Global temperatures and green house gases - A common features approach

Li Chen[§], Jiti Gao[*] and Farshid Vahid[*]

Xiamen University, China[§]
Monash University, Australia[*]

February 16, 2019

Abstract

Despite the difficulty in determining the exact nature of trend in global temperatures, we can establish that the trend in global temperatures and green house gases is common. We then estimate the long-run relationship between global temperatures and green house gases and provide theoretically justified measures of confidence in our estimates. We explore the direction of causation in this relationship and compare our results with the existing results in the literature.

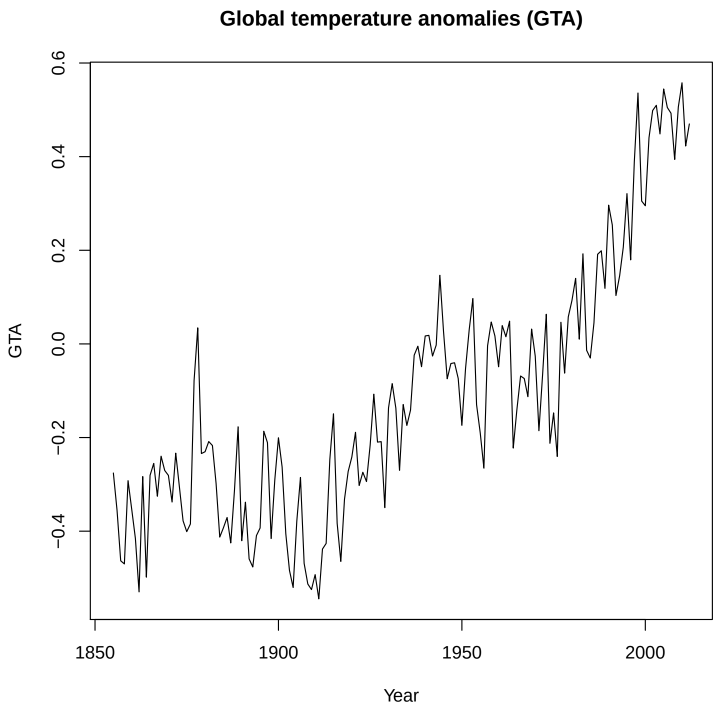
Keywords: Global warming, radiative forcing, common features, stochastic trends, nonlinear trends

1 Introduction

No one understands trends but every one sees it in the data.
Unit roots always cause problems.

from "Laws and Limits of Econometrics", Phillips (2003)

What is the exact nature of trend in global temperatures? It is extremely difficult, if not impossible, to determine the nature of the trend in a time series from a finite sample of observations. In particular it is hard to differentiate deterministic trends from extremely persistent stochastic components. This question has attracted considerable attention in recent years (see Harvey & Mills 2001, Gao & Hawthorne 2006, Magnus et al. 2011, Breusch & Vahid 2011, Estrada et al. 2013, Holt & Teräsvirta 2012, Kaufmann et al. 2013, Werner et al. 2015, Storelvmo et al. 2016, for examples from the econometrics standpoint).

