Is Radiative Forcing Cointegrated with Temperature? A Further Examination Using a Structural Time Series Approach

March 2019

Kelvin Balcombe
University of Reading

Iain Fraser
University of Kent

Abhijit Sharma*
Bradford University School of Management

Abstract

Purpose

This paper re-examines the long-run relationship between radiative forcing (including emissions of carbon dioxide, sulphur oxides, methane and solar radiation) and temperatures from a structural time series modelling perspective. We assess whether forcing measures are cointegrated with global temperatures using the structural time series approach.

Design/methodology/approach

A Bayesian approach is used to obtain estimates that represent the uncertainty regarding this relationship. Our estimated structural time series model enables alternative model specifications to be consistently compared by evaluating model performance.

Findings

Our results confirm that cointegration between radiative forcing and temperatures are consistent with the data. However, our results find less support for cointegration between forcing and temperature data than found previously.

Research limitations/implications

Given considerable debate within the literature relating to the 'best' way to statistically model this relationship and explain results arising as well as model performance, there is uncertainty regarding our understanding of this relationship and resulting policy design and implementation. There is a need for further modelling and use of more data.

Practical implications

There is divergence of views as to how best to statistically capture, explain and model this relationship. Researchers should avoid being too strident in their claims about model performance and better appreciate the role of uncertainty.

Originality/value

The results of this study make a contribution to the literature by employing a theoretically motivated framework in which a number of plausible alternatives are considered in detail, as opposed to simply employing a standard cointegration framework.

Key Words: Radiative forcing, cointegration, structural time series.

***Corresponding Author:** Bradford University School of Management, Emm Lane, Bradford BD9 4JL, UK.
Email: A.Sharma12@bradford.ac.uk, Tel: +44 (0)1274 234781.

Is Radiative Forcing Cointegrated with Temperature? A Further Examination Using a Structural Time Series Approach

March 2019

Abstract

Purpose

This paper re-examines the long-run relationship between radiative forcing (including emissions of carbon dioxide, sulphur oxides, methane and solar radiation) and temperatures from a structural time series modelling perspective. We assess whether forcing measures are cointegrated with global temperatures using the structural time series approach.

Design/methodology/approach

A Bayesian approach is used to obtain estimates that represent the uncertainty regarding this relationship. Our estimated structural time series model enables alternative model specifications to be consistently compared by evaluating model performance.

Findings

Our results confirm that cointegration between radiative forcing and temperatures are consistent with the data. However, our results find less support for cointegration between forcing and temperature data than found previously.

Research limitations/implications

Given considerable debate within the literature relating to the 'best' way to statistically model this relationship and explain results arising as well as model performance, there is uncertainty regarding our understanding of this relationship and resulting policy design and implementation. There is a need for further modelling and use of more data.

Practical implications

There is divergence of views as to how best to statistically capture, explain and model this relationship. Researchers should avoid being too strident in their claims about model performance and better appreciate the role of uncertainty.

Originality/value

The results of this study make a contribution to the literature by employing a theoretically motivated framework in which a number of plausible alternatives are considered in detail, as opposed to simply employing a standard cointegration framework.

Key Words: Radiative forcing, cointegration, structural time series.

1. Introduction

There is an emerging consensus that statistical evidence supports the relationship between radiative forcing measures (e.g. carbon dioxide, sulphur oxides, methane, solar radiation) and temperatures. In this paper we follow the established definition of radiative forcing or forcing following, for exmaple, Kaufmann et al. (2010: 398) who define this as "the forcing ... due to carbon dioxide, methane, CFC11 [chlorofluoro-carbons], CFC12, nitrous oxide, sulfur emissions, and solar activity". The particular statistical framework used by Mills (2009), Kaufmann and Stern (2002) and Kaufmann et al. (2006) is that of cointegration. Cointegration between two variables implies that the variables have stochastic trends, but a linear combination between the variables exists that has no stochastic trend. Equivalently, the existence of cointegration between two variables implies that they share a common stochastic trend. Where one of the variables is weakly exogenous, this variable may be causally responsible for the trend in the other. Thus, the implication that radiative forcing is cointegrated with temperature provides evidence consistent with some scientific models that imply forcing measures play a possible positive causal role in relation to warming trends.

While the earlier work of Stern and Kaufmann (2000) employed structural time series models, more recent research has employed conventional tests for cointegration, within an autoregressive framework. Our aim is to re-examine whether forcing measures are cointegrated with global temperatures using the structural time series approach. We do not dispute the methodological rigour or specific findings of the studies above. On the contrary, the previous finding of cointegration between temperatures and global warming are easily replicated. However, there continues to be considerable interest in examining the statistical properties of causal relationships between global temperatures and human activity. For example, there is an ongoing debate as to whether global temperatures are stationary or best represented by some other more complex process. Recent econometric evidence on this is provide by Lai and Yoon (2018). There have also been studies that suggest that the length of dataset in this context matters. For example, McMillana and Wohar (2013) report a weak relationship between temperature and CO2 and no statistically significant evidence of a trend when employing a much longer time series of data. At the same time there are studies, for example, Stern and Kaufmann (2014), that do report causal relationships between certain types of forcing (e.g. natural and anthropogenic) and temperature change. Stern and Kaufmann (2014) arrive at these conclusions by employing Granger causality tests as opposed to developing a time series model of the relationship between the variables of interest. They argue that time series models are dependent on assumptions regarding the time series properties of the data and as is well documented in the literature there is far from a consensus of opinion on this issue. This in part stems from the fact that it is hard (impossible) to know the underlying data generating process which then makes model selection difficult. This has, however, not stopped further developments in this context. For example, an alternative approach to examining this issue is presented by Gallegati (2018) who use wavelet analysis. This approach identifies that different data series can have different time scales that is only partially resolved when employing cointegration analysis because of how the methods deals with non-stationarity arising from stochastic trends. In addition, although our study is conducted at global level there are links with research conducted at the country and regional levels. For example, time series data and methods have been employed in empirical studies of the Environmental Kuznets Curve (EKC) by Cialani (2007) for Italy and Mohapatra and Giri (2009) for India. There is good reason to think that such studies consider the type of advanced econometric methodology employed in the current paper.

