Carbon Emissions and Economic Growth: Homogeneous Causality in Heterogeneous Panels

By David Maddison[a,]* and Katrin Rehdanz

[a] Department of Economics, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom, Email: d.j.maddison@bham.ac.uk.
[b] Department of Economics, Christian-Albrechts University of Kiel, Olshausenstraße 40, 24118 Kiel, Germany.
[c] The Kiel Institute for the World Economy, Düsternbrooker Weg 12, 24105 Kiel, Germany, email: katrin.rehdanz@ifw-kiel.de.

**Corresponding author.*

June, 2008

FNU Working Paper 163

Abstract

This paper introduces the concept of homogeneous non-causality in heterogeneous panels. This concept is used to examine a panel of data for evidence of a causal relationship between GDP and carbon emissions. The technique is compared to the standard test for homogeneous non-causality in homogeneous panels and heterogeneous non-causality in heterogeneous panels. In North America, Asia and Oceania the homogeneous non-causality hypothesis that CO_2 emissions does not Granger cause GDP cannot be rejected if heterogeneity is allowed for in the data-generating process. In North America the homogeneous non-causality hypothesis that GDP does not cause CO_2 emissions cannot be rejected either.

Keywords Energy; Carbon Emissions; Granger Causality; and Heterogeneous Panels

1. Introduction

Researchers have long attempted to determine the existence and direction of causal linkages between energy and income. The original impetus to this literature was concern over energy price rises, the finite nature of key energy resources and the presumed importance of providing energy to facilitate the development process. Apprehension about the environmental consequences of energy use has provided further stimulus to the literature testing for causality between energy use and income although given the widespread concern about climate change, the relationship between carbon and income is now presumably, of greater relevance.

A causal link running from energy use to income is usually interpreted as implying that attempts to reduce energy consumption would hamper economic growth, whereas causality running in the other direction implies that reducing economic growth will effect a reduction in energy use. Unidirectional causality from GDP to energy also indicates that energy forecasts can be prepared assuming that income is exogenous. No causal relationship indicates that development takes place independently of energy use and vice-versa. Bidirectional causality demonstrates the mutual interdependence of energy and income.[1]

For recent reviews of these studies the interested reader can consult Lee and Chang (2007), Yoo (2006) or Lee (2005). Virtually all recent contributors to the literature have however commented on the ambiguous findings generated by the research. More specifically, it appears that evidence on causality and the direction of causality depends critically on the countries considered, the time period under scrutiny and the empirical techniques employed.[2]

---

[1] The idea that energy causes income emerges from the neoclassical paradigm exemplified by the work of Jorgenson and Wilcoxen (1993). This approach emphasises the importance of energy in the production function alongside capital, labour and raw materials. Alternative paradigms such as the ecological approach also suggest that energy causes income. From an ecological perspective energy is required to extract low entropy materials and turn them into goods and high entropy waste (Cleveland et al, 1984).

[2] In his review of the literature Lee (2005) calls the evidence on causality "mixed and conflicting" and cites a number of papers providing conflicting results of the nature and direction of causality for the identical-same countries. Mahadevan and Asafu-Adjaye (2007) refer to the evidence as "mixed". Referring to tests of the

Granger causality tests also appear to produce mixed and conflicting results in other areas too, for example the question of whether increases in GDP causes growth in Government spending (sometimes known as the Wagner hypothesis), or whether the causality runs from Government spending to GDP growth as suggested by Keynesian economic theory. A further application of Granger causality testing has been to try to elucidate the causal relationship between FDI and economic growth. Here too, the results are often contradictory.

In an attempt to identify the reason underlying the often conflicting results Zachariadis (2006) compares some of the methods used by researchers to test for bivariate causality between energy and GDP including the two-step error correction technique (Engle and Granger, 1987), the autoregressive distributed lag method (Pesaran and Shin, 1999) and the vector autoregressive method (Toda and Yamamoto, 1995). Using identical datasets he confirms that different techniques can lead researchers to different conclusions when the sample size is limited.

It is precisely the problem of limited sample size that has encouraged researchers to combine cross-sectional and time-series data on energy and GDP hoping that more consistent causal relationships might emerge. In such applications fixed dynamic effects (FDE) techniques are typically used in tests of Granger causality involving nonstationary panel data. These analyses almost invariably constrain the parameters governing the short run dynamics and the long run cointegrating relationship to be identical.

But although the availability of panel data offers important advantages it may be that greater caution is required regarding its use. For example, in the dynamic panel data context it is important to ask whether different panels can be considered homogenous

relationship in transition countries Al-Iriani (2006) refers to the evidence as “mixed and sometimes conflicting” and in developing countries “not more conclusive than [that] obtained from developed and transition economies”. Soytas and Sari (2007) describe the research as having “failed to achieve unanimous results”.

because failure to account for heterogeneity may result in a heterogeneity bias (Pesaran and Smith, 1995).[3,4]

Admitting the existence of panel heterogeneity should also prompt questions about what precisely, the concept of Granger causality means in the context of panels. Hurlin and Venet (2005) distinguish between homogeneity in causal relationships and homogeneity in the data-generating process.

The purpose of this paper is to explore the concept of homogeneous Granger causality (or non-causality) in heterogeneous panels, to add to the existing concepts of homogeneous Granger causality (or non-causality) in homogeneous panels, and heterogeneous Granger causality (or non-causality) in heterogeneous panels. We propose a test of homogenous Granger non-causality in heterogeneous panels by assuming that the panels are drawn from a universe of possible panels which may or may not be characterised by Granger causal relationships. Each panel is separately tested for Granger causality, thereby avoiding any potentially unwarranted assumptions regarding parameter homogeneity. These probabilities are then pooled using techniques similar for example, to those used to combine the results of epidemiological studies.

