Response to Referee Reports and Open Comments

David F. Hendry and Felix Pretis*

*Institute for New Economic Thinking at the Oxford Martin School

University of Oxford, UK

Abstract

Response to Referee Reports and Open Comments on "Comment on 'Polynomial cointegration tests of anthropogenic impact on global warming' by Beenstock et al. (2012) - Some fallacies in econometric modelling of climate change"

*

1 Overall Response

We would like to thank the three anonymous referees and Beenstock et al. for providing comments on our submission. In response to the comments we extend our critique by first addressing some overall points mentioned by the reviewers and then address individual points in turn. Beenstock et al. (2012) will refer to their initial paper, and Beenstock et al (2013) to their response.

1.1 More direct reference to Beenstock et al. (2012) Paper:

Referees pointed out that some of the critiques mentioned in the first part of our comment should be more directly focused on the application of Beenstock et al. (2012). Here we provide extensions on omitted variables, incorrectly modelled relations and a brief paragraph on un-modelled heterogeneity and aggregation to address this point. These will be included in the revised paper.

1.1.1 Omitted Variables

We mention omitted variables being a problem but should have been more specific in how this applies to Beenstock et al. (2012). Here we briefly list a few of these variables that have been omitted and may play an important role, though they should be considered with caution as some of these gases are not as well mixed (but spatially varied) and thus may not be appropriate in a zero-dimensional model. Myhre et al. (2001) provide a good overview of available time series for the time period used. These include CFCs (Chlorofluorocarbons), as used by Stern & Kaufmann (2000), which together with tropospheric ozone likely exhibit a positive forcing, as well as stratospheric ozone (see Myhre et al., 2001) which likely acts as negative forcing. Further, it is unclear whether Beenstock et al. (2012) included volcanic aerosol forcing. It may have been included in their initial analysis as stratospheric aerosols, but this is unclear as the link provided in their data appendix appears not to work. Since there is no graph of the stratospheric aerosol series provided in their paper it is difficult to verify whether volcanic emissions were considered.

1.1.2 Incorrectly Modelled Relations

In our initial critique we pointed out that the assumption that all anthropogenic forcing variables are I(2) is flawed (this is further discussed in section 1.2 here). We expand our initial section on incorrectly modelled relations by assessing their method of constructing a measure of anthropogenic activity. Their measures of anthropogenic activity (as given by equations 9 and 10 in their paper, reproduced here as equations (1) and (2) are the residuals g_1 and g_2 of a single regression of radiative forcing of CO2 on the forcing of other greenhouse gases.

$$\text{rfCO}_2 = 10.972 + 0.046\text{rfCH}_4 + 10.134\text{rfN}_2\text{O} + g_1 \tag{1}$$

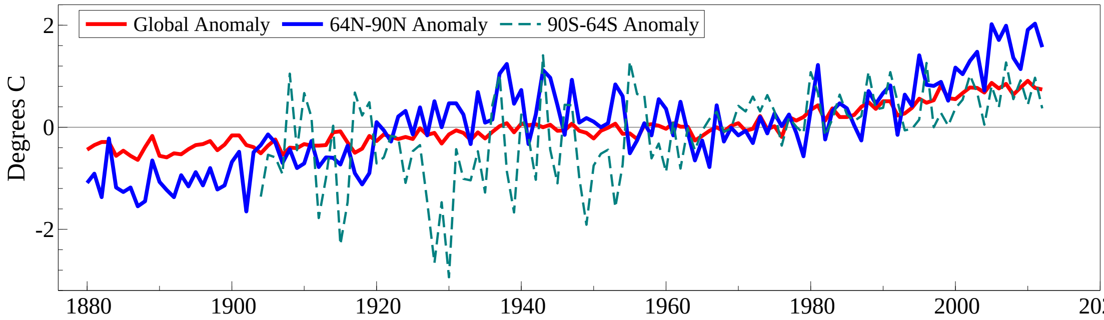
$$\text{rfCO}_2 = 12.554 + 0.345\text{rfCH}_4 + 9.137\text{rfN}_2\text{O} + 1.029\text{BC} + 0.441\text{Reflaer} + g_2 \tag{2}$$

Such regressions are a variant for possibly I(2) variables of the approach in Engle & Granger (1987). Banerjee et al. (1986) demonstrated that this type of test imposed 'common factor' restrictions of the form criticized by Hendry & Mizon (1978), the test lacks power and is substantively inferior to the system approach in Johansen (1988), which we describe below. Lets consider the anthropogenic measure of Beenstock et al. (2012) in two cases.