Another example of an econometric development applied to this topic is Chevillon (2017) who employs a procedure that offers a robust test for the rank of cointegration within a VAR that may have misspecified local linear trends. Using this approach it is reported that temperature and greenhouse gases appear to be cointegrated. This paper also provides an overview of the proceeding literature that once again illustrates the ongoing debate regarding the extent to which statistical models can truly reveal the relationships of interest.

However, we believe it is worth investigating how robust previous findings are to alternative model specifications. These lead us to our main research question which is investigating the presence or otherwise of a stable long run relationship between radiative forcing and global temperature by employing classical and Bayesian methods, and explicitly considering alternative model specifications, both cointegrated and non-cointegrated, which is a departure from the prior literature and makes a contribution to the literature. We do this in three ways.

First, we conduct cointegration tests introduced by Shin (1994) which adopt cointegration as the null hypothesis rather than the alternative hypothesis as is the case on other empirical work carried our so far. Tests that adopt a the presence of a unit root or no-cointegration as the null hypothesis have commonly been found to obtain different findings to tests that have a null hypothesis of stationarity or cointegration respectively (see Maddala and Kim, 1998 Chap 4). Therefore, we believe it would be useful to investigate whether Shin's (1994) approach supports previous findings concerning cointegration.

Second, by explicitly estimating a structural time series model, alternative model specifications (i.e. cointegrated and non-cointegrated) can be consistently compared by evaluating model performance. Structural time series models are particularly useful for this purpose as they can nest both cointegrated and non-cointegrated models as special cases.

Third, we estimate structural time series models using Classical and Bayesian methods. This dual approach to estimating the structural time series model is revealing. Whereas the Classical approach to estimation will be based on only one mode of the likelihood, the Bayesian approach to inference can reflect multiple high density points. As we will explain, there is strong evidence that the posterior density has a number of high density points. This requires us to place important qualifications on the results we report.

Overall, the results of our study make a contribution to the literature by employing a theoretically motivated framework in which a number of plausible alternatives are considered in detail, as opposed to simply employing a standard cointegration framework. Our research fits in well within the context of research such as Romero-Avila (2008) who examines convergence within carbon dioxide emissions for 23 countries between 1960-2002, Lee and Chang (2009) who investigate stochastic convergence of per capita carbon dioxide emissions and multiple structural breaks for OECD countries, as well as Ajmi et al. (2013) who study relationships between energy consumption and income for G7 countries using nonlinear causality tests. Marrero (2008) considers global greenhouse gas emissions within his study of emissions, growth and energy usage mix for Europe. Anger (2008) generalises this type of analysis by considering the economic impact of emissions trading schemes and likely impact on emissions. Our research approach helps provide an important backdrop to studies in related areas such as decomposition of carbon dioxide emissions (Sun, 1999) and studies of non-CO_2 greenhouse gas emissions (e.g. Shukla et al, 2006). Our research also links with studies outlining policies aimed at reducing CO_2 emissions such as Gerlagh and Zwaan (2006), Lu et al. (2013) who consider CO_2 emission efficiency in OECD countries, as well as in developing countries facing issues such as poverty alleviation and growth promotion (Van Heerden et al, 2006). Sam et al. (2009) study the effectiveness of voluntary emissions programmes in the US. Our paper extends work such as that on emission in the US and evidence on convergence patterns for pollutants using unit root tests (see List, 1999).

Our paper proceeds by outlining the statistical models we employ in Section 2. Section 3 describes our approach to model estimation. In Section 4 we briefly discuss the data and present our empirical results. Section 5 provides a discussion of our results and their implications. Finally, in Section 6 we offer conclusions.

2. Econometric Models

As our preceding discussion suggests, there is considerable and ongoing debate with regard to choice of econometric models and use of an appropriate empirical strategy in order to study our main relationships of interest (e.g. Chevillon (2017), Lai and Yoon (2018), Stern and Kaufmann (2014), Gallegati (2018) and McMillana and Wohar (2013)). Based on our consideration of the empirical literature we believe that our results make a contribution to the literature due to our use of a theoretically motivated framework in which a number of plausible alternatives are considered in detail, as opposed to simply employing a basic cointegration framework, as outlined in detail below. As a result, we mainly limit our discussion to the structural time series approach. We also employ a standard vector autoregressive (VAR) approach, where cointegration is treated as the alternative hypothesis so as to ensure that, should our results radically differ from previous findings, this would be due to the modelling approach adopted and not driven by slight differences in the data employed in the analysis. For details on the VAR approach readers are referred to Johansen (1995).

The model introduced in Shin (1994) is of the structural time series form:

$$\begin{aligned} y_t &= \mu_t + \beta_t + x_t\alpha + e_t \\ \mu_t &= \mu_{t-1} + v_t \end{aligned} \tag{1}$$

where y_t is temperature at time t, x_t is a $m \times 1$ vector of covariates (in this case radiative forcing) and e_t and v_t are stationary innovations that can be serially correlated. The model in equation (1) above contains a time trend (t), but this can be removed from the regression if there is no deterministic trend in the data generating process. If y_t and x_t are integrated of order 1 (see Johansen, 1995, p35), cointegration between y_t and x_t implies the variance of v_t is zero.

The tests outlined in Shin (1994) do not require explicit estimation of the variance components within equation (1). A test for cointegration can be constructed by obtaining estimates of the long run variance of e_t (ie, ω_e) and then constructing the following test statistic

$$C = T^{-2} \sum_{t=1}^{T} S_t^2 / \omega_e \tag{2}$$

where S_t is the estimate of the partial sum process $S_t = \sum_{i=1}^{t} e_i$. The distribution of this test statistic has been tabulated in Shin (1994), but it can be simulated using Monte Carlo methods.