[3] More specifically if the parameter of interest is the averaged effect of some exogenous variable on the dependent variable then the pooled estimator is not consistent in dynamic models even for large N and T. The reason is that when the regressors are serially correlated incorrectly ignoring parameter heterogeneity generates serial correlation in the disturbances which in turn generates inconsistent estimates in models with lagged dependent variables. This source of inconsistency is distinct from that suffered by the fixed effects estimator in panels with small T. In contrast to the FDE approach the mean group (MG) estimator provides consistent although inefficient estimates of the average effect of exogenous variables on the dependent variable. Each panel is estimated separately and the mean values of the parameter estimates calculated along with their variances. This estimator allows for heterogeneity in both the short run dynamics and the cointegrating parameters. Somewhere between the FDE and the MG estimator, the pooled mean group (PMG) estimator (Pesaran et al, 1999) allows only for the possibility of heterogeneous short run dynamics. The MG estimator provides consistent estimates of the average effect of the independent variables. But a test of Granger causality cannot be conducted on the basis of the average effect since such averaging may conceal causal relationships in individual panels.

[4] Interestingly the heterogeneous panel approach has been used by Martinez-Zarzoso and Bengochea-Morancho (2004) to estimate an environmental Kuznets curve for carbon but not for the purposes of causality testing.

A number of papers invoke the concept of heterogeneous non-causality in heterogeneous-panels Granger causality after first testing for (and finding) panel heterogeneity e.g. Butt et al (2005) and Hood et al (2008) neither of which involve questions related to energy. Nair-Reichert and Weinhold (2001) also explore the question causality in heterogeneous panels using an approach based on the technique of random coefficients. Researchers examining causal relations between energy and income by contrast, frequently commence testing for heterogeneous non-causality in heterogeneous panels without first testing the hypothesis of panel homogeneity. Other papers test for homogeneous non-causality in homogeneous panels without first testing the assumption of panel homogeneity inviting the kind of biases discussed in footnote 3.[5]

Compared to the test of heterogeneous non-causality in heterogeneous panels, our test seems more attractive since it avoids having to admit the possibility that any Granger causality detected might exist in only one panel. There may moreover be some instances in which the assumption of homogeneous causality or non-causality in heterogeneous panels is justified. Indeed, we would argue that the causal relation between energy and GDP is likely to be characterised by homogeneity even when the data generating process is not e.g. because countries' economic structures differ.

We test the hypothesis of a heterogeneous non-causal relationship between carbon emissions and GDP in a heterogeneous panel data context against the alternative, that at least one of the panel members exhibits a Granger causal relationship. We also test the null of homogeneous non-causality against the hypothesis of causality in heterogeneous panels. This latter test is apparently, new to the literature (see Hurlin, 2008 for a recently proposed alternative). These results are presented alongside the more common tests of homogenous non-causality in homogeneous panels and a test for panel homogeneity.

To anticipate the main findings, it appears that the null hypothesis of no homogenous or heterogeneous causal relation between per capita carbon emissions and GDP per capita can usually be rejected. Nevertheless, there are several important geographical regions in

[5] Remarkably, none of the papers exploring causal relations between energy and income cited here conducted a test of panel homogeneity. For an example of a paper that tests for homogeneous non-causality in homogeneous panels without first testing for panel homogeneity see Al-Iriani (op cit).

which the results of the non-causality tests disagree with one another even when accounting for panel heterogeneity.

The remainder of the paper is organised as follows. The next section presents a review of the literature analysing causal links between GDP and carbon emissions. Section three describes panel data on GDP per capita and carbon emissions per capita for 134 countries from 1990 to 2005 and examines its time series properties. Section four describes in detail the approach to testing for causality in homogeneous and heterogeneous panels. Section five compares the results of conventional tests of Granger causality with the new test for homogeneous Granger non-causality in heterogeneous panels. The final section concludes.

2. Literature review

Given the fact that there are several recent reviews of the literature on the causal relationship between energy use and GDP, and the fact that the causal relationship between carbon and GDP is arguably of greater use, we confine our literature review to the latter.

So far only four papers have examined the causal relationship between carbon emissions and income. The evidence is thus extremely limited when compared to that dealing with either the existence of causal relationships between energy and income, or to the existence of Kuznets curve-type relationships between carbon and income.[6]

Using VAR methods on differenced data Coondo and Dinda (2002) examine the existence of causal relationships between income and carbon emissions using cross country panel data. The results do not provide much evidence for the existence of a universal causal relationship between income and carbon emissions. Instead, the authors find that in some regions there is a causal relationship running from carbon to income and in other regions the causal relationship runs from income to carbon. In Africa and Asia bidirectional causality is observed.

Dinda and Coondoo (2006) re-examine the evidence using more modern time series econometric techniques. Surprisingly, they reject the null hypothesis of a long run cointegrating relationship for the North America, South America, Asia and Oceania country groupings. For the remaining country groupings (as well as for the panel in its entirety), the evidence strongly points to the existence of bidirectional causality.

Using the two-step error correction approach, Soytas et al (2007) study long run causality between carbon emissions, energy use and income in the US. Once again, they find no evidence of causality between either income and carbon emissions, or income and energy use. By means of the method of Toda and Yamamoto (1995) Soytas and Sari (2007) discover that carbon emissions Granger cause energy use in Turkey. As the authors

[6] Unfortunately, just like the literature on causal relations between energy and income, the environmental Kuznets curve literature has similarly failed to reach firm conclusions regarding whether pollution bears an inverse U-shaped relationship with income (e.g. Stern, 2004). The key difference between the two strands of literature is that even if it exists, a Kuznets curve relationship does not reveal the direction of causality.

themselves remark, this result appears to be counter intuitive. The same study suggests that neither carbon emissions nor energy use Granger cause income.

