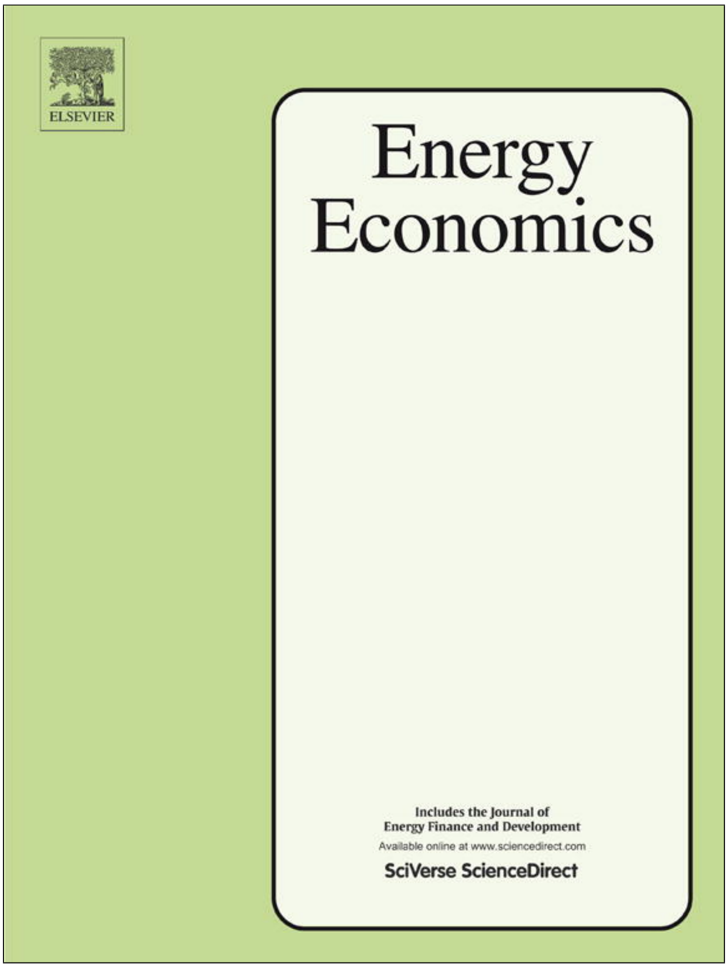
ELSEVIER
Energy Economics
Includes the Journal of Energy Finance and Development
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Energy Economics 39 (2013) 135–146

Contents lists available at SciVerse ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

Causality between energy and output in the long-run

David I. Stern [a,*], Kerstin Enflo [b]

[a] Crawford School of Public Policy, Australian National University, Canberra, ACT 0200, Australia
[b] Department of Economic History, Lund University, P.O. Box 7083, SE-220 07 Lund, Sweden

ARTICLE INFO

Article history:
Received 16 August 2012
Received in revised form 4 May 2013
Accepted 5 May 2013
Available online 12 May 2013

JEL classification:
Q43
C32
C52

Keywords:
Energy
Macroeconomics
Granger causality
Cointegration
Sweden

ABSTRACT

Though there is a very large literature examining whether energy use Granger causes economic output or *vice versa*, it is fairly inconclusive. Almost all existing studies use relatively short time series, or panels with a relatively small time dimension. We apply Granger causality and cointegration techniques to a Swedish time series dataset spanning 150 years to test whether increases in energy use and energy quality have driven economic growth or *vice versa*. We show that these techniques are very sensitive to variable definition, choice of additional variables in the model, sample periods and size, and the introduction of structural breaks. The relationship between energy and growth may also have changed over time – energy causes output in the full sample while output causes energy use in recent smaller samples. Energy prices have a more robust causal impact on both energy use and output.

1. Introduction

Does growth in energy use cause economic growth? Or does economic growth drive increasing energy consumption? There is a very large literature investigating these questions, but it is fairly inconclusive (Stern, 2011). In this paper, we apply Granger causality and cointegration techniques to a dataset covering 150 years of Swedish economic history. This time series is longer than any others that have been used previously in this literature. We show that these techniques are very sensitive to variable definition, choice of additional variables besides energy and output, sample periods, and structural breaks. All of the following appear to make a finding that energy causes growth more likely: using multivariate models, defining variables to better reflect their theoretical definition, using larger samples, and including appropriate structural breaks. However, it is also possible that the relationship between energy and growth has changed over time and that results from recent smaller samples reflect this. We find that energy prices have a more robust causal impact on both energy use and output.