The model in equation (1) can be generalised to allow for autoregressive components with a local linear trend intercept as follows:

$$\begin{aligned} y_t &= \mu_t + \sum_{i=1}^{p} \gamma_i y_{t-i} + x_t' \alpha + e_t \\ \mu_t &= \mu_{t-1} + \beta_t + v_t \\ \beta_t &= \beta_{t-1} + w_t \end{aligned} \tag{3}$$

where y_t and x_t' are as defined above. Because the autoregressive components are assumed to "soak up" any serial correlation, e_t, v_t and w_t are assumed to be independent normal innovations.[1] The intercept in this model μ_t is able to evolve in a stochastic manner if either v_t or w_t have non-zero variances. The trend in the intercept at time t is β_t. Cointegration between y_t and x_t requires that both are non-stationary and that

$$\left| \sum_{i=1}^{p} \gamma_i \right| < 1 \text{ and } Var\left(v_t\right) = Var\left(w_t\right) = 0. \tag{4}$$

The unknowns within model (1) are of two types:

i) the 'latents' $\Gamma = \left(\{\mu_t\}, \{\beta_t\}, \{\gamma_i\}, \alpha\right)$ (along with the errors that can be constructed given knowledge of these quantities); and

ii) the 'hyper parameters' $\Psi = \left(\sigma_e^2, \sigma_v^2, \sigma_w^2\right)$.

Additionally, there are initial conditions (priors) for the latents Γ_0 which are the prior mean and covariance for the value of the latents at $t = 0$.

3. Model Estimation

3.1. Classical Estimation

The test introduced by Shin (1994) only requires a standard ordinary least squares regression to estimate the null (cointegrated) model. Alternatively, an estimator that allows for serial correlation in the error and exogeneity can be employed, such as the Fully Modified (FM) estimator outlined in Phillips and Hansen (1991). The FM estimator is employed here, since it may yield less biased and more efficient estimates, and it also requires the component ω_e to be estimated. Therefore, the test statistic described in Section 2 only needs the additional construction of the partial sum component.

Classical estimation of the general model described by equation (3) can proceed in a number of ways. Harvey (1989) outlines Classical approaches in detail. For example, the 'time domain' approach outlined in Harvey (1989) employs the Kalman Filter, that enables the likelihood to be calculated using the prediction error decomposition. Using

[1] We also initially incorporated errors of a moving average nature, but found no significant correlation of this form, having allowed for lagged dependents in our covariates.

this approach, the likelihood is expressed as a function of the hyper parameters and the priors for the latent components only.[2] This likelihood is denoted here as $L(\{y_t\}, \Gamma_0, \Psi)$.

Classical estimation usually proceeds by finding the estimated value of $\hat{\Psi}$ that maximises the likelihood. Subject to regularity conditions, inference about the parameters Ψ can then be performed using likelihood ratio, Wald or Lagrange Multiplier tests. Likelihood ratio and Wald tests have distributions that are non-standard (see Harvey, 1989, p.234). The estimates of the latents (along with their (co)variances) can be obtained (at $\hat{\Psi}$) by the Kalman Smoother.

3.2. Bayesian Estimation

Bayesian inference uses the posterior distribution of the parameters. Unlike the Classical approach all the parameters are treated as random variables. The likelihood is therefore viewed as the density of the data conditional on these parameters. For example, the likelihood above, can be denoted as a marginalised likelihood $f(\{y_t\}|\Psi; \Gamma_0)$. Consequently, using the Bayes theorem implies that the prior distribution for Ψ is $f(\Psi)$. It then follows that the posterior distribution is $f(\Psi|\{y_t\}; \Gamma_0) \propto f(\{y_t\}|\Psi; \Gamma_0) f(\Psi)$. Providing this posterior can be mapped, Bayesians will report the mean and variance of the posterior distribution as point estimates.

Bayesian estimation with 'flat' priors, delivers a posterior density that is, over a certain range, approximately proportional to the likelihood. Therefore, Bayesian inference can often give results that are similar to those derived using maximum likelihood. However, in some situations, Bayesian and Classical estimates may diverge. For example, if there are two distinct local maximums for the likelihood, then there may be two distinct parameterisations of the model that equally well represent the sample information. Unlike Classical procedures, Bayesian inference is not based on the behaviour of the likelihood function locally around a single point where it has been maximised. From a Bayesian perspective, the values of the parameters at the maximum of the likelihood, and the curvature of the likelihood at that point, do not fully reflect the sample information. Should the likelihood be multimodel all high density points are reflected in the final estimates (the mean and variance of the posterior distribution). Just as importantly, however, we can examine the entire posterior distribution of key parameters to learn about the data generating process.

An introduction to a Bayesian approach to estimating structural time series models is presented in Koop (2003). Unlike the approach outlined in Koop, we map the posterior distribution for the parameters of the hyper parameters Ψ using a random walk

[2] The latents have been integrated out of the likelihood function as opposed to the 'concentrated' or 'profile' likelihood. See Harvey (1989) for details.

Metropolis-Hastings algorithm (see Koop 2003, p.92). This is a simple and efficient computational tool for the posterior distribution given that that the number of hyper parameters are few. Given draws of Ψ from the posterior distribution of the hyper parameters, the smoothed estimates of the latents can then be generated using the Kalman Smoother along with an estimate of the covariance matrix for the latents. This could be done by simply plugging in a point estimate $\hat{\Psi}$ and obtaining conditional estimates of the latents at that value. However, using the Kalman Smoother to generate the latents in this way is not fully Bayesian, because a fully Bayesian estimate of the latents would embody the parameter the posterior uncertainty (variability) in the estimates of the latents (Ψ). A fully Bayesian approach requires a draw for each of the latents which needs to be made for every posterior draw of Ψ within the sampler. While we follow this latter approach, we note that it yields similar results in most cases in comparison to where the former approach is followed.

3.3 Bayesian Model Comparison

By employing Bayesian methods we are also able to compare model performance very easily. Working directly with the marginalised likelihood $f(\{y_t\} : \Psi, \Gamma_0)$ has the advantage that the values of $f(\{y_t\} : \Psi^g, \Gamma_0)$ can, for posterior draws of Ψ^g (where $g = 1, ..., G$), be recorded directly within the estimation process facilitating model comparison. In order to compare models we use the Deviance Information Criteria (DIC) of Spiegelhalter et al. (2002). The DIC provides a measure of model performance in terms of the balance between goodness of fit and model complexity. In the literature, a model with the smallest DIC is considered the preferred model according the DIC criteria. The DIC is a Bayesian analog of the Classical information criteria (e.g. Akaike). Numerically it computes a value of K, which is an estimate of the 'effective number of parameters'. The DIC rewards a high average log likelihood, but penalizes each model according the effective number of parameters.