The presumption that carbon emissions cause economic growth is thus hardly borne out by the limited evidence available.

3. Description of the data

The data is taken from the IEA (2007). It includes both GDP per capita (GDPPC) measured in year 2000 USD and CO_2 per capita (CO2PC) measured in metric tonnes. The data covers 134 countries and runs from 1990 to 2005. The countries are furthermore grouped into income quartiles as well as by geographic area. These include: Western Europe; Eastern Europe; Africa; Asia; Middle East; Latin America; North America; and Oceania. The data is described in Table 1 and the membership of the country groupings is rendered explicit in Annex 1.[7]

The data are subject to a variety of tests for stationarity. Consistent with our concerns about panel heterogeneity, the test of Im et al (2003) and the test of Hadri (2000) are both employed. The Im et al test indicates that when individual effects and deterministic trends are included the null hypothesis of a unit root in all panels is occasionally strongly rejected. The test of Hadri by contrast, indicates that the null hypothesis of stationarity is always strongly rejected in favour of the alternative hypothesis indicating a common unit root. Henceforth it is assumed that the CO2PC and GDPPC data is characterised by a unit root with a deterministic trend, but it is clear that the tests for stationarity yield mixed messages.

The Kao (1999) and Pedroni (1999) tests are used to establish the existence of a long run cointegrating relationship between GDPPC and CO2PC. Although Pedroni provides a suite of seven different tests for cointegration in panel data, Pedroni (2004) establishes the greater power of the Panel ADF and Group ADF tests in small samples. The results of these tests are displayed in Table 3. With the exception of countries in the third income quartile, Western Europe and North America, the tests indicate that the null hypothesis of no cointegration in at least one panel can be rejected. The test of Kao by comparison uniformly rejects the null hypothesis of no cointegration. In what follows a cointegrating relationship is assumed to be present. It is interesting to note that when the Pedroni test is conducted on the dataset as a whole the finding is that every single panel is cointegrated. But when the data are subdivided into groups, the same test suggests that there are panels which are not cointegrated.

[7] A small number of observations have been linearly interpolated or extrapolated.

Finally we have also conducted tests for non-stationarity and cointegration on each panel separately. These indicate that the null hypothesis of non stationarity and the null hypothesis of no cointegration are very seldom rejected, no more frequently than might occur through chance in a dataset containing 134 panels. These tests typically are available on request from the authors.

These findings are very different to those obtained by Dinda and Coondoo (2006) who failed to find cointegration for several large country groupings. This may be because the authors use country groupings that differ in terms of membership to those employed here or because the data itself is different. Also, unlike this paper the authors do not use the now standard Pedroni test for panel cointegration. Instead, they obtain the residuals from an estimate of the panel specific long run relationships linking GDP and carbon emissions per capita and then subject these to the Im et al (op cit) unit root test. This procedure provides an invalid test of cointegration.

The failure to find a cointegrating relationship for major country groupings meant that Dinda and Coondoo were unable to proceed to the next stage of estimating the error correction model used to test for causality for several important regions.

Table 1. Description of the data

Number of countries = 134

Number of time periods = 16

Variable	Definition	Mean	Std Dev	Min	Max
GDPPC	GDP per capita (000s USD)	9811.08	9661.87	490.23	59250.00
CO2PC	CO_2 per capita (tonnes)	5.51	6.68	0.03	57.92
YEAR	Calendar year	1997.5	4.61	1990	2005

Source: IEA (2007).

Table 2. Results of Panel Stationarity Tests

Group	Variable	IPS Test	Hadri Test
All	log GDPPC	-5.07***	23.31***
	log CO2PC	-2.58***	20.97***
Q1 Income	log GDPPC	-0.36	12.04***
	log CO2PC	-1.17	10.08***
Q2 Income	log GDPPC	-1.58*	12.16***
	log CO2PC	-0.14	11.30***
Q3 Income	log GDPPC	-6.37***	10.73***
	log CO2PC	-1.10	10.04***
Q4 Income	log GDPPC	-1.76**	8.99***
	log CO2PC	-2.63***	10.36***
W Europe	log GDPPC	-1.88**	4.05***
	log CO2PC	-2.50***	7.79***
E Europe	log GDPPC	-6.88***	11.51***
	log CO2PC	0.52	10.16***
Africa	log GDPPC	-0.23	8.84***
	log CO2PC	-2.44***	7.71***
Asia	log GDPPC	0.10	6.62***
	log CO2PC	-0.21	8.46***
Middle East	log GDPPC	0.27	4.59***
	log CO2PC	-1.17	6.73***
Latin America	log GDPPC	-2.28**	7.51***
	log CO2PC	0.26	6.64***
N America	log GDPPC	-0.26	1..89**
	log CO2PC	0.31	2.62***
Oceania	log GDPPC	-1.39*	1.51**
	log CO2PC	-2.51***	1.75**

Note: The Im, Pesaran and Shin (IPS) test includes individual effects and individual trends. The optimal lag length is chosen on the basis of the Schwartz Information Criterion. The null hypothesis is the existence of individual unit root processes in all *panels. The Hadri test includes individual effects and individual trends. The null hypothesis is stationarity. *** denotes significance at the one percent level of confidence; ** denotes significance at the five percent level; and * denotes significance at the one percent level.*

