of tonnes (Mt)	[1], [2].
O_t	=	Net oil usage, millions of tonnes	[3]: 1 tonne = 0.984 imperial tons.
C_t	=	Coal volumes in millions of tonnes	[4].
G_t	=	real GDP, £10 billions, 1985 prices	[5], [7], p.836, [8]a (1993), [10] code:YBHH.
K_t	=	total capital stock, £billions, 1985 prices	[6], [7], p.864, [8]c (1972,1979,1988,1992)
P_t	=	implicit deflator of GDP, (1860=1	[7], p.836, [8]a (1993), [10] code:ABML.
$P_{o,t}$	=	price index, raw materials & fuels	[9]
$\mathbf{1}_{abcd}$	=	impulse indicator equal to unity in year $abcd$	
S_{abcd}	=	step indicator equal to unity up to year $abcd$	
Δx_t	=	$(x_t - x_{t-1})$ for any variable x_t	
$\Delta^2 x_t$	=	$\Delta x_t - \Delta x_{t-1}$	

Sources:

- [1] World Resources Institute <http://www.wri.org/our-work/project/cait-climate-data-explorer>
and <https://www.gov.uk/government/collections/final-uk-greenhouse-gas-emissions-national-statistics>;
- [2] Office for National Statistics (ONS)
<https://www.gov.uk/government/statistics/provisional-uk-greenhouse-gas-emissions-national-statistics-2015>;
- [3] Crude oil and petroleum products: production, imports and exports 1890 to 2015 Department for Business, Energy and Industrial Strategy (Beis);
- [4] Beis and Carbon Brief <http://www.carbonbrief.org/analysis-uk-cuts-carbon-record-coal-drop>;
- [5] ONS <https://www.ons.gov.uk/economy/nationalaccounts/uksectoraccounts#timeseries>;
- [6] ONS <https://www.ons.gov.uk/economy/nationalaccounts/uksectoraccounts/bulletins/capitalstocksconsumptionoffixedcapital/2014-11-14#capital-stocks-and-consumption-of-fixed-capital-in-detail>;
- [7] Mitchell (1988) and Feinstein (1972);
- [8] Charles Bean (from (a) *Economic Trends Annual Supplements*, (b) *Annual Abstract of Statistics*, (c) *Department of Employment Gazette*, and (d) *National Income and Expenditure*);
- [9] UN Statistical Yearbook and Christopher Gilbert;
- [10] ONS, *Blue Book* and *Annual Abstract of Statistics* and *Economic Trends Annual Supplement*.