3.4. Priors

As discussed above, the use of the Kalman Filter requires priors to be specified for the latents (Γ_0). These need to be specified in both a Bayesian or Classical context. While, in principle, these can be specified using prior information, an alternative, 'non informative' approach is to use the first few observations of the explanatory variables in order to construct a proper prior, after which we exclude these observations in estimation. This can be done more easily, but equivalently, by setting the mean of the latents to zero and the covariance of the latents can be set to be equal to a very large value (e.g. $I \times 10^8$

where I is an identity matrix). The predictive error likelihood (Harvey, 1989, p126) is then summed from $t = n + d + 1$, where $n + d$ is equal to the number of regressors in the model and d is equal to the number of non-stationary components in the transition equation. Within the Classical approach no further priors are required.

However, if a Bayesian approach is used, then priors are also required for the hyper parameters Ψ. These can be set in a reasonably non-informative way by reparameterising the model as $\Psi^* = (\ln \sigma_e^2, \ln \sigma_v^2, \ln \sigma_w^2)$ and then adopting a flat (improper) prior

$$p(\Psi^*) = I_{[-u,\infty]}\left(\ln \sigma_e^2\right) \times I_{[-u,\infty]}\left(\ln \sigma_w^2\right) \times I_{[-u,\infty]}\left(\ln \sigma_w^2\right) \quad (5)$$

where $I_{[-u,\infty]}(x)$ denotes an indicator function which is equal to one if $x \in_{[-u,\infty]}$ and zero otherwise. The finite bottom bound is required because as the variance goes to zero, then the logged variance becomes near unidentified (which means that the likelihood become invariant to smaller values) below a small value $-u$. Here we set $u = 25$ (results are negligibly different to those we present if we set $u = 10$ or 50).

4. Empirical Section

4.1. Data

The temperature data that we employ in this Section are obtained from the CRU website. These are global temperature anomalies from 1850 to 2009[3]. The forcing measures are those used in Mills (2009) available on David Stern's website[4] from 1850 to 2000. Therefore, our estimated models are over this shorter time period, 1850 to 2000. The construction of this data has been discussed in a number of places and therefore we do not repeat this here.

In this paper, as in Mills (2009), we employ the aggregate forcing measure that is a linear sum of the greenhouse gases, sulphur dioxides, and solar components. The use of aggregative or total forcing can be justified from a theoretical view since they are constructed in such a way that the measures should be summable. Moreover, previous work using total forcing has suggested that this measure is cointegrated with temperature anomalies, and Mills (2009) also finds that a test for equality of the forcing measures accepts this restriction. As can be seen from this plot, there is an evident rise in both radiative forcing and temperatures over most of the later part of last century.

[3]The series we use is Hadcrut3gl. We note that the variance adjusted version of the series available from the Website http://www.cru.uea.ac.uk/cru/data/temperature/ gives very similar results in the models we estimated.

[4]http://www.sterndavidi.com/datasite.html

Table 1: Null of Cointegration

No Trend	Trend
C_μ=0.218	C_τ=0.219
P value (0.133)	P value (0.005)

4.2 Tests for Integration.

Since the unit root behaviour and tests under the null of no cointegration have been presented in the preceding literature we will not repeat this analysis herein. However, briefly, as in previous work (e.g. Mills, 2009) unit root tests indicated that both temperature and radiative forcing series are non-stationary. Both series (temperature and radiative forcing) are consistent with being integrated of order one according to Augmented Dickey Fuller tests along with other tests including those that adopt a unit root as the alternative hypothesis. The results of these tests are available from the authors on request.

4.3 Null Hypothesis of No Cointegration

Tests for cointegration (Johansen rank test) allowing for a restricted trend and intercept in the long run relationship, indicate that forcing and temperature series are cointegrated. The VAR analysis suggests that, using a model with an intercept and a time trend, two lags are appropriate (on the basis of an F test of the significance of a third lag, and according to both Akaike and Bayes information criteria), and that the hypothesis of no cointegration is rejected at below the 1% level of significance. This was also supported by Bayesian estimation of the VAR with and without rank restrictions. Regardless of lag length, the DIC criteria supported cointegrated models over a fully differenced VAR or VAR without rank restrictions. Again due to length constraints these are not reported here.

4.4 Null Hypothesis of Cointegration

The tests for cointegration, adopting cointegration as the null rather than the alternative, is less definitive. The critical values for the tests of no cointegration C_μ and C_τ are given in Shin (1994). However, we simulated the p-values for our sample size (151) using 10,000 Monte Carlo trials[5]. These results are presented in Table 1.

As can be seen from Table 1, we cannot reject the cointegration hypothesis at the 10% level of significance if a trend is not included in the regression, but if a trend is included,

[5]Our simulated critical values are very similar to those produced in Shin, (1994), therefore, we believe that our p-values should be accurate.

we would reject the cointegration hypothesis at a very low level of significance. The trend in the FM regression is not significant, therefore we may as well conclude that the No Trend result is reliable (i.e. preferred). Nonetheless, the rejection of the cointegration hypothesis when a trend is included needs to be given some weight. Therefore, it is not completely clear that the null of cointegration between the two variables cannot be rejected using these tests.

4.5 Structural Time Series Results

Moving on to an analysis of the structural time series model represented by equation (3) under alternative restrictions, we first discuss the Classical Maximum Likelihood results, before assessing the results from the Bayesian analysis. In all models (containing a random trend β_t), the estimate for the parameter σ_w^2 was indistinguishable from zero and a p-value for this restriction based on an adjusted likelihood ratio test was close to one. Therefore, for subsequent analysis we imposed the restriction that $\sigma_w^2 = 0$ (the trend term β_t in the equation is time invariant) for all models.

Therefore, we have three models:

- M1: σ_e^2 and σ_v^2 (unrestricted model)
- M2: $\sigma_v^2 = 0$ (cointegrated model)
- M3: $\sigma_e^2 = 0$ (random walk error model)

M1 contains both a stochastic intercept and a random error and nests both models M2 and M3 as special cases. M2 is equivalent to a standard regression with a stationary error (a cointegrated model). M3 has a non-stationary error, with only a random walk intercept.