Granger causality and cointegration methods have been extensively used to test for causal relations between the time series of energy, GDP, and other variables since the late 1970's (Kraft and Kraft, 1978; Ozturk, 2010). Early studies relied on Granger causality tests on unrestricted vector autoregressions (VAR) in levels of the variables, while more recent studies tend to use cointegration methods. Studies can also be distinguished by whether they use bivariate or multivariate models.

The results of early studies that tested for Granger causality using bivariate models were inconclusive (Stern, 1993). Where there were nominally significant results, they mostly indicated that output causes energy use. However, results differed across time periods, the countries investigated, and model specifications. Most economists believe that capital, labor, and technological change play a significant role in determining output, yet early studies implicitly assumed that energy is the only input to production. If this is not true, it will lead to omitted variables bias and, in the case of stochastically trending variables, non-cointegration and hence spurious and often sample dependent regression results (Stern and Common, 2001). In addition, samples were small, which results in higher sampling variability. These factors may explain the very divergent nature of much of the literature. In order to address the first of these issues, Stern (1993) estimated a VAR for GDP, capital, labor, and a Divisia index of energy use, finding that energy Granger caused GDP. But this was not the case for bivariate models or when the heat equivalent of energy was used in place of the quality adjusted index.

* Corresponding author. Tel.: +61 2 6125 0176.
E-mail addresses: david.stern@anu.edu.au (D.I. Stern), kerstin.enflo@ekh.lu.se (K. Enflo).

0140-9883/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.eneco.2013.05.007

Yu and Jin (1992) conducted the first cointegration study of the energy-GDP relationship. Again, the results of subsequent research vary widely. Stern (2000) estimated a cointegrating VAR for the variables included in Stern (1993), showing that there is a cointegrating relation between the four variables and that energy Granger causes GDP either unidirectionally or possibly bidirectionally. Warr and Ayres (2010) replicate this model for the U.S. using their measures of exergy and useful work in place of Stern's Divisia index of energy use. They find both short and long-run causality from either exergy or useful work to GDP but not *vice versa*. Oh and Lee (2004) and Ghali and El-Sakka (2004) apply Stern's (1993, 2000) methodology to Korea and Canada, respectively, coming to exactly the same conclusions. Lee and Chang (2008) and Lee et al. (2008) use panel data cointegration methods to examine the relationship between energy, GDP, and capital in 16 Asian and 22 OECD countries over three and four decade periods respectively. Lee and Chang (2008) find a long-run causal relationship from energy to GDP in the group of Asian countries while Lee et al. (2008) find a bidirectional relationship in the OECD sample. Taken together, this body of work suggests that the inconclusive results of earlier research are possibly due to the omission of non-energy inputs. By contrast, in recent bivariate panel data studies, Joyeux and Ripple (2011) find causality flowing from income to energy consumption for 56 developed and developing economies, while Chontanawat et al. (2008) find causality from energy to GDP to be more prevalent in the developed OECD countries compared to the developing non-OECD countries in a panel of 100 countries.

Other researchers have estimated multivariate VARs that include energy prices. Hamilton (1983) and Burbridge and Harrison (1984) found that changes in oil prices Granger-cause changes in GNP and unemployment whereas oil prices are exogenous. More recently, Blanchard and Galí (2008) used a VAR of GDP, oil prices, wages, and two other price indices, to argue that the effects of oil price shocks have reduced over time. Hamilton (2009a) deconstructs their arguments to show that past recessions would have been mild or have merely been slowdowns if oil prices had not risen. Furthermore, he argues that the large increase in the price of oil that climaxed in 2008 was a major factor in causing the 2008–2009 recession in the US. However, because it is hard to substitute other inputs for energy, the short-run elasticity of demand for oil and other forms of energy is low and the main short-run effects of oil prices on output are expected to be through reducing spending by consumers and firms on other goods, services, and inputs rather than through reducing the input of energy to production (Edelstein and Kilian, 2009; Hamilton, 2009a). Therefore, models using oil prices in place of energy quantities may not provide much evidence regarding the effects of energy use itself on economic growth.

Using a panel vector error correction model (VECM) model of GDP, energy use, and energy prices for 26 OECD countries (1978–2005), Costantini and Martini (2010) find that in the short run energy prices cause GDP and energy use and that energy use and GDP are mutually causative. However, they find that in the long-run GDP drives energy use and energy prices. Other researchers who model a cointegrating relation between GDP, energy, and energy prices for individual countries produce mixed results. For example, Glasure (2002) finds very similar results to Costantini and Martini (2010) for Korea, while Masih and Masih (1997) and Hondroyiannis et al. (2002) find mutual causation in the long run for Korea and Taiwan and Greece respectively. Following Stanley et al. (2010), we should probably put most weight on the study with the largest sample – Costantini and Martini (2010) – concluding that these models identify a demand function relationship where, in the long-run, GDP growth drives energy use.