Table 3. Results of Panel Cointegration Tests

	Kao Test	Pedroni Test	
Group		Panel ADF	Group ADF
All	5.03***	-11.51***	-12.43***
Q1 Income	3.40***	-4.38***	-6.49***
Q2 Income	1.46*	-5.71***	-4.52***
Q3 Income	6.74***	-0.38	-0.88
Q4 Income	5.18***	-9.43***	-4.15***
W Europe	10.02***	-3.36	-1.23
E Europe	2.43***	-4.59***	-7.56***
Africa	6.80***	-6.56***	-8.82***
Asia	4.23***	-3.76*	-3.51***
Middle East	6.55***	-5.77***	-4.03***
Latin America	7.59***	-2.16***	-4.19***
N America	7.15***	-0.84	-0.64
Oceania	6.60***	-2.84**	-2.03*

*Note: The Pedroni tests assume an individual deterministic trend and intercept. The null hypothesis is that of no cointegration in at least one panel. The Kao test assumes an individual intercept. The null hypothesis is no cointegration. *** denotes significance at the one percent level of confidence; ** denotes significance at the five percent level; and * denotes significance at the one percent level.*

3. Testing for homogeneous and heterogeneous non-causality

The basis for assessing causal relationships in this paper is the pooled two-step Engle Granger procedure. The majority of analyses examining causal relationships between carbon or energy, and GDP also employ this approach allowing for FDE in a limited attempt to control for panel heterogeneity. Remaining parameters are assumed to be homogeneous across panels.

Testing for homogeneous non-causality against the alternative hypothesis of homogeneous causality whilst assuming FDE and a logarithmic relationship between GDPPC and CO2PC entails running the following regression:

$$\Delta \ln GDPPC_{jt} = \alpha_j + \beta YEAR_t + \sum_{i=1}^{i=n} \gamma_i \Delta \ln \mathrm{GDPPC}_{\mathrm{jt\text{-}i}} + \sum_{\mathrm{i}=1}^{i=n} \delta_i \Delta \ln CO2PC_{jt-i} + \zeta ECT_{jt-1} + \varepsilon_{jt}$$

Where j signifies the country, t the time period and the maximum number of lagged variables is i=n.[8] The Greek letters α_j, β, γ_i, δ_i and ζ are parameters to be estimated. The error correction term ECT corresponds to the residual u from the long run relationship:

$$\ln GDPPC_{jt} = \phi_j + \varphi \ln CO2PC_{jt} + u_{jt}$$

Testing the hypothesis of homogeneous non-causality involves testing whether $\zeta = \delta_i = 0$.

In order to test reverse causality i.e. the hypothesis that in the long run GDPPC causes CO2PC the following regression is required:

$$\Delta \ln CO2PC_{jt} = \alpha_j + \beta YEAR_t + \sum_{i=1}^{i=n} \gamma_i \Delta \ln \mathrm{GDPPC}_{\mathrm{jt\text{-}i}} + \sum_{\mathrm{i}=1}^{i=n} \delta_i \Delta \ln CO2PC_{jt-i} + \zeta ECT_{jt-1} + \varepsilon_{jt}$$

Testing for homogeneous non-causality against the alternative hypothesis of homogeneous causality once more involves determining whether $\zeta = \gamma_i = 0$.

---

[8] In view of the relatively short time period covered by the data we generally assume that i = 1 in the analysis that follows. We will also on occasion investigate longer lag lengths.

All this presupposes that the panel is indeed homogenous. Slope homogeneity can be readily assessed using conventional F-tests based on the sum of squared residuals from the constrained and unconstrained models:

$$\Delta \ln GDPPC_{jt} = \alpha_j + \beta YEAR_t + \sum_{i=1}^{i=n} \gamma_i \Delta \ln \mathrm{GDPPC}_{\mathrm{jt\text{-}i}} + \sum_{\mathrm{i=1}}^{i=n} \delta_i \Delta \ln CO2PC_{jt-i} + \zeta ECT_{jt-1} + \varepsilon_{jt}$$

Compared to:

$$\Delta \ln GDPPC_{jt} = \alpha_j + \beta_j YEAR_t + \sum_{i=1}^{i=n} \sum_{j=1}^{j=m} \gamma_{ji} \Delta \ln \mathrm{GDPPC}_{\mathrm{jt\text{-}i}} + \sum_{\mathrm{i=1}}^{i=n} \sum_{j=1}^{j=m} \delta_{ji} \Delta \ln CO2PC_{jt-i} + \zeta_j ECT_{jt-1} + \varepsilon_{jt}$$

Where m is the number of countries. If the panel is found to be heterogeneous then a test for heterogeneous non-causality may once more, be based on comparing the sum of squared residuals from the constrained and unconstrained models:

$$\Delta \ln GDPPC_{jt} = \alpha_j + \beta_j YEAR_t + \sum_{i=1}^{i=n} \sum_{j=1}^{j=m} \gamma_{ji} \Delta \ln \mathrm{GDPPC}_{\mathrm{jt\text{-}i}} + \varepsilon_{jt}$$

Compared to:

$$\Delta \ln GDPPC_{jt} = \alpha_j + \beta_j YEAR_t + \sum_{i=1}^{i=n} \sum_{j=1}^{j=m} \gamma_{ji} \Delta \ln \mathrm{GDPPC}_{\mathrm{jt\text{-}i}} + \sum_{\mathrm{i=1}}^{i=n} \sum_{j=1}^{j=m} \delta_{ji} \Delta \ln CO2PC_{jt-i} + \zeta_j ECT_{jt-1} + \varepsilon_{jt}$$

The alternative hypothesis is that causal relationships are present in at least one of panels.