4.6 Classical Results

The results presented in this paper include up to three lags of the temperature variable as explanatory variables in equation (3). A fourth lag is insignificant in all models that we estimated. The significance of the lags depended on the restrictions that were placed on the variance terms. For models that have the restriction $\sigma_e^2 = 0$ imposed, all three lags are highly significant. For models that imposed :$\sigma_v^2 = 0$ only the first lag is highly significant. Therefore, we present results for one, two and three lags.

Due to failure of detectability and stabilisability conditions (see Harvey, 1989 for details) if $\sigma_w^2 = 0$, then a formal test of $\sigma_v^2 = 0$ cannot be constructed using likelihood ratio, Lagrange Multiplier or Wald statistics. Thus, for performing a formal Classical test of the cointegration hypothesis, we rely on the Shin tests statistics presented in Table 1.

A valid test can be constructed for $\sigma_e^2 = 0$ using a likelihood ratio test provided the significance is adjusted to take account that it is on the edge of the parameter space. The p-values results for testing restrictions $\sigma_e^2 = 0$ are presented in Table 2 for M3 containing one, two and three lags.

Table 2. Likelihood Ratio Tests and Likelihood Comparisons

	1 lag	2 lags	3 lags
Null Model= M3*	1E-5	0.0005	0.1529
$\ln(L_{M2})$-$\ln(L_{M3})$	8.468	3.25	-0.450

*The alternative model is M1. Values represent P-Values for the null hypothesis

As we noted above, the significance of the lags implies that three lags are definitely required for a valid test of M3. As we can see from Table 1 when the model contains one or two lags only, $\sigma_e^2 = 0$ is rejected. However, where there are three lags in the model we cannot reject the null at the 10% level. In other words, provided three lags are included in the model (all of which are significant), a model with a pure random walk cannot be rejected.

The importance of the number of lags to include is also apparent when comparing models M1, M2 and M3. Comparing the log-likelihoods for each of the models, the likelihood function M2 is higher than for M3 for one and two lags, but if three lags are included then the likelihood function for the pure random walk error model (M3) is in fact slightly higher than for M2. In summary, if three lags are included, then a model which has a random walk error cannot be rejected and, this model has a higher likelihood function than the cointegrated model.

Henceforth, we only report the results for models with three lags. The reason for this is that for the Classical results the coefficients of explanatory variables are almost identical for models M1 and M2 regardless of whether one, two or three lags are included. However, as we have outlined above for M3, the third lag is highly significant. Therefore, results for M3 would be biased unless three lags are included.

We now present, in Table 3, our Classical estimates for the structural time series models.

Table 3. Classical Estimates of Coefficients

	M1	M2	M3
temp_{t-1}	.424	0.541	-.338
	(.0829)	(0.082)	(.083)
temp_{t-2}	-.1466	-0.105	-.361
	(.089)	(0.093)	(.079)
temp_{t-3}	.0253	.111	-.2603
	(.0829)	(.081)	(.080)
forcing	.2548	.2434	.164
	(.0963)	(.063)	(.226)
trend	.0011	.0001	.0082
	(.0017)	(.004)	(.0091)
σ_e	.0958	.1021	
σ_v	.0170	.	.1073
-2LogL	-642.818	-640.87	-641.77
Prediction error variance	.011795	.011862	.011656
AIC	-4.3209	-4.3152	-4.3327
BIC	-4.1411	-4.1354	-4.1529
Normality P-Values*	0.3829	0.3026	0.6075

Numbers in parentheses are standard errors

*From Bowman Shenton Statistic (Harvey, p.260)

Considering the Classical results in Table 2, the coefficients for the explanatory variables are presented along with estimates of the variances of innovations that drive the irregular and random walk components plus other summary statistics. All models appear to have normal errors, and according to both the information criteria used, M3 is the preferred model. In M1 it is evident that the variances σ_e^2 is estimated to be considerably larger than for σ_v^2. However, this change in M2 and M3 whereby setting one of the variances to zero, yields a variance estimate of a similar magnitude for the other. This may seem surprising, given that the effects of innovations of v_t are cumulative and would generally therefore be expected to have smaller variance. However, examination of the coefficients for the lagged temperatures in the models reveal that coefficients are very different in models M1 and M2 compared with M3. The lag coefficients in M3 are all negative and sum to around -0.96. This means that the apparently irregular component in the series is being captured by negative correlations from period to period, even though each shock is treated as having a permanent impact. Notably, the estimates from unrestricted model M1 are much more similar to M2 than M3. Importantly, both

M1 and M2 have a highly significant positive coefficient on the forcing variable (0.2548 and 0.2434 respectively).

Given the magnitudes of the lag coefficients, the long-run multipliers, which are defined as $\frac{\alpha}{1-\sum_{i=1}^{3}\gamma_i}$ for the impact of forcing on temperatures, are approximately 0.34 and 0.54 for M1 and M2 respectively. This is in contrast to the smaller and insignificant coefficient from M3 (0.1649) and a corresponding long run multiplier at just over half that value. Therefore, the findings with regard to the impact of radiative forcing are substantively different if we use M3 rather than M1 or M2. Furthermore, as discussed above, the restriction of M1 to M2 cannot be rejected on the basis of a Likelihood Ratio test, and M3 has a slightly higher likelihood (providing 3 lags are included in the model).

These results may seem confusing since the unrestricted model M1 yields rather similar results to the restricted cointegrated model M2, yet M3 which yields very different estimates seems to be marginally preferred to M2 (if three lags are included). The reason for this outcome is that there is a global maximum likelihood which has a relatively small variance in the random walk component and a larger variance in the irregular component. However, the evidence here suggests another local maximum with a small irregular component and larger random walk component. Maximum likelihood estimation reflects only the former (global maximum). However, from a Bayesian perspective, point estimates should be derived from the full posterior density, not just a single mode. For this reason we now consider Bayesian estimation.