Until very recently, all papers in this literature examined annual time series of a few decades at most, which is a small sample size for time series analysis, though researchers have also used panel data to try to increase statistical power. Two recent papers use much longer time series.[1] Vaona (2012) tests for causality between Malanima's (2006) data on Italian energy use and GDP from 1861 to 2000 using the Toda and Yamamoto (1995) procedure, the Johansen cointegration test, and Lütkepohl et al.'s (2004) cointegration test that allows for a shift in the mean of the process at an unknown time. Vaona disaggregates energy into renewable and non-renewable energy but only estimates bivariate VARs. The causality tests find mutual causation between non-renewable energy and GDP and from one measure of renewable energy to GDP. While the standard Johansen procedure does not find cointegration between GDP and non-renewable energy, the Lütkepohl et al. approach does find cointegration between these variables with a structural break in 1947.

Stern and Kander (2012) use 150 years of data for Sweden to estimate an econometric model with two equations – a nonlinear constant elasticity of substitution production function for the logarithm of gross output and capital, labor, and energy inputs, and an equation for the logarithm of the ratio of energy costs to non-energy costs. They estimate two specifications – one assumes that the rate of technological change was constant over the 150-year period and the other allows the rate to differ in each 50-year period. Using Choi and Saikkonen's (2010) nonlinear cointegration test, they find that the latter model cointegrates but the former does not. This implies that there is a causal relationship between the variables, but the direction of causality is unknown. In the current paper, we test for the direction of causality between energy and output in this Swedish dataset.

2. Granger causality testing

As is well known, correlation alone does not imply causation and so, without additional information, simple static regression analysis of observational data can only be used to estimate the partial correlations between variables or to compactly represent the joint probability distribution (Chen and Pearl, 2012). In this context, researchers must use theory to establish potential causal mechanisms (Gerring, 2010; Heckman, 2008), determine if variables are truly exogenous, and ensure that there are no confounding omitted variables. If the classical regression conditions do hold true, then the static regression model can be interpreted causally. More sophisticated techniques, including Granger causality testing, instrumental variables regression, and the potential outcomes framework (Ferraro and Hanauer, 2011), can be used to determine causal relationships under weaker conditions, though some assumptions are still needed.

Granger causality testing has been the most common approach to determining the causal validity of energy-output models. A variable, x, is said to Granger cause another variable, y, if its past values help predict the current level of y given all other relevant information. This definition is based on the concept of causal ordering. Two variables may be contemporaneously correlated by chance but it is unlikely that the past values of x will be useful in predicting y, given all the past values of y and other relevant information, unless x does actually cause y in a philosophical sense. Similarly, if y in fact causes x, then given the past history of y it is unlikely that information on x will help predict y. However, where a third variable, z, drives both x and y, but is omitted from the conditioning information, x might still appear to drive y, though there is no actual causal mechanism directly linking the variables. The simplest test of Granger causality requires estimating the bivariate VAR:

$$y_t = \beta_{1,0} + \sum_{i=1}^{p} \beta_{1,i} y_{t-i} + \sum_{i=1}^{p} \beta_{1,p+i} x_{t-i} + \varepsilon_{1t} \tag{1}$$

[1] The downside of using larger samples is that it potentially increases heterogeneity. The data generating process may change over time for long time series and vary across countries in the case of panel data. Though both Stern and Kander (2012) and Vaona (2012) allow for structural breaks in the deterministic time trend, other parameters may also change. Similarly, though panel data studies allow for country effects, other parameters may also vary across countries.

$$x_t = \beta_{2,0} + \sum_{i=1}^{p} \beta_{2,i} y_{t-i} + \sum_{i=1}^{p} \beta_{2,p+i} x_{t-i} + \varepsilon_{2t} \quad (2)$$

where p is the number of lags that adequately models the dynamic structure so that the coefficients of further lags of variables are not statistically significant and the error terms ε approximate white noise processes, which may be correlated across the equations. Deterministic time trends can also be added to the model. If the p parameters $\beta_{1,p+1}$ to $\beta_{1,2p}$ are jointly significant, then we can reject the null hypothesis that x does not Granger cause y. Similarly, if the p parameters $\beta_{2,1}$ to $\beta_{2,p}$ are jointly significant then the null hypothesis that y does not Granger cause x can be rejected. There are several other variants of this Granger causality test including the Sims (1972) causality test and the Toda and Yamamoto (1995) procedure discussed below.