The method employed to test for homogeneous non-causality in heterogeneous panels is somewhat different. It involves estimating separately the long run relationship for each unit and obtaining the residuals in the conventional manner. More specifically we allow the long run cointegrating parameter φ to vary across panels:

$$\ln GDPPC_{jt} = \phi_j + \varphi_j \ln CO2PC_{jt} + u_{jt}$$

For each individual panel the error correction model is then estimated and subjected to a test of Granger causality by means of a conventional F-test. This process results in m separate P-values denoted P_j. These P values are then combined to test the null hypothesis of homogeneous non-causality against the alternative hypothesis of homogeneous causality. The technique used to combine the P values is that proposed by Fisher (1948):

$$\chi^2_{2j} = -2\sum_j \log P_j$$

Note that this technique does not rely on the assumption of homoscedastic error variances across panels. This assumption can itself be tested using Bartlett's test (Bartlett, 1937). Note also that the test can be statistically significant even when none of the constituent P-values are statistically significant and statistically insignificant when some of the P-values are significant.

In the next section we compare different tests of panel causality using either adjusted F-tests that are robust to the problem of heteroscedastic panels, or which do not depend on homoscedasticity.

4. Results

The results of the test of homogeneous non-causality in homogeneous panels are presented in Table 4. Remember that these results assume slope homogeneity in the short run dynamics as well as in the long run cointegrating parameters.

Pooling all 134 countries the data suggests a bidirectional causal relationship between CO2PC and GDPPC. Examining the results for countries divided into income quartiles also reveals strong evidence in favour of a bidirectional causal relationship between CO2PC and GDPPC. Turning to the geographical groupings there is once more a strong bidirectional relationship between CO2PC and GDPPC in all regions apart from North America and Latin America where the hypothesis that GDPPC does not CO2PC cannot be rejected even at the ten percent level of confidence. We note in passing that such findings already differ greatly from those reported in Dinda and Coondoo (op cit).

We have also explored the possibility that the statistical insignificance of the test that GDPPC causes CO2PC for Latin America might be due to the short lag length. Increasing the lag length to n=3 however, results in a test statistic of $F(4, 245) = 1.79$ which is still statistically insignificant at the 10 percent level. Increasing the lag length for North America results in an F statistic of only $F(4, 14) = 0.39$. The corresponding test that CO2PC causes GDPPC in North America with an assumed lag length of n=3 is $F(4, 14) = 1.50$ which is again, not significant.

Table 5 presents a series of tests for parameter homogeneity. Other than in the case of Western Europe the results indicate that the hypothesis of parameter homogeneity can be readily rejected, typically at the one percent level of confidence. This suggests that it is inappropriate to base any results concerning causal relations based on the assumption of parameter homogeneity.

Table 6 presents results for the heterogeneous-panels test of heterogeneous non-causality. These indicate that the hypothesis of heterogeneous non-causality can be rejected not only for the World as a whole, but also for the four income quartiles. Heterogeneous non-causality can also be rejected for the vast majority of country groupings except for North America where the hypothesis that CO2PC does not cause GDPPC cannot be rejected at any conventional level of significance. Unlike with the test for homogeneous causality in homogeneous panels the heterogeneous non-causality test statistics for Latin America are now both highly significant. Note that we do not speak of bidirectional causal relationships in heterogeneous non-causality in heterogeneous panels: In bivariate regressions some panels may exhibit Granger causality running in one direction, whereas an entirely different set of panels may exhibit Granger causality running in the other direction. Such panels cannot be described as demonstrating 'bidirectional' causality.

Finally in Table 7 we test the hypothesis of homogeneous non-causality in heterogeneous panels. Once more the results indicate that the null hypothesis of homogeneous non-causality can typically be rejected with a high degree of confidence. There is however one important change: the test statistic for Asia is now statistically insignificant even at the ten percent level of confidence.

Although this is the only result that changes, with China and India this grouping includes the world's two most populous countries. It is expected that these two countries will be responsible for much of the anticipated global increase in CO2 emissions. If we believe that causal relations are homogeneous over that region then the surprising implication is that curtailing carbon emissions there will not reduce economic growth (or at least would not have done so over the period 1990-2006). The homogeneous non-causality test statistics for North America are statistically insignificant at the 10 percent level of confidence.

We have once more examined the possibility that the statistically insignificant results for Asia and for North America might be altered by selecting a longer lag length.[9] With a lag length of n=3 the Fisher test of the hypothesis that CO2 causes GDP for Asia is statistically insignificant at the 10 percent level of confidence with a chi-square statistic of 41.06. The equivalent result for North America is 3.15, which is also insignificant at the 10 percent level. We also explored the hypothesis that GDP causes CO2 in North America using a lag length of n=3. The chi-squared statistic of 2.50 is once more statistically insignificant at the 10 percent level of confidence.

[9] The Fisher test can easily combine P-statistics generated by regressions employing different lag lengths whose dimensions might be suggested by reference to some criterion.

Table 4. Test of Homogeneous Non-Causality in Homogeneous Panels

	GDPPC → CO2PC	CO2PC → GDPPC
All	$F(2, 1738) = 27.15$***	$F(2, 1738) = 11.51$***
Q1 Income	$F(2, 438) = 10.79$***	$F(2, 438) = 8.29$***
Q2 Income	$F(2, 438) = 14.92$***	$F(2, 438) = 2.36$*
Q3 Income	$F(2, 425) = 30.48$***	$F(2, 425) = 11.07$***
Q4 Income	$F(2, 425) = 18.77$***	$F(2, 425) = 5.90$***
W Europe	$F(2, 282) = 16.40$***	$F(2, 282) = 4.66$**
E Europe	$F(2, 347) = 21.97$***	$F(2, 347) = 7.91$***
Africa	$F(2, 334) = 3.51$**	$F(2, 334) = 11.06$***
Asia	$F(2, 243) = 10.76$***	$F(2, 243) = 3.05$**
Middle East	$F(2, 165) = 6.30$***	$F(2, 165) = 6.47$***
Latin America	$F(2, 295) = 1.48$	$F(2, 295) = 13.17$***
North America	$F(2, 22) = 2.47$	$F(2, 22) = 0.73$
Oceania	$F(2, 22) = 8.02$***	$F(2, 22) = 2.58$*