4.7 Bayesian Results

The Bayesian estimates of all three models above are presented in Table 4. These are presented along with the DIC for each model, which should be at a minimum for the best performing model.

First, it is evident that the two restricted models (M2 and M3) yield virtually identical estimates to the Classical results reported in Table 2. This is because by restricting either of the variances the values of σ_e or σ_v derived from the mean of the posterior are almost the same as their maximum likelihood components and we have only a very small standard deviation.

Second, the unrestricted model, when estimated using a Bayesian approach, yields quite different results from the Classical approach. Examining the coefficients of M1 it becomes clear that the estimates sit in between M2 and M3. This is because, in effect, it averages over M2 and M3, since both these models have reasonably high posterior densities. This can best be seen by the contour plots of the joint posterior densities for σ_e and σ_v displayed in Figure 1. There are two clear posterior modes, one where σ_e is very small and σ_v is around 0.10 and one where σ_v is very small and σ_e is around 0.10.

Between these two modes there is a ridge of a lower density region that highlights the negative covariance between σ_e and σ_v. The point estimates obtained by taking the mean posterior values are 0.0614 and 0.0575, but neither of these points are high density points as such. Rather they sit somewhere in the middle between the two posterior modes. The rest of these coefficients also reflect this tendency to average between these two highly competing models.

Third, according the DIC, the models are ranked M3 (top) followed by M1 and then M2 which concurs with the Classical information criteria. Thus, the cointegrated model is less preferred as compared to the unrestricted model of random walk errors. Radiative forcing retains its positive coefficient estimate, but the standard deviations for this coefficient are as large or larger for both M1 and M3. Thus a (Bayesian) credible interval would contain considerable mass below zero. Interpreting this in Classical terms would suggest that the forcing variable is insignificant.

Table 4. Bayesian Coefficient Estimates

	M1	M2	M3
temp_{t-1}	.053 (.351)	0.541 (.082)	-.338 (.089)
temp_{t-2}	-.251 (.128)	-.1056 (.093)	-0.3609 (.0801)
temp_{t-3}	-0.115 (.159)	.1109 (.082)	-0.2603 (.0811)
forcing	.194 (.180)	0.253 (.063)	0.1649 (.229)
trend	.0047 (0.007)	.0001 (.0004)	.0082 (.009)
σ_e	.0614	.1027	.
σ_v	.0575		.1080
DIC	-206.8472	-206.727	-207.60

Numbers in Parentheses are standard deviations.

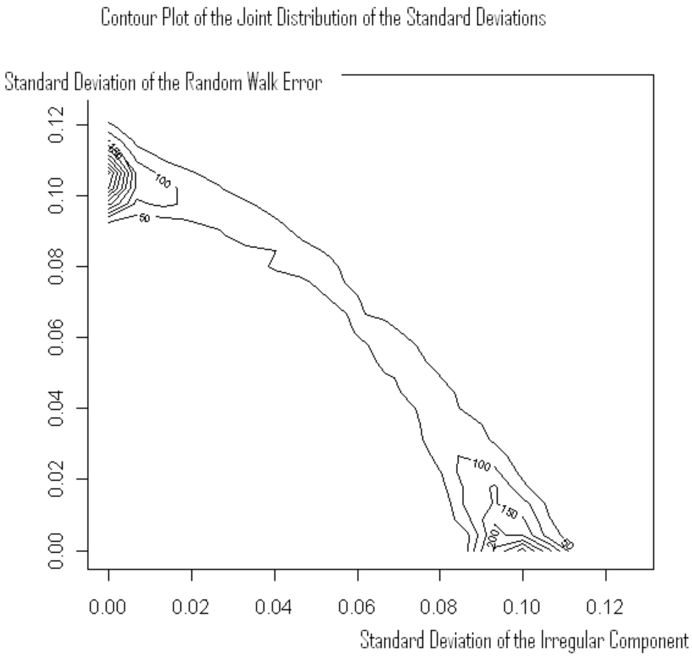

Figure 1: Contour plot for the joint posterior densities

5. Discussion

Most recently Mills (2009) presented statistical evidence that there is a 'long-run equilibrium' between radiative forcing measures and temperature using data from 1850-2000. Mills (2009) builds upon work by Brohan et al. (2006), Stern and Kaufmann (2000), Kaufmann and Stern (2002), Kaufmann et al. (2006) that broadly supports the contention that forcing measures have a quantitative impact on global temperatures. Dergiades at al. (2016) provide results on long-run changes in radiative forcing and surface temperature consistent with the theory of anthropogenic climate change. Folberth et al. (2012) find that increased emissions and radiative forcing have a significant (negative) impact globally on megacities.

The results we have presented in our analysis are an example of how there remains uncertainty about the form and strength of the statistical relationship between global temperatures and human activity. Indeed, our results are not suggesting that there is no direct human role in climate change, far from it. But from a strategic perspective they speak to the idea that there is still uncertainty as to the specific causal mechanisms and as such we need to be somewhat cautious when it comes to how we might best articulate the specific type of policy interventions required. This point is also discussed by Tol (2018) in an excellent overview of the economics of climate change. As is explained, climate change is universally agreed to be a negative externality and that policy does need to be put in place to deal with the effects. However, what remains a hotly debated issues is the impact of climate change. The reason why this is so contentious is that the impact will in turn inform the price that is placed on carbon and differences opinion of this have serious ramifications for policy. Therefore, econometric analysis showing causal links between temperature and human activity still matter as this can inform the focus of such policy interventions and also the likely price of carbon required to induce the necessary changes in behaviour. So remaining uncertainty about the relationships, such as the one examined in the paper matter.[6]

The impact of any statistical results in this domain can be considered in terms of how they may or may not influence policy making. Although there is a general consensus about the impact of human economic activity on the climate there still remains much uncertainty as to the precise mechanisms through which this works (for examples of this within areas of environmental policy making, see Touza and Perrings (2011), Eichner and Pethig (2018) and Kersting (2018)). This uncertainty, however, can causes problem for government when it comes to making credible commitments. Clearly, if governments are able to state a credible position regarding climate policy this can reduce uncertainty for economic agents. However, credible commitments by government cannot typically be enforced and as such economic agents will always place some positive probability on a policy change and as a result a loss of some degree of credibility. But there is also the need for policy responses to be flexible especially as new information becomes available. An obvious and well understood consequence of this resulting uncertainty will be impacts in terms of investment directed to dealing with aspects of climate change. Zetland (2017) argues that within the context of group cooperation in the provision of public goods it may be easier to promote cooperation in the provision public goods within a more competitive setting whereby teams (or coalitions) are encouraged to beat other teams (or coalitions) rather than cooperating with them. Of course, there are steeps that can be taken to minimize the impact of credible commitment whilst retaining flexibility. But, the econometric results we present and the literature we add to demonstrate clear that there remains aspects of uncertainty and that this means that there must be flexibility in policy making even if this impacts on policy makers ability to credible commitment to policy options today.