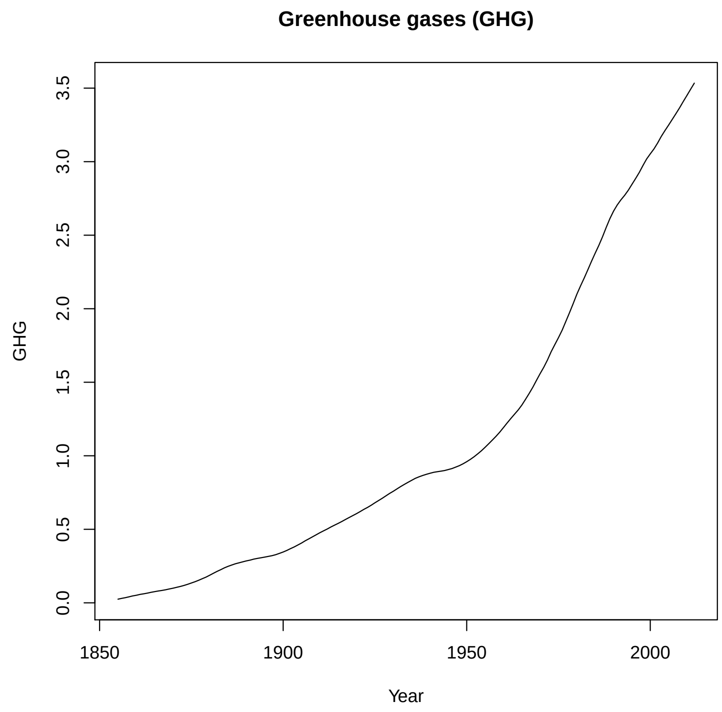


The graph shows one of the commonly used measures of global temperatures[1] compiled by the Climatic Research Institute from 1850 to 2016. Breusch & Vahid (2011) provide a discussion of whether the trend in this data series is deterministic or stochastic, and the difficulty of determining that from the observed data. In particular, they explain that it will be even harder to determine if this series has a stochastic trend against the alternative that it has a broken (piecewise linear) deterministic trend.

The importance of knowing the nature of the trend is that different assumptions about trend lead to different inferential apparatus when we move to the question of relationship between

[1] These data are annual combined land and ocean temperatures. They are temperature anomalies computed as deviations from the 1960-1990 average. This series is the HadCRUT4 series downloadable from https://crudata.uea.ac.uk/cru/data/temperature/

greenhouse gases and global temperatures. Consider the following graph of time series of well mixed green house gases downloaded from NASA:



The graph is much smoother than the graph of global temperatures and one may expect the series to be dominated by a trend with a greater degree of integration (perhaps I(2)). And indeed statistical tests cannot reject the null of a unit root in the first difference of this series. The question is that with all this uncertainty in the nature of the trend in these series, how can we establish if they have a long-run relationship, and if so, how to estimate this relationship and determine a measure of confidence in the estimates.

We approach this issue from a common features perspective (Engle & Kozicki 1993). A *feature* is a dominant statistical property of a variable that cannot be eliminated by linearly transforming that variable[2]. For example, a trend, whether deterministic or stochastic, is a feature. Two series have a *common feature* if each one of them have a certain feature, but a linear combination of them does not have that feature. For example, cointegrated series have a common stochastic trend, and variables that move concurrently with the business cycle have a common cycle (see Vahid & Engle 1993, for a formal definition). Also, nonlinearity can be regarded as a feature as in Anderson & Vahid (1998).

Given that trend is the dominant feature of both global temperatures and the radiative forcing of well-mixed greenhouse gases, we ask the question if there is a linear combination of

[2] A "feature" is a dominant statistical property that satisfies:

- if x_t has (does not have) the feature, then cx_t where c is a non-zero constant will also have (not have) it,
- if x_t and y_t do not have the feature, $x_t + y_t$ will not have it,
- if x_t has the feature and y_t does not have it, then $x_t + y_t$ will have the feature.

these two time series that does not have a trend. That is, do these series have a common trend? If so, can we estimate the linear combination that eliminates the trend, and more importantly, can we provide an interval estimate without determining the nature of the common trend?

2 Common Trends

Consider the following time series representation for global temperature (GTA) and well-mixed greenhouse gases (GHG):

$$GTA_t = f_t + u_t$$
$$GHG_t = g_t + v_t$$

where f and g are deterministic or stochastic trend functions and $(u_t, v_t)'$ is a stationary vector time series. We only require that trend is not mean-reverting. So, it can be a linear or finite order polynomial function of time, it can be an integrated stochastic process of order 1 or larger, it can be a piecewise linear trend, or a combination of these. As such, the trend satisfies the definition of a "feature". We then say that GTA and GHG have a common trend if only if $f_t = \alpha g_t$ for all t because $GTA_t - \alpha\ GHG_t$ will have no trend. It is important to be clear that the possibility that these series may have a common stochastic trend but different deterministic trends is of no interest to us.