*Note: These results assume slope homogeneity in the short run coefficients and the long run cointegrating parameters. The null hypothesis is homogeneous non-causality and the alternative is homogeneous causality. *** denotes significance at the one percent level of confidence; ** denotes significance at the five percent level; and * denotes significance at the one percent level.*

Table 5. Test of Panel Heterogeneity

	Dependent Variable	
	GDPPC	CO2PC
All	F(532, 1206) = 3.17***	F(532, 1206) = 1.51***
Q1 Income	F(132, 306) = 2.66***	F(132, 306) = 1.39**
Q2 Income	F(132, 306) = 2.86***	F(132, 306) = 1.50***
Q3 Income	F(128, 297) = 3.25***	F(128, 297) = 1.60***
Q4 Income	F(128, 297) = 2.49***	F(128, 297) = 2.04***
W Europe	F(84, 198) = 0.64	F(84, 198) = 0.91
E Europe	F(104, 243) = 3.64***	F(104, 243) = 0.96
Africa	F(100, 234) = 1.72***	F(100, 234) = 1.53***
Asia	F(72, 171) = 1.70***	F(72, 171) = 2.28***
Middle East	F(48, 117) = 1.59**	F(48, 117) = 3.97***
Latin America	F(88, 207) = 1.24	F(88, 207) = 1.50***
North America	F(4, 18) = 3.60**	F(4, 18) = 0.39
Oceania	F(4, 18) = 2.62*	F(4, 18) = 2.97**

*Note: The null hypothesis is homogeneity for the short run coefficients and the alternative is slope heterogeneity for at least one of the short run coefficients. *** denotes significance at the one percent level of confidence; ** denotes significance at the five percent level; and * denotes significance at the one percent level.*

Table 6. Test of Heterogeneous Non-Causality in Heterogeneous Panels

	Dependent Variable	
	GDPPC	CO2PC
All	F(268, 1206) = 7.10***	F(268, 1206) = 5.37***
Q1 Income	F(68, 306) = 6.25***	F(68, 306) = 4.52***
Q2 Income	F(68, 306) = 6.24***	F(68, 306) = 7.09***
Q3 Income	F(66, 297) = 7.28***	F(66, 297) = 4.58***
Q4 Income	F(66, 297) = 6.68***	F(66, 297) = 6.04***
W Europe	F(44, 198) = 1.80***	F(44, 198) = 3.51***
E Europe	F(54, 243) = 9.41***	F(54, 243) = 5.57***
Africa	F(52, 234) = 4.47***	F(52, 234) = 4.46***
Asia	F(38, 171) = 3.37***	F(38, 171) = 5.55***
Middle East	F(26, 117) = 4.91**	F(26, 117) = 9.33**
Latin America	F(46, 207) = 3.26***	F(46, 207) = 3.70***
North America	F(4, 18) = 3.97**	F(4, 18) = 0.82
Oceania	F(4, 18) = 3.53**	F(4, 18) = 6.61***

*Note: These results assume slope heterogeneity in the short run coefficients and the long run cointegration parameters. The null hypothesis is heterogeneous non-causality and the alternative is heterogeneous causality for at least one of the panels. *** denotes significance at the one percent level of confidence; ** denotes significance at the five percent level; and * denotes significance at the one percent level.*

Table 7. Fisher Test of Homogeneous Non-Causality in Heterogeneous Panels

	GDPPC → CO2PC	CO2PC → GDPPC
All	$\chi^2(268) = 602.07$***	$\chi^2(268) = 514.13$***
Q1 Income	$\chi^2(68) = 176.83$***	$\chi^2(68) = 118.61$***
Q2 Income	$\chi^2(68) = 108.90$***	$\chi^2(68) = 105.11$***
Q3 Income	$\chi^2(66) = 147.09$***	$\chi^2(66) = 161.03$***
Q4 Income	$\chi^2(66) = 169.25$***	$\chi^2(66) = 129.39$***
W Europe	$\chi^2(44) = 109.37$***	$\chi^2(44) = 67.59$**
E Europe	$\chi^2(54) = 110.04$***	$\chi^2(54) = 161.94$***
Africa	$\chi^2(52) = 135.60$***	$\chi^2(52) = 102.19$***
Asia	$\chi^2(38) = 75.86$***	$\chi^2(38) = 42.38$
Middle East	$\chi^2(26) = 69.61$***	$\chi^2(26) = 64.18$***
Latin America	$\chi^2(46) = 84.84$***	$\chi^2(46) = 67.42$**
North America	$\chi^2(4) = 4.34$	$\chi^2(4) = 4.39$
Oceania	$\chi^2(4) = 10.58$**	$\chi^2(4) = 4.04$

*Note: These results allow for slope heterogeneity in both the long run cointegrating parameters and the short run coefficients. The null hypothesis is homogeneous non-causality and the alternative homogeneous causality. *** denotes significance at the one percent level of confidence; ** denotes significance at the five percent level; and * denotes significance at the one percent level.*

5. Conclusions

Previous research into causal relationships between energy or carbon and GDP has reached differing conclusions regarding both the existence and the direction of causality. Some of these discrepancies might be explained by differences in the time period under scrutiny or by the application of differing empirical techniques to small, noisy samples. But some of the discrepancies might be caused by inappropriately pooling data from different countries thereby implicitly assuming that countries share the same long run cointegrating relationship and short run dynamics.

This paper introduces the concept of homogeneous non-causality in heterogeneous panels to set off the existing concepts of homogeneous causality in homogeneous panels, and heterogeneous causality in heterogeneous panels. Using data on GDPPC and CO2PC the paper compares various tests of causality and demonstrates that they can produce different results.