Mills (2010) provides a useful take on this issue. Essentially, statistically arguments alone will not provide definitive evidence or singularly resolve many of the most highly

[6]We contend that there needs to be less strident expression of opinion about this topic and the research published. This is neatly illustrated in the response to Tol (2016) by Cook et al. (2016) who criticise Tol because of a specific use of data on opinions about climate change. Exchanges such as this distract attention from the very real and important issues that climate change presents.

debated issues within the literature. As he neatly explains, when it comes to examining these issues:

"Statistical arguments alone are unlikely to settle issues such as these, but neither are appeals to only physical models or the output of computer simulations of coupled general circulation models. In such circumstances it would appear that, to quote another ageless proverb, 'you pays your money and you takes your choice'. Indeed, it could be argued that such a proverb is particularly apposite given the ongoing debate concerning the potential costs of combating global warming and climate change, the most notable recent protagonists being Stern (2007) and his reviewers, for example, Nordhaus (2007), Tol and Yohe (2006) and Weitzman (2007)." (p. 424).

Finally, there is scope for micro and context specific studies of greenhouse gas emissions to provide greater context to the type of global study we have presented here. For example, the single and multiple country level studies, such as those undertaken by Guntin-Araujo et al (1999), Fereidouni (2013), Yusuf et al. (2014) and Raheem and Ogebe (2017) can help to empirically link global carbon dioxide emissions with specific sources.

The way to address this is as follows. If going to further our understanding of global temperatures there is a need to examine and challenge existing hypotheses - this is not to refute global warming but to at least raise the prospect that current mechanisms as they are understood may need to be redefined in light of alternative model results. This is a very important result while formulating appropriate environmental policy at national and international levels, to realise more effective outcomes.

6. Conclusions

In this paper, we have presented further empirical investigation of the relationship between radiative forcing and global temperature anomalies. Unlike other recent work exploring this relationship, we used a structural time series approach comparing alternative models as well as adopting cointegration as the null hypothesis. Our findings suggest that previous findings of cointegration between forcing measures and temperatures should be treated tentatively. While the data is consistent with a positive impact of radiative forcing on temperatures, the significance of the impact of forcing was model dependent. While a model that assumes cointegration between forcing temperatures performs reasonably well, a non-cointegrated model performs just as well, or on the basis of the tests conducted here, even better. This was particularly evident when examining the posterior density of the standard deviations in the irregular and random walk errors.

The reason for this finding has been explained using Bayesian methods. Specifically, a contour plot of the posterior densities showed two peaks, one in a cointegrated region and another in a cointegrated region. In addition, the DIC model selection criteria also suggested that restricting the model to one with only a random walk error improved the performance of the model. Finally, in models where temperatures and total forcing are not treated as being cointegrated, then the evidence that total forcing has an impact on temperatures is reduced. However, we would contend that given the ongoing debates within the literature regarding how best to statistically capture, explain and model this relationship, that researchers should avoid being too strident in their claims about model performance. This then inevitably implies uncertainty regarding our understanding of the relationship which in turn has implications for how policy makers respond to and use statistical results of this in policy design and implementation.

References

Ajmi, A. N., Montasser, El, G., and Nguyen, D. K. (2013). Testing the relationships between energy consumption and income in G7 countries with nonlinear causality tests. Economic Modelling, 35, 126-133.

Anger, N. (2008). Emissions trading beyond Europe: Linking schemes in a post-Kyoto world. *Energy Economics* 30, 2028-2049.

Cialani, C. (2007). Economic growth and environmental quality: An econometric and a decomposition analysis, Management of Environmental Quality, 18(5): 568–577.

Chevillon, G. (2017). Robust cointegration testing in the presence of weak trends, with an application to the human origin of global warming, Econometric Reviews, 36:5, 514-545.

Cook, J., Naomi Oreskes, Peter T Doran, William R L Anderegg, Bart Verheggen, EdW Maibach, J Stuart Carlton, Stephan Lewandowsky, Andrew G Skuce, Sarah A Green, Dana Nuccitelli, Peter Jacobs, Mark Richardson, Bärbel Winkler, Rob Painting and Ken Rice (2016). Consensus on consensus: a synthesis of consensus estimates on human-caused global warming Environmental Research Letters 11 048002

Dergiades, T., R. K. Kaufmann and T. Panagiotidis. (2016). Long-run changes in radiative forcing and surface temperature: The effect of human activity over the last five centuries. *Journal of Environmental Economics and Management* 76, 67–85.

Eichner, T. and R. Pethig. (2018). Global Environmental Agreements and International Trade: Asymmetry of Countries Matters, *Strategic Behavior and the Environment:* 7(3-4), 281–316.

Fereidouni, H. G. (2013). Foreign direct investments in real estate sector and CO2 emission, Management of Environmental Quality:, 24(4): 463–476.

Folberth, G. A., S. T. Rumbold, W. J. Collins and T. M. Butler. (2012), Global radiative forcing and megacities. Urban Climate. 1, 4–19.

Gallegati, M. (2018). A systematic wavelet-based exploratory analysis of climatic variables. Climatic Change,148:325–338

Gerlagh R. and Zwaan B. Van Der. 2006. Options and Instruments for a Deep Cut in CO2 emissions: Carbon dioxide capture or renewables, taxes or subsidies?. *The Energy Journal* 27(3): 25–48.