Assuming both series have a common trend, the we can write the system as:

$$GTA_t = \alpha GHG_t + e_t \tag{1}$$
$$GHG_t = g_t + v_t$$

where g_t is a trend process. We call this a pseudo-structural system because e_t and v_t are correlated, and hence in the first equation the right hand side variable is endogenous. We first focus on consistent estimation of α in this pseudo-structural system.

Proposition: The powers of t are legitimate instruments that can produce an instrumental variable (IV) estimator of α in equation (1), *regardless of whether the trend (g_t) is deterministic or stochastic.*

That powers of t can be used as instruments in estimating cointegrating relationships is evident in Park & Phillips (1988), and the performance of this IV estimator in the case of I(1) variables has been studies in Phillips & Hansen (1990). Here we propose that this applies more generally to all forms of common trends. The rationale is based on Phillips (1998).

Phillips (1998):

- For some $\delta > 0$, $n^{-\delta} g_t$ has an expansion in powers of t,

$$\frac{g_t}{n^\delta} = \boldsymbol{a}'\boldsymbol{\phi}\left(\frac{t}{n}\right)$$

  where $\boldsymbol{a}$ is an infinite vector and $\boldsymbol{\phi}(\frac{t}{n})$ is an infinite dimensional vector of orthogonal functions of $\frac{t}{n}$.

- Let $\boldsymbol{\phi}_K(\frac{t}{n})$ be the first K element of $\boldsymbol{\phi}(\frac{t}{n})$. In the regression

$$g_t = \hat{\boldsymbol{a}}_K'\boldsymbol{\phi}_K\left(\frac{t}{n}\right) + \hat{r}_t$$

  the usual OLS t and F statistics related to $\hat{\boldsymbol{a}}_K$ diverge.

We take advantage of these "spurious" correlations. Since we have

$$\frac{1}{n}\sum_{t=1}^{n}\boldsymbol{\phi}_K\left(\frac{t}{n}\right)e_t \xrightarrow{p} \mathbf{0}$$

we are able to use $\boldsymbol{\phi}_K\left(\frac{t}{n}\right)$ as instruments to estimate α. While it is true that under the null of common trends even the OLS estimator will be consistent, the importance of the IV estimator is in its over-identifying restrictions.

Suppose we simply use powers of t as instruments. We know (for $K = 3$):

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\begin{pmatrix}\frac{t}{n}\\ \frac{t^2}{n^2}\\ \frac{t^3}{n^3}\end{pmatrix}e_t \xrightarrow{d} N\left[\mathbf{0}, \boldsymbol{S} = \lambda^2\begin{pmatrix}\frac{1}{3} & \frac{1}{4} & \frac{1}{5}\\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6}\\ \frac{1}{5} & \frac{1}{6} & \frac{1}{7}\end{pmatrix}\right]$$

where λ^2 is the long-run variance of e_t. We show that as long as the trend is more than linear, the test statistic

$$J(\hat{\alpha}(\hat{\boldsymbol{S}}^{-1}), \hat{\boldsymbol{S}}^{-1}) \xrightarrow{d} \chi^2_{K-1}$$

when trends are common. This gives us a way to obtain a confidence interval for α by inverting this statistic. More specifically, by changing α while keeping the weight matrix $\hat{\boldsymbol{S}}^{-1}$ fixed, we can find the set of α where the value of the GMM objective function is below the critical value of χ^2_{K-1} at the desired level of significance.

While this distribution is correct under the null, we need to pay attention to some important considerations. The value of λ^2 does not play a role in the IV estimator, and the GMM estimator here is the 2SLS estimator. However, the estimate of λ^2 is required for the computation of the test for overidentifying restrictions. Unlike the usual GMM context where variables are considered to be stationary and ergodic, which implies that the errors will always be stationary and ergodic, in our case the errors are stationary and ergodic only under the null. Under the alternative, the errors will be trending, and their long run variance is infinite. This highlights the importance of careful choices of K (number of instruments) and appropriate calculation of λ^2.