What does this research imply for current practice? The single most important advice is always to test for panel heterogeneity prior to conducting any kind of test for causality. The second piece of advice is for researchers to distinguish between homogeneous non-causality and heterogeneous non-causality both of which can characterise heterogeneous panels. Depending on the context one of these concepts might be more appropriate than the other, and could yield different results. Indeed, this paper finds that whilst the hypothesis of heterogeneous non-causality between CO2PC and GDPPC can be rejected for Asia the hypothesis of homogeneous non-causality cannot be rejected. Overall it seems likely that future researchers might wish to test for more than one type of causality in panels if these turn out to be heterogeneous.

References

Al-Iriani, A.M. (2006), Energy-GDP relationship revisited: An Example from the GCC Countries using Panel Causality, *Energy Policy*, Vol. 34, pp. 3342-3350.

Bartlett, M. S. (1937), Properties of Sufficiency and Statistical Tests, *Proceedings of the Royal Statistical Society Series A*, Vol. 160, pp. 268–282.

Butt, M.S., Iqbal, A., Ahmad, N., Akmal, M.S. and Ahmad, Q.M. (2006) Financial Development and Economic Growth in Asian Countries: a Heterogeneous Dynamic Panel Data Approach. *Journal of Social and Economic Policy*, Vol. 3, pp. 213-233.

Cleveland, C.J., Costanza, R. Hall, C.A.S., and Kaufman, R.F. (1984), Energy and the US Economy: A Biophysical Perspective, *Science*, Vol. 225, pp. 890-897.

Coondoo, D. and Dinda, S. (2002), Causality between Income and Emission: A Country Group Specific Econometric Analysis, *Ecological Economics*, Vol. 40, pp. 351-367.

Dinda, S. and Coondoo, D. (2006), Income and Emissions: A Panel Based Cointegration Analysis, *Ecological Economics*, Vol. 57, pp. 167-181.

Engle, R.F. and Granger, C.W.J. (1987), Cointegration and Error Correction: Representation, Estimation and Testing, *Econometrica*, Vol. 55, pp. 251-276.

Fisher, R.A. (1948), Combining Independent Tests of Significance, *American Statistician*, Vol. 2, pp. 30.

Hadri, K. (2000), Testing for Stationarity in Heterogeneous Panel Data, *Econometrics Journal*, Vol. 3, pp. 148-161.

Holtz-Eakin, D. and Selden, D.M. (1995), Stoking the Fires? CO2 Emissions and Economic Growth, *Journal of Public Economics*, Vol. 57, pp. 85-101.

Hood, M.V., Kidd, Q. and Morris, I.L. (2008) Two Sides of the Same Coin? Employing Granger Causality Tests in a Panel Framework. Forthcoming in *Political Analysis*.

Hurlin, C. (2008), Testing for Granger Non-Causality in Heterogeneous Panels, Mimeo. Department of Economics: University of Orleans.

Hurlin, C. and Venet, B. (2005) Testing for Granger Causality in Heterogeneous Panel Data Models, *Revue Economique*, Vol. 56, pp. 1-11.

Im, K.S., Pesaran, M.H., Shin, Y. (2003), Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics*, Vol. 115, pp. 53-74.

Jorgenson, D.W. and Wilcoxen, P.J. (1993), Reducing US Carbon Emissions: An Econometric General Equilibrium Assessment, *Resource and Energy Economics*, Vol. 15, pp. 7-25.

Kao, C. (1999), Spurious Regression and Residual-Based Tests for Panel Cointegration, *Journal of Econometrics*, Vol. 90, pp. 1-44.

Kraft, J. and Kraft, A. (1978), Relationships between Energy and GNP, *Journal of Energy and Development*, Vol. 3, pp. 401-403.

Lee, C-C. (2005), Energy Consumption and GDP in Developing Countries: A Cointegrated Panel Analysis, *Energy Economics*, Vol 27, pp. 415-427.

Lee, C-C. (2006), The Causality Relationship between Energy Consumption and GDP in G-11 Countries Revisited, *Energy Policy*, Vol. 34, pp. 1086-1093.

Lee, C-C. and Chang, C-P. (2007), Energy Consumption and Economic Growth in Asian Economies: A More Comprehensive Analysis using Panel Data, *Resource and Energy Economics* (forthcoming).

Levin, A., Lin, C. F. and Chu, C.S. (2002), Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties, *Journal of Econometrics*, Vol. 108, pp. 1-24.

Mahadevan, R. and Asafu-Adjaye, J. (2007), Energy Consumption, Economic Growth and Prices: A Reassessment using Panel VECM for Developed and Developing Countries, *Energy Policy*, Vol. 35, pp. 2481-2490.

Martinez-Zarzoso, I. and Bengochea-Morancho, A. (2004), Pooled Mean Group Estimation of an Environmental Kuznets Curve for CO_2, *Economics Letters*, Vol. 81, pp. 121-126.

Nair-Reichert, U. and Weinhold, D. (2001), Causality Tests for Cross-Country Panels: A New Look at FDI and Economic Growth in Developing Countries, *Oxford Bulletin of Economics and Statistics*, Vol. 63, pp. 153-171.

Pedroni, P. (1999), Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors, *Oxford Bulletin of Economic Statistics,* (Special Issue) pp. 653-669.

Pesaran, M.H. and Smith, R. (1995), Estimating Long Run Relationships from Dynamic Heterogeneous Panels, *Journal of Econometrics*, Vol. 68, pp. 79-113.

Pesaran, M.H., Shin, Y. and Smith, R. (1999), Pooled Mean Group Estimators of Dynamic Heterogeneous Panels, *Journal of the American Statistical Association*, Vol. 94, pp. 621-634.