First, suppose we accept their starting point that all anthropogenic variables are I(2). They state that equations (1) and (2) are to test for cointegration between the anthropogenic series. However, cointegration is a system property and thus the variables need to be treated as such. To establish cointegration between the variables in Equation (1) (rfCO2 regressed on rfCH4 and rfNO2), the full system of three variables needs to be considered (see Hendry & Juselius, 2001). The system has at most full rank (=3) or if there is cointegration, the system may exhibit reduced rank of one or two. The rank of the system needs to be tested, then the system can be estimated with the reduced rank imposed and the cointegrating relationships estimated. If the reduced rank happens to equal two, then there are actually two cointegrating relations between the three series and thus there are two potential anthropogenic anomaly measures. The single anthropogenic anomaly given in their equation (1) is then a linear combination of measures of anthropogenic activity. The same problem generalizes to their equation (2), with there being five variables in the system and a much larger set of potential cointegrating relations. The system of five variables may have full rank (=5), or reduced rank between one and four if there is cointegration, implying up to four cointegrating relations and up to four measures of the anthropogenic anomaly. Thus, even if their starting point that all anthropogenic are I(2) is accepted, then their measure of the anthropogenic anomaly is likely only one of many, given the large number of potential cointegrating relations. There could well be a residual (anthropogenic anomaly) that does cointegrate with temperature and solar irradiance. This point had been hinted at by a referee for their initial submission but clearly not been addressed.

Second, given that their starting point of assuming all anthropogenic variables are I(2) is flawed, the measure of the anthropogenic anomaly is inappropriate. Their

Figure 1: Temperature Anomalies relative to 1951-1980 Average



measure of anthropogenic activity is the residual of a regression of rfCO2 on radiative forcing of the other greenhouse gases. This means that the measure of anthropogenic activity used in their analysis is really the variation in radiative forcing of CO2 that is unexplained by the variation in other greenhouse gases. Given the basic energy balance model, radiative forcings are mostly considered additively. The total effect of all forcings together (while taking feedbacks into account) is what is important. Taking the unexplained variation in radiative forcing of CO2 as a measure of anthropogenic activity is then simply incorrect and does not measure what Beenstock et al. (2012) claim it does. The main test of anthropogenic global warming in Beenstock et al (the regression of temperature on solar irradiance and the anthropogenic anomaly in their table 3) is then really just a regression of temperature on solar irradiance and a residual. It is very questionable whether this captures any anthropogenic component at all, and it does not capture the main anthropogenic forcing component.

1.1.3 Unmodelled Heterogeneity and Aggregation Bias

Temperature trends vary spatially and rejecting a relationship between a single aggregated series and anthropogenic emissions does not imply there is no relationship in general. The time series literature studying radiative forcing and its effect on temperature primarily relies on the global temperature anomaly as a single temperature series. In practice this may be a necessary simplification, however, even if we assume that the analysis in Beenstock et al. (2012) is valid, finding no cointegration between a global aggregate and global anthropogenic forcing does not imply there exist no relationships overall. To illustrate some of this spatial variation, Figure 1 shows the global anomaly together with approximate Arctic (averaged over 64N-90N degrees latitude) and close to Antarctic (averaged over 90S-64S degrees latitude) anomalies (data from NASA Goddard Institute for Space Studies (GISS), 2011).

1.2 Time Series Properties and the "Standard Result"

Beenstock et al.'s (2013) response to our comment primarily appeals to the so-called "standard result" of anthropogenic forcings being stationary in second differences but not in first differences. However, as we show here this is not a "standard result" at all.

The two works cited by Beenstock et al. (2013) to support their "standard result" are: Kaufmann, Kauppi, & Stock (2006) and Kaufmann, Kauppi, & Stock (2010). However these two papers do not investigate the stationarity properties of anthropogenic forcings and are therefore inappropriately cited in Beenstock et al's (2013) comment. These papers contain no test of the time series properties of the anthropogenic forcing series. Further, the first sentences in the first paper (Kaufmann et al., 2006) already contradicts the overall conclusions of Beenstock et al. (2012) by stating that anthropogenic forcing cointegrates with temperature. The paper does not test the time series properties of anthropogenic forcings, but rather analyses temperature series from global climate models. The second paper, Kaufmann et al. (2010), also does not test the time series properties of anthropogenic forcings, but even states that radiative forcing associated with greenhouse gases rises irregularly throughout the historical record (p. 402), which is not consistent with Beenstock et al.'s (2013) claim of a "standard result".