The number of functions of t must be sufficiently large to create a coordinate system that approximates the trends in each variable "sufficiently well". For our purpose, we require that the remainders are stationary. This can be easily investigated in practice by increasing K such that the residuals of regressions of GTA and GHG on these functions are considered stationary by a Dickey-Fuller test. While Dickey-Fuller test has low power in distinguishing between stochastic and deterministic trends, the rejection of it will provide sufficient evidence for stationarity.

When K is determined in this way, then the trend in any linear combination of the two variables is likely to be well approximated by these K instruments. Only if the null of common trends is true, then there is a combination that has no correlation with these K functions. Hence, if we take the IV residuals and regress them on these K instruments, the remainder of this regression will always be stationary, regardless of whether the null hypothesis is true or not, and if the null is true, the long-run variance of this remainder will be a consistent estimator of the long-run variance of the true errors $\{e_t\}$ in (1).

Another consideration that makes the theoretical derivations easier and has no effect on empirical results is that rather than powers of t, we use orthogonal polynomials of $\tau = \frac{t}{n}$. In particular we use the shifted *Legendre polynomials* defined on [0,1] that are normalised such that they are orthonormal, i.e.,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \Phi_k(\tau_t)\Phi_k(\tau_t)' = \int_0^1 \Phi_k(\tau)\Phi_k(\tau)' d\tau = I_k, \tag{2}$$

for $\tau_t = t/n$. The first 5 polynomials are given below and are plotted in Figure 1.

$$\begin{aligned}
\phi_0(\tau) &= 1,\\
\phi_1(\tau) &= \sqrt{3}(2\tau_t - 1),\\
\phi_2(\tau) &= \sqrt{5}(6\tau_t^2 - 6\tau + 1),\\
\phi_3(\tau) &= \sqrt{7}(20\tau_t^3 - 30\tau_t^2 + 12\tau_t - 1),\\
\phi_4(\tau) &= \sqrt{9}(70\tau^4 - 140\tau^3 + 90\tau^2 - 20\tau + 1),\\
\phi_5(\tau) &= \sqrt{11}(252\tau^5 - 630\tau^4 + 560\tau^3 - 210\tau^2 + 30\tau - 1),\\
\phi_6(\tau) &= \sqrt{13}(924\tau^6 - 2772\tau^5 + 3150\tau^4 - 1680\tau^3 + 420\tau^2 - 42\tau + 1).
\end{aligned}$$

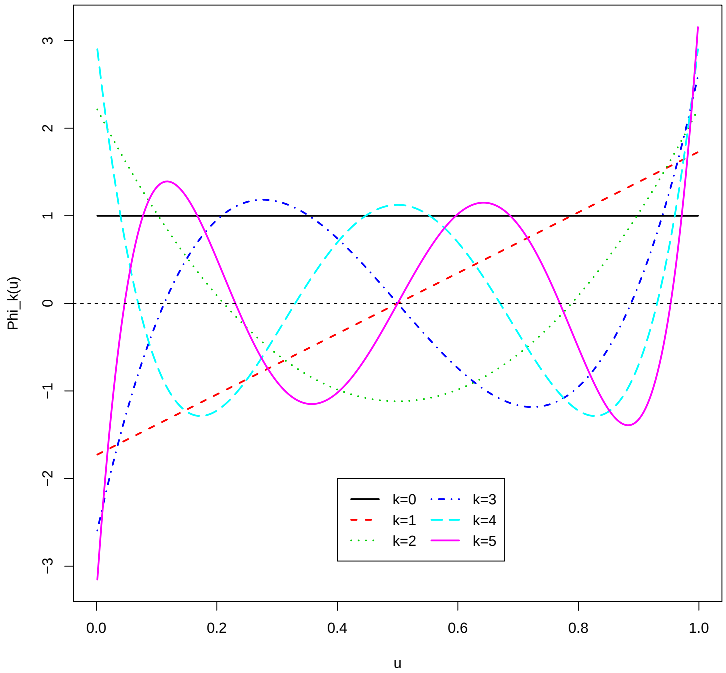


Figure 1: Normalized Legendre polynomials up to the fifth order

3 Monte Carlo Simulations

We consider the data generating process

$$\begin{aligned} y_t &= w_{1t} + u_t \\ x_t &= (1-q) \times w_{1t} + q \times w_{2t} + v_t \\ w_{it} &= w_{it-1} + \varepsilon_{it},\ i = 1, 2 \\ u_t &= \rho u_{t-1} + \eta_t \\ \{\varepsilon_{1t}, \varepsilon_{2t}, \eta_t, v_t\} &\sim i.i.d.N(0,1) \end{aligned}$$