Guntin-Araujo, X., M.L. Chas-Amil, M.C. Lorenzo-Diaz. 1999. The atmospheric emissions in Spain: a regional analysis, Environmental Management and Health, 10(3): 155–164.

Harvey A.C. (1989). *Forecasting Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Johansen S. (1995). *Likelihood-Based Infernce in Cointegrated Vector Auto-Regressive Models, Advanced Tests in Econometrics.* Oxford University Press. Oxford.

Kaufmann R.K., Kauppi, H. and Stock, J.H. (2006). The Relationship Between Radiative Forcing and Temperature: What do Statistical Analyses of the Instrumental Temperature Record Measure? *Climatic Change*. 77: 279–289.

Kaufmann R.K., Kauppi, H. and Stock, J.H. (2010). Does temperature contain a stochastic trend? Evaluating conflicting statistical results. *Climatic Change*. 101: 395-405.

Kaufmann R.K. and Stern D.I. (2002). Cointegration Analysis of Hemispheric Temperature Relations. *Journal of Geophysical Research*, 107 (D2), article No. 4012.

Kersting, J. (2018). Stability of Global Climate Cooperation Under Uncertainty, *Strategic Behavior and the Environment:* 7(3-4), 351–396.

Koop G. (2003). *Bayesian Econometrics.* Wiley, Chichester.

Lai, K.S. and Yoon, M. (2018). Nonlinear trend stationarity in global and hemispheric temperatures, Applied Economics Letters, 25:1, 15-18.

Lee, C.-C., and Chang, C.-P. (2009). Stochastic convergence of per capita carbon dioxide emissions and multiple structural breaks in OECD countries. Economic Modelling, 26(6), 1375-1381.

Lee, D. M. and Lyon, K. S. (2004). A dynamic analysis of the global timber market under global warming: An integrated modeling approach. *Southern Economic Journal*, 70(3): 467–489.

List, J. A. (1999). Have air pollutant emissions converged among US regions? Evidence from unit root tests. *Southern Economic Journal*, 66(1): 144–155.

Lu, C.-C., Chiu, Y.-H., Shyu, M.-K., and Lee, J.-H. (2013). Measuring CO2 emission efficiency in OECD countries: Application of the Hybrid Efficiency model. Economic Modelling, 32, 130-135. doi:10.1016/j.econmod.2013.01.047

Maddala G.S. and Kim, I.M. (1998). *Unit Roots, Cointegration and Structural Change.* Cambridge University Press. Cambridge.

Marrero, G. A. (2010). Greenhouse gases emissions, growth and the energy mix in Europe. *Energy Economics* 32, 1356–1363.

Mann M.E., Bradley R.S. and Malcolm K. (1998). Hughes Global-scale Temperature Patterns and climate forcing over the past six centuries. *Nature,* 392: 779–787.

McMillana, D.G. and Wohar, M.E. (2013). The relationship between temperature and CO2 emissions: evidence from a short and very long dataset. Applied Economics, 45, 3683–3690.

Mills, T.C. (2010). 'Skinning a cat': alternative models of representing temperature trends. An editorial comment. Climatic Change, 101:415–426.

Mills T.C. (2009) How Robust is the Long-Run Relationship Between Temperature and Radiative Forcing? *Climatic Change,* 94:351–361.

Mohapatra, G. and A.K. Giri. (2009). Economic development and environmental quality: an econometric study in India, Management of Environmental Quality, 20(2): 175–191.

Phillips, P.C.B. and Hansen, B. (1990). Statistical Inference in Instrumental Variable Regressions with I(1) Processes. *Review of Economics Studies,* 57: 99–125.

Raheem, I. D. and J. O. Ogebe. (2017). CO2 emissions, urbanization and industrialization: Evidence from a direct and indirect heterogeneous panel analysis, Management of Environmental Quality, 28(6): 851–867.

Romero-Avila, D. (2008). Convergence in carbon dioxide emissions among industrialised countries revisited. *Energy Economics* 30, 2265–2282.

Shin Y (1994) A Residual-Based Test of the Null of Cointegration Against the Alternative of No Cointegration: *Econometric Theory*, 10(1): 91–115

Shukla P. R., Garg A., Kapshe M. and Nair R. (2006). IndiaÕs Non-CO2 GHG Emissions: Development Pathways and Mitigation Flexibility. *The Energy Journal* 27: 461–483.

Spiegelhalter D.J., Best N.G., Carlin, B.P. and van der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society: Series B* (Statistical Methodology) 64 (4): 583 - 639

Stern, D.I. and Kaufmann, R.K. (2014). Anthropogenic and natural causes of climate change. Climatic Change, 122:257–269

Stern D. and Kaufmann R.K. (2000). Detecting a Global Warming Signal in Hemipeheric Temperature Series: A Structural Time Series Analysis. *Climatic Change* 47: 411–438.

Sun J. W. (1999). Decomposition of aggregate CO2 emissions in the OECD: 1960-1995. *The Energy Journal* 20(3): 147–156.

Tol, R.S.J. (2016). Comment on 'Quantifying the consensus on anthropogenic global warming in the scientific literature'. Environmental Research Letters, 11, 048001.

Tol, R.S.J. (2018). The Economic Impacts of Climate Change. Review of Environmental Economics and Policy, 12 (1): 4–25.

Touza, Julia M. and Charles Perrings (2011). Strategic Behavior and the Scope for Unilateral Provision of Transboundary Ecosystem Services that are International Environmental Public Goods, *Strategic Behavior and the Environment:* 1(2), 89–117.

Van Heerden J., Gerlagh R., Blignaut, J., Horridge M., Hess S., Mabugu R. and Mabugu M. 2006. Searching for triple dividends in South Africa: Fighting CO2 pollution and poverty while promoting growth. *The Energy Journal* 27(2): 113–142.

Yusuf, R.O. , Z. Z. Noor, A.H. Abba, M. A. A. Hassan, M. R. Majid and N. I. Medugu. (2014). Predicting methane emissions from livestock in Malaysia using the ARIMA model, Management of Environmental Quality, 25(5): 585–599.

Zetland, D. (2017), Exploring Group Cooperation in the Provision of Public Goods, *Strategic Behavior and the Environment:* 7(1-2): 109–133.