Pesaran, M.H., Shin, Y. and Smith, R.J. (2001), Bounds Testing Approaches to the Analysis of Level Relationships, *Journal of Applied Econometrics*, Vol. 16, pp. 289-326.

Soytas, U. and Sari, R. (2007), Energy Consumption, Economic Growth and Carbon Emissions: Challenges Faced by an EU Candidate Member, *Ecological Economics* (forthcoming).

Soytas, U., Sari, R. and Ewing. B.T. (2007), Energy Consumption, Income and Carbon Emissions in the United States, *Ecological Economics*, Vol. 60, pp. 482-489.

Stern, D. (2004), Economic Growth and Energy. In Cleveland, C.J. et al (Eds.) *Encyclopedia of Energy Volume 2*, Elsevier, Amsterdam.

Toda, H.Y. and Yamamoto, T. (1995), Statistical Inference in Vector Autoregression with Possibly Integrated Processes, *Journal of Econometrics*, Vol. 66, pp. 225-250.

Yoo, S-H. (2006), The Causal Relationship between Electricity Consumption and Economic Growth in the ASEAN Countries, *Energy Policy*, Vol. 34, pp. 3573-3582.

Zachariadis, T. (2006), On the Exploration of Causal Relationships between Energy and the Economy, Discussion Paper 2006-05, Department of Economics, University of Cyprus.

Annex 1. Definitions of Country Groupings

Q1 low income countries: Angola; Bangladesh; Benin; Bolivia; Cameroon; Congo; Cote d'Ivoire; Dem, Rep, of Congo; DPR of Korea; Eritrea; Ethiopia; Georgia; Ghana; Haiti; India; Iraq; Kenya; Kyrgyzstan; Mozambique; Myanmar; Nepal; Nigeria; Pakistan; Republic of Moldova; Senegal; Serbia and Montenegro; Sudan; Tajikistan; Togo; United Rep, of Tanzania; Uzbekistan; Vietnam; Yemen; Zambia.

Q2 medium low income countries: Albania; Algeria; Armenia; Azerbaijan; Belarus; Bosnia and Herzegovina; Dominican Republic; Ecuador; Egypt; El Salvador; FYR of Macedonia; Guatemala; Honduras; Indonesia; Islamic Rep, of Iran; Jamaica; Jordan; Kazakhstan; Lebanon; Morocco; Namibia; Nicaragua; Panama; Paraguay; People's Rep, of China; Peru; Philippines; Sri Lanka; Syria; Tunisia; Turkmenistan; Ukraine; Venezuela; Zimbabwe.

Q3 medium high income countries: Argentina; Bahrain; Brazil; Bulgaria; Chile; Colombia; Costa Rica; Croatia; Cuba; Czech Republic; Estonia; Gabon; Hungary; Korea; Latvia; Libya; Lithuania; Malaysia; Malta; Mexico; Netherlands Antilles; Oman; Poland; Romania; Russia; Saudi Arabia; Slovak Republic; Slovenia; South Africa; Thailand; Trinidad and Tobago; Turkey; Uruguay.

Q4 high income countries: Australia; Austria; Belgium; Brunei; Canada; Chinese Taipei; Cyprus; Denmark; Finland; France; Germany; Gibraltar; Greece; Hong Kong, China; Iceland; Ireland; Israel; Italy; Japan; Kuwait; Luxembourg; Netherlands; New Zealand; Norway; Portugal; Qatar; Singapore; Spain; Sweden; Switzerland; United Arab Emirates; United Kingdom; United States.

Western European countries: Austria; Belgium; Denmark; Finland; France; Germany; Greece; Iceland; Ireland; Italy; Luxembourg; Netherlands; Norway; Portugal; Spain; Sweden; Switzerland; Turkey; United Kingdom; Cyprus; Gibraltar; and Malta.

Eastern European countries: Czech Republic; Hungary; Poland; Slovak Republic; Albania; Bulgaria; Romania; Bosnia and Herzegovina; Croatia; FYR of Macedonia; Slovenia; Serbia and Montenegro; Azerbaijan; Belarus; Estonia; Georgia; Kazakhstan; Kyrgyzstan;

Latvia; Lithuania; Republic of Moldova; Russia; Tajikistan; Turkmenistan; Ukraine; and Uzbekistan.

African countries: Algeria; Angola; Benin; Cameroon; Congo; Democratic Republic of Congo; Cote d'Ivoire; Egypt; Eritrea; Ethiopia; Gabon; Ghana; Kenya; Libya; Morocco; Mozambique; Namibia; Nigeria; Senegal; South Africa; Sudan; United Republic of Tanzania; Togo; Tunisia; Zambia; and Zimbabwe.

Asian countries: Bangladesh; Brunei; Chinese Taipei; India; Indonesia; Democratic People's Republic of Korea; Malaysia; Myanmar; Nepal; Pakistan; Philippines; Singapore; Sri Lanka; Thailand; Vietnam; People's Republic of China; and Hong Kong.

Middle Eastern countries: Bahrain; Islamic Republic of Iran; Iraq; Israel; Jordan; Kuwait; Lebanon; Oman; Qatar; Saudi Arabia; Syria; United Arab Emirates; and Yemen.

Latin American countries: Argentina; Bolivia; Brazil; Chile; Colombia; Costa Rica; Cuba; Dominican Republic; Ecuador; El Salvador; Guatemala; Haiti; Honduras; Jamaica; Mexico; Netherlands Antilles; Nicaragua; Panama; Paraguay; Peru; Trinidad and Tobago; Uruguay; and Venezuela.

North American countries: Canada and United States of America.

Oceania countries: Australia and New Zealand.