where we let $\rho = 0.1$ and $n = 200, 400, 800$, respectively. We use the power functions $(t/n)^j$ as instruments for $j = 0, 1, 2, ..., k$ in which $k = 3, 5, 7$ respectively. When $q = 0$, x_t and y_t share a common stochastic trend, while $q > 0$, the trends in them could not be canceled out for any linear combinations of them. We repeat the simulation process for 10000 times and report the percentages of rejecting the null hypothesis of *common trend* respectively under 5% significance level in the Table below.

	$\text{Prob}(J_n > \chi^2_{k-1}(0.95))$								
	$k=3$			$k=5$			$k=7$		
q	200	400	800	200	400	800	200	400	800
0.00	0.0621	0.0638	0.0615	0.0649	0.0716	0.0693	0.0616	0.0723	0.0690
0.01	0.0864	0.1634	0.3684	0.0796	0.1578	0.3803	0.0795	0.1505	0.3703
0.02	0.1521	0.3519	0.6426	0.1412	0.3662	0.7051	0.1340	0.3537	0.7051
0.05	0.4306	0.7148	0.9052	0.4601	0.7880	0.9613	0.4478	0.8088	0.9781
0.10	0.7096	0.8998	0.9715	0.7789	0.9583	0.9946	0.7892	0.9738	0.9982

In the simulation, we let the two time series share a common trend w_{1t} by setting $q = 0$. Under 5% significance level, the probability that the J-test rejects the null hypothesis of the existence of common trend is only $6\% \sim 7\%$. Hence, the test does not suffer from size distortions.As we gradually increase q, an extra stochastic trend w_{2t} enters x_t with growing influence. The power of the common trend test reaches a satisfactory level when tje extra trend becomes only as 1/9 of the scale of the original common trend component.

4 Empirical results

We begin our empirical example by determining the number of basis functions that can approximate the trends of GTA and GHG in our sample to the extent that the remainder is stationary. The following table summarises the results of this stage of analysis.

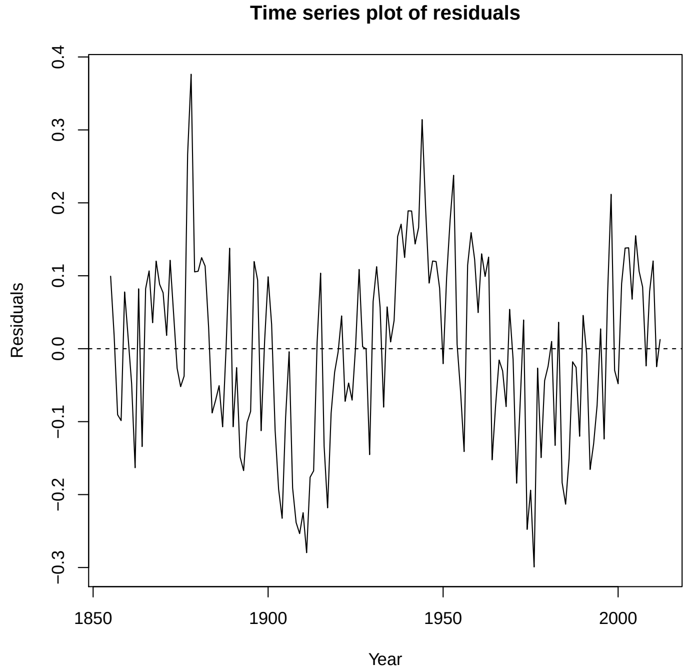
Table 1: Determination of the number of instruments

	trend specification	residuals
GTA	$\alpha_0 + \alpha_1 t$	I(1)
	$\alpha_0 + \alpha_1 t + \alpha_2 t^2$	I(0)
	$\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3$	I(0)
GHG	$\alpha_0 + \alpha_1 t$	I(2)
	$\alpha_0 + \alpha_1 t + \alpha_2 t^2$	I(2)
	$\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3$	I(0)