Thus, these two works are inappropriately referenced and do not provide any evidence of this so-called "standard result". Looking further, two papers by Kaufmann and co-authors that actually do investigate the time series properties (but were not cited by Beenstock et al. (2013) in their response) actually show quite the opposite of what Beenstock et al. (2013) are claiming to be a "standard result". First, Stern & Kaufmann (2000) show that when using univariate tests, the results can vary considerably depending on which type of test is used and also vary across different anthropogenic gases. For example, CO2 appears to be I(1) in two out of the four tests, and I(2) in the others (Table 1 in their paper). Similarly, N2O appears I(1) in three out of the four test types. Due to these conflicting results they therefore then employ a different, structural, time series approach. Second, Kaufmann & Stern (2002) test the time series properties of the aggregate of the radiative forcing of all major greenhouse gases (CO2, CH4, N2O, CFC11, CFC12) and find it to be I(1).

The results that Beenstock at al. (2013) claim to be standard are therefore not standard. None of the references provided in Beenstock et al. (2013) actually show or test this "standard result". The two papers they cite do not address this question, two further papers together with our critique show that it is not straight forward to come to a conclusion on the integratedness of these series. As our quick analysis shows, Beenstock et al. (2012) results are not very robust and there are clear structural breaks. Our main critique holds that there are obvious changes to the series, in particular at the point where measurement changes. This break point is not "cherry-picked" as criticised by Beenstock et al. (2013), but given. There appear to be additional step changes as kindly pointed out by one of our anonymous referees. Regardless of the underlying reason for structural change, these breaks need to be addressed. They especially need to be addressed if, as is the case in the Beenstock et al. (2012) paper, the entire analysis hinges on alleged differences in time series properties of various series. Their claim "In summary, the time series properties of anthropogenic forcings are fundamentally different to the time series properties of temperature and solar irradiance" is simply incorrect.

This directly leads to the next main point which is crucial and Beenstock et al. (2013) appear to misunderstand. The level of integration and thus stationarity of data is not intrinsic to the data itself and can change over time. There is nothing inherent in the physical data generating process that makes anthropogenic forcings or other variables I(1) or I(2). The observational data may, over some period, be consistent with an empirical model that is I(1) or I(2) but this is not an intrinsic property that cannot change. There are many examples of this, two that come to mind is that the level of CO2 emissions is related to economic activity which varies over time, or emissions of CFCs which only arose in the latter part of the 20th century. Both of these may be stationary in second differences from the 1950s onwards but because of the underlying processes they may well have been stationary in first differences before then, and in the case of CFCs non-existent before their discovery (Myhre et al. (2001) provide some examples concerning CFCs). To then claim that all greenhouse gases are always I(2) is incorrect. This result is also inconsistent with our analysis and the tests conducted in the previous literature. It is therefore especially dangerous to come to overly strong conclusions just based on the integration properties of the data given the changing nature and the short period of observations.

1.3 Lack of Physical Model and Spurious Relationships

The main point of our comment was to illustrate the fallacies that can be made, and were made by Beenstock et al. (2012), when using econometric modelling of climate series. Thus, the criticisms of our analysis making the same mistakes as Beenstock et al. (2012) in terms of using a too simple model do not apply, since we made no attempt at modelling the series. Our comment contains no attempt to provide a complete model, but rather we show that the statistical modelling approach by

Beenstock et al. (2012) et al is invalid. With this in mind, our Figure 4 plotting temperature against CO2 was meant as an illustration only, but we understand that this can easily lead to misunderstandings and will thus remove it.

2 Direct Response to Comments

2.1 Anonymous Referee #1

- Comment 2. Sentence will be changed to 'Indeed it does not even imply that the elements of $\{Z_t\}$ are not...'.

- Comment 3. We will clarify that this relates to the mentioned example and reword the sentence, it should read 'a linear approximation to a non-linear relation (see Ch. 6 in Hendry, 1995) or an unmodelled location shift (see Castle & Hendry, 2013)'

- Comment 4. "It would be interesting to also test the unit root hypothesis for $\Delta \mathrm{rfCO}_2$ test using the data from 1978–2011 [...] with the constant". Table 1 shows the test results of an ADF test on $\Delta \mathrm{rfCO}_2$ for 1978–2011 on just a constant. The null hypothesis of unit-root non-stationarity is rejected. Thank you for pointing out more dramatic step shifts from 1960 to 1978, this adds to our critique that there may be many underlying structural changes that require modelling in the data.

Table 1: ADF Unit root test for $\Delta \mathrm{rfCO}_2$ using a constant

1978-2011		
D-lag	**t-adf**	**Reject H0**
2	-3.13	*
1	-4.23	**
0	-4.74	**

- Typographical errors: Will be changed.