The table shows that $K = 3$ would be sufficient for both series. The 2SLS estimate of the relationship between global temperature anomalies and green house gases using t, t^2 and t^3 as instruments produces

$$GTA_t = -0.384 + 0.240 GHG_t + \widehat{e}_t$$

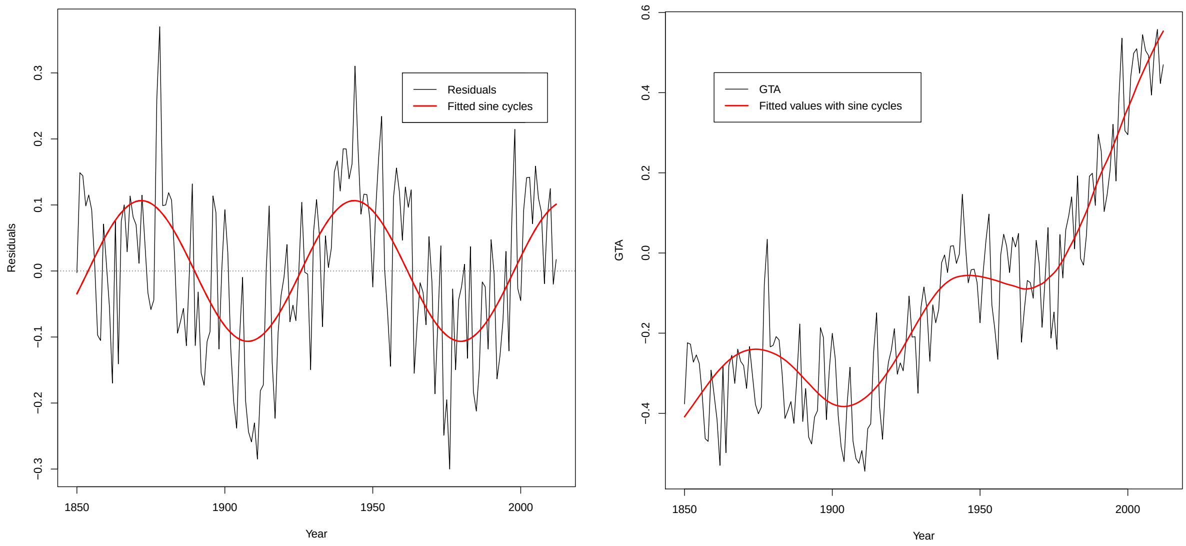
The p-value of the Sargan test of overidentifying restrictions is 0.4. These results show that regardless of what the trend is, there is a strong evidence that global temperatures and greenhouse gases share a common trend. As the following table shows these IV estimates are close to the OLS estimates and change very little as we increase K



The plot of the residuals of this relationship shows a very regular pattern. We add a single periodic function of time to the specification of equation (1) and we obtain the following:

$$\widehat{GTA}_t = -0.399 + 0.249GHG_t + 0.113\sin(-6.086 + 4.453\frac{\pi t}{T})$$

This corresponds to a cycle with a period of 70 years, which is almost half of the sample size. While this is arguably too long to believe that it can be reliably estimated from such a relatively short sample, it is within the range of the Atlantic Multidecadal Oscillation (AMO) that several climate scientists have documented using 8000 years of climate proxy records (corals, three rings, ice cores). See, for example, Knudsen et al. (2011). The fitted values for GTA are plotted in the following figure.



(a) Fitted Sine cycles

(b) Fitted values of GTA

Alternatively, the dynamics in the residuals can be modelled as an ARMA process. The specification that fits these best is

$$GTA_t = -0.391 + 0.243GHG_t + \widehat{e}_t \tag{3}$$

$$\widehat{e}_t = 0.499\widehat{e}_{t-1} + 0.216\widehat{e}_{t-4} + \widehat{\varepsilon}_t \tag{4}$$

which is similar to the time series specification found in Breusch & Vahid (2011).

The question of how useful such a pseudo-structural model is in forecasting temperature into the future. The best one can do is to provide conditional (scenario) forecasts: how GTA will evolve for a given GHG trajectory. The validity of such an exercise relies on the ability to keep GHG on the given trajectory. If there is strong feedback from the current GTA to future GHG, then it is unlikely that the trajectory of GHG can be controlled exogenously. With that caveat, we perform a small pseudo-out-of-sample conditional forecasting exercise. We estimate the models using data from 1855 to 1982, then evaluate them based on their forecasts of GTA in the last 30 years given the observed GHG trajectory. Of course this exercise suffers from data snooping and not surprisingly, the forecasts of the trend plus periodic deterministic cycle performs better than other forecasts. The conclusion that we want to make from this exercise is that the simple time series model can produce credible prediction intervals for up to 30 year ahead conditional forecasts.

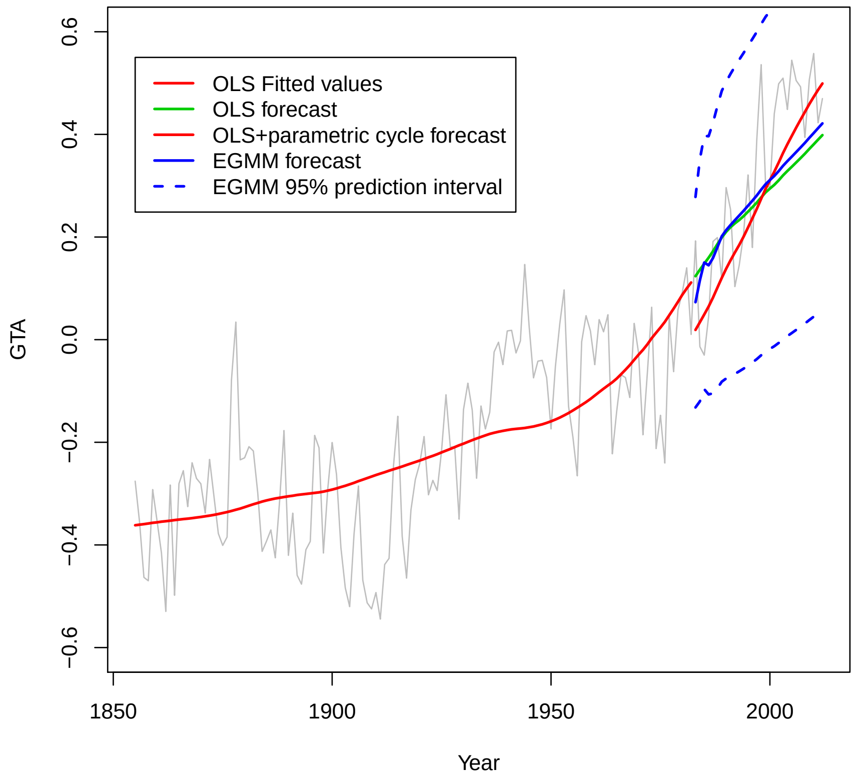


Figure 3: Comparison on out-of-sample forecasts

Engle and Granger (1987) established that I(1) variables that are cointegrated have an error correction representation. While in co-trending relationship, it is still reasonable to believe that at least one of ΔGTA_t or ΔGHG_t, or both, must respond to previous equilibrium error $GTA_t - \beta GHG_t$.

$$\widehat{\Delta GTA}_t = -\underset{(0.0916)^{***}}{0.3119}(GTA_{t-1} - 0.2339GHG_{t-1}) + controls \tag{5}$$

$$\widehat{\Delta GHG}_t = -\underset{(0.0022)}{0.0003}(GTA_{t-1} - 0.2339GHG_{t-1}) + controls \tag{6}$$

where $controls$ are the lagged values of ΔGTA_t and ΔGHG_t. The non-response of GHG to the equilibrium error is a necessary condition for exogeneity of GHG.