2.2 Anonymous Referee #2

- i) "Fail to specify a plausible physical model". Following the point made in the previous section, we do not rely on a basic model because main difference to Beenstock et al. (2012) is that we did not make an attempt to model the system. Rather we show that using the statistical tools that Beenstock et al. (2012) use is invalid on top of their basic of physical model.

- ii) "Beenstock et al [...] come to overly hasty conclusions on the implcations for climate change". We agree with reviewer that Beenstock et al. (2012) come to overly hasty conclusions and show that many can easily be overturned.

- iii) "For the length of time series we have it is hard to come to very strong conclusions about the nature of the series using simple tests". We very much agree that for the lengths of time series it is difficult to come to conclusions, we added the additional point that these properties are not intrinsic and can easily change over time. In particular, and this is our point, that the measurement changes and there are very clear structural breaks. As pointed out above, the "standard result" Beenstock et al. (2013) appeal to is anything but standard.

- iv) "A possible explanation is an increase in uptake of heat by the ocean". We agree that this may be a potential reason, unfortunately Beenstock et al. (2012) do not address this point due to the short data series available.

- v) "There are various reasons why there might not actually be warming or warming might be difficult to observe". Point taken. While physically there is a clear relationship and global coupled climate models all point in this direction, it is true that empirically we might not observe this direct link, especially given the short time series we have available.

- vi) "It would be good if the critique was more focused". We agree, we added a section on omitted variables, incorrectly modelled relations, unmodelled heterogeneity and a section on the "standard result" together with an important comment on the fact that integratedness is not intrinsic. While overall the comment was aimed at the work by Beenstock et al. (2012), the broader aim is to introduce readers to the wider scope of fallacies that can be made.

- vii) "Figure 4, which shows this plot is not very convincing". We agree, as mentioned above it was meant to illustrate rather than model but appeared to add confusion. We will remove this plot.

- viii) "It doesn't make physical sense to split the series in two and assume different models". We agree, however, this is not what we do. We made no attempt at modelling it in our original comment but pointed out that if there are underlying breaks (which there are, e.g. due to the measurement change) then simply concluding that something is I(1) or I(2) as Beenstock et al. (2012) do, is not valid.

- ix)"The brief discussion of unit roots could be explained better". We agree, we will add a short section on introducing the concepts of trend/difference stationarity.

2.3 Anonymous Referee #3

- "the use of the word 're-radiation' is confusing": Will be clarified.
- "what is meant by non-linear approximation [...] also a reference might be in order" The sentence should read 'a linear approximation to a non-linear relation (see Hendry, 1995) Ch6. or an unmodelled location shift (see Castle & Hendry, 2013). Concerning the reference for μ, we believe that no reference is needed, since μ here refers to the mean of road deaths in the above graph, thus it is notable from graphs that μ is changing over time. We will make this more clear in the paper.
- "p8, l1-5 sentence is unclear, please consider revising". Will be revised. We meant the composition of the climate, as well as the distribution of series for both spatial (vertical as well as horizontal) and temporal scales. Sinks and sources refer to drivers and reducing factors of various greenhouse gases and heat.
- "not the best strategy to use scatterplots of log(CO2) against temperature". Yes we agree, as mentioned above, it was meant as an illustration rather than implying causality but created additional confusion. It will be removed.
- "Tables 1 and 2. it might be useful to state what the null hypothesis is...". We agree, we will add additional descriptions under the tables.
- "Figure 1 & 4". We will add units and labels for the graphs.
- "Typographical errors". Will be changed.

2.4 Reply by Beenstock et al.

We address most of Beenstock et al.'s (2013) comments above in section (1). Some additional points are being addressed in this section.

- i) "Our conclusions that temperature is stationary in first differences whereas anthropogenic forcings are stationary in second differences is not original and is standard". Please see the above section 1.2 on the "standard result". The result is anything but standard and the two references provided by Beenstock et al. (2013) are inappropriately referenced as they do not address or test the time series properties of anthropogenic forcings.

- ii) "We clearly stated [...] [that] differences in orders of integration do not necessarily refute AGW because temperature and anthropogenic forcings might be polynomially cointegrated Our original contribution was to show that temperature, solar irradiance and anthropogenic forcings are not polynomially cointegrated". We do not imply that Beenstock et al. (2012) are trying to show this, we state that this is an implied result of their analysis. It is very clear that they test for polynomial cointegration in the context of data analysis.