5 Conclusion

We establish a coordinate system consists of orthogonal basis functions of polynomials, and propose estimation and inference methods for common trends among nonstationary time series regardless of the nature of trends. The over-identification test could also be used to construct confidence intervals for the common trend coefficients. Simulation results indicate that the over-identification test performs quite well for both deterministic and stochastic trends. We also find common trend in the global temperature and green house gases representing a long-run relationship between the two time series. We also find evidence of regular cycles in global temperatures compatible with the 70 year North Atlantic Multidecadal Oscillations. If such cycles really exist, our estimation tells us that we are at the peak of such cycle at the moment, and even if green house gases continue to grow at the current rate, we are likely to see a flattening of the temperatures as the cycle goes from its peak to its trough, and disguises the underlying trend. This might mean that it may become even harder to convince the median voter of the reality of global warming.

6 References

References

Anderson, H. M. & Vahid, F. (1998), 'Testing multiple equation systems for common nonlinear components', *Journal of Econometrics* **84**, 1–36.

Breusch, T. & Vahid, F. (2011), 'Global temperature trends', *Department of Econometrics and Business Statistics Working Paper 4-11* .
URL: *https://ideas.repec.org/p/msh/ebswps/2011-4.html*

Engle, R. F. & Kozicki, S. (1993), 'Testing for common features', *Journal of Business & Economic Statistics* **11**, 369–380.

Estrada, F., Perron, P. & Martínez-López, B. (2013), 'Statistically derived contributions of diverse human influences to twentieth-century temperature changes', *Nature Geoscience* **6**, 1050–1055.

Gao, J. & Hawthorne, K. (2006), 'Semiparametric estimation and testing of the trend of temperature series', *The Econometrics Journal* **9**, 332–355.

Harvey, D. I. & Mills, T. C. (2001), 'Modelling global temperature trends using cointegration and smooth transitions', *Statistical Modelling* **1**, 143–159.

Holt, M. T. & Teräsvirta, T. (2012), 'Global hemispheric temperature trends and coshifting: A shifting mean vector autoregressive analysis', *CREATES Research Paper 2012-54* .

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

Knudsen, M. F., Seidenkrantz, M.-S., Jacobsen, B. H. & Kuijpers, A. (2011), 'Tracking the Atlantic Multidecadal Oscillation through the last 8,000 years', *Nature Communications* **2**, 178.
URL: *http://dx.doi.org/10.1038/ncomms1186*

Magnus, J. R., Melenberg, B. & Muris, C. (2011), 'Global warming and local dimming: The statistical evidence', *Journal of the American Statistical Association* **106**, 452–464.

Park, J. Y. & Phillips, P. C. B. (1988), 'Statistical inference in regressions with integrated processes: Part 1', *Econometric Theory* **4**(3), 468–497.

Phillips, P. C. B. (1998), 'New tools for understanding spurious regressions', *Econometrica* **66**, 1299–1325.

Phillips, P. C. B. (2003), 'Laws and limits of econometrics', *The Economic Journal* **113**, C26–C52.

Phillips, P. C. B. & Hansen, B. (1990), 'Statistical inference in instrumental variables regression with I(1) processes', *Review of Economic Studies* **57**(1), 99–125.

Storelvmo, T., Leirvik, T., Lohmann, U., Phillips, P. C. B. & Wild, M. (2016), 'Disentangling greenhouse warming and aerosol cooling to reveal earth's climate sensitivity', *Nature Geoscience* **9**, 286–289.

Vahid, F. & Engle, R. F. (1993), 'Common trends and common cycles', *Journal of Applied Econometrics* **8**, 341–360.

Werner, R., Valev, D., Danov, D. & Guineva, V. (2015), 'Study of structural break points in global and hemispheric temperature series by piecewise regression', *Advances in Space Research* **56**, 2323–2334.
URL: *http://dx.doi.org/10.1016/j.asr.2015.09.007*