- iii) "Our clearly stated motivation was not to refute this well established [greenhouse gas] theory" We accept the point that Beenstock et al. (2012) are not trying to refute physical theory, we just show that on top of a basic physical model, the statistical starting points are not valid and thus the conclusions have to be re-considered. Our example on road fatalities remains valid as it shows the general problems that can be found when doing analysis of the type of Beenstock et al. (2012). We extended the critique above to show important omitted variables, unmodelled heterogeneity and a failure to handle important distributional shifts.

- iv) "HP need to be more specific which variables we have omitted". See section (1.1.1) above. We added a short section on specific variables that may have been omitted.

- v)"Just because the method of measurement changed does not itself mean that the timer series properties of rfCO2 had to change" We very much agree that just because the type of measurement changed does not imply that the results have changed. However, just by inspection of the data, as we illustrated, it is obvious (as shown by our graph) that the change in measurement induced a change in the series. This is then further supported by our tests on the measured sub-samples.

- vi) "Also, structural breaks might have ocurred even if the method of measurement had not changed" & "the break point in Table 1 has been 'cherry picked'". We agree, structural breaks may have occurred at any point in time, which is why a more rigorous approach than used by Beenstock et al is necessary. For example Impulse Indicator Saturation (Hendry et al. (2008) and Castle et al. (2011)) could be used. However, the break point was not cherry picked but precisely defined by the date when the measurement system changed. As Beenstock et al. should be aware, changes in measurement can alter the time series properties of observed variables, as the recent revisions to US national accounts (adding 3% to GNP) show, and were already analyzed in Hendry (1995).

- vii) "Whereas we have a unified model of anthropogenic forcings, HP have idiosyncratic models for each forcing." First, we did not attempt to provide a full model, we illustrate the invalid statistical starting point of Beenstock et al. (2012). Second, while Beenstock et al. (2012). claim to have a unified model, as pointed out in our section 1.1.2, actually they miss the very important unified system nature of cointegration by using only a single equation analysis.

- viii) "[In reference to the bi-variate plot] This is precisely an example of the sort of spurious correlation result that we sought to expose in our paper". As mentioned above, the graph was meant as an illustration rather than infer causality, it seems to have lead to misunderstanding and will be removed.

- ix) "Also, their criticisms apply to the existing literature [...]". First, Beenstock et al. (2013) inappropriately reference the existing literature claiming a "standard result" that is not standard (see section 1.2). Second, our criticisms of working with this data would then naturally extend to other papers making the same strong assumptions as Beenstock et al. (2012) make. This particular point raised by Beenstock et al. (2013) re-enforces the fact that our critique is relevant to their work. Most of the papers cited here do not make this strong claim of all anthropogenic series being I(2) and do not over-rely on this assumed property.

References

Banerjee, A., Dolado, J., Hendry, D. F., & Smith, G. (1986). Exploring equilibrium relationships in econometrics through statistical models: some monte carlo evidence. *Oxford Bulletin of Economics and Statistics*, *48*, 253-277.

Beenstock, M., Reingewertz, Y., & Paldor, N. (2012). Polynomial cointegration tests of anthropogenic impact on global warming. *Earth System Dynamics*, *3*, 173-188.

Castle, J. L., Doornik, J. A., & Hendry, D. F. (2011). Model selection when there are multiple breaks. *Journal of Econometrics*, forthcoming.

Castle, J. L., & Hendry, D. F. (2013). Model selection in under-specified equations with breaks. *Journal of Econometrics*, forthcoming.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55:2*, 251-276.

Hendry, D. F. (1995). *Dynamic econometrics.* Oxford: Oxford University Press.

Hendry, D. F., Johansen, S., & Santos, C. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, *23*, 337-339.

Hendry, D. F., & Juselius, K. (2001). Explaining cointegration analysis: Part II. *Energy Journal*, *22*, 75–120.

Hendry, D. F., & Mizon, G. E. (1978). Serial correlation as a convenient simplification, not a nuisance: A comment on a study of the demand for money by the bank of england. *The Economic Journal*, *88*, 549-563.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, *12*, 231-254.

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

Kaufmann, R. K., & Stern, D. I. (2002). Cointegration analysis of hemispheric temperature relations. *Journal of Geophysical Research*, *107:2*, -.

Myhre, G., Myhre, A., & Stordal, F. (2001). Historical evolution of radiative forcing of climate. *Atmospheric Environment*, *35*, 2361-2373.

NASA Goddard Institute for Space Studies (GISS). (2011). *GISS - Surface Temperature Analysis.* Available on-line [http://data.giss.nasa.gov/gistemp/].

Stern, D. I., & Kaufmann, R. K. (2000). Detecting a global warming signal in hemispheric temperature series: A structural time series analysis. *Climatic Change*, *47*, 411-438.