Sargent (1979) and Sims (1980) introduced the VAR modeling approach as a method of carrying out econometric analysis with a minimum of *a priori* assumptions about economic theory (Qin, 2011). The VAR model generalizes the model given by Eqs. (1) and (2) to a multivariate setting. Standard multivariate Granger causality tests are identical to those described above except that there are lags of additional variables in each regression. The advantage of multivariate Granger tests over bivariate Granger tests is that they can help avoid spurious correlations by including additional variables that may be responsible for causing x or y or whose effects might obscure the effect of x on y or *vice versa* (Lütkepohl, 1982; Stern, 1993). Though a VAR cannot, due to limited degrees of freedom, include all variables that may be causally related to the principal variables under investigation, some attempt can be made to include as many as possible. There may also be indirect channels of causation from x to y and *vice versa*, which VAR modeling could uncover.

Of course, failure to reject the null hypothesis that x does not cause y, does not necessarily mean that there is in fact no causality. Low statistical power could be due to a misspecified lag length, insufficiently frequent observations (Granger, 1988), too small a sample (Wilde, 2012) or equivalently too small an effect size relative to the noise, omitted variables bias (Lütkepohl, 1982), and nonlinearity (Sugihara et al., 2012).

When some or all of the variables are non-stationary, a standard Granger causality test on a VAR in levels is invalid as the distribution of the test statistic is not the standard chi-square distribution (Ohanian, 1988; Toda and Phillips, 1993). This means that the significance levels reported in early studies of the Granger-causality relationship between energy and GDP may be incorrect, as both variables are generally integrated series. If there is no cointegration between the variables then the causality test should be carried out on a VAR in differenced data, while if there is cointegration, standard chi-square distributions apply when the cointegrating restrictions are imposed (Toda and Yamamoto, 1995). Toda and Yamamoto (1995) developed a modification of the standard Granger causality test on the variables in levels that is robust to the presence of unit roots. This method, described in detail below, adds additional lags of the variables to the VAR but these lags are not restricted in the Granger causality tests. This simple modification restores standard inference. Clarke and Mirza (2006) show that, despite the additional parameters, the Toda–Yamamoto test shows little loss of power compared to the alternative of testing the restrictions on a VECM that imposes cointegrating restrictions. Bauer and Maynard (2012) suggest an alternative test where only an extra lag of the variable being tested for exclusion is added. This procedure is robust to a wide array of data generating processes including structural breaks in the explanatory variables but not to I(2) variables. They find that the reduction in parameters increases power relative to the Toda–Yamamoto procedure across the data generating processes that they test in a Monte Carlo exercise.

Nonlinear Granger causality testing procedures exist such as the frequently used Hiemstra and Jones (1994) approach as well as more recent methods (e.g. Hassani et al., 2010; Sun, 2008). The Hiemstra and Jones test has been used to test for nonlinear causality between energy and output (Chiou-Wei et al., 2008) but is not applicable to non-stationary data (Hassani et al., 2010). As a result, Chiou-Wei et al. (2008) difference their data, but this throws away information on the long-run relationship between the variables.

3. Data

Some background on the Swedish economy will be useful for interpreting the data and the econometric results discussed below. Swedish industrialization and modern economic growth took off roughly around 1850 (Greasley et al., 2013), which is also the starting year for the analysis in this paper. Sweden went from being one of the poorest European economies in the early nineteenth century to one of the richest 150 years later.

Sweden is a small open economy and exports have contributed greatly to its economic success. Natural resources, such as the charcoal-based Swedish bar iron were the traditional export good and completely dominated Swedish exports until the mid-nineteenth century. The upswing in industrial growth in Western Europe in the 19th Century, led to an expansion of Swedish exports in three commodities – bar iron, wood, and oats. In order to connect Sweden's vast natural resources with the international market, state-sponsored railways were built starting in the 1850s. From the 1890s, the focus shifted towards new enterprises, which were closely related to the so-called Second Industrial Revolution. This meant that scientific knowledge and more complex engineering skills replaced the earlier dependence on natural resources. The electric motor became especially important and new companies such as ASEA (later ABB) were formed that combined engineering skills with the large supply of hydropower in Sweden (Schön, 2008).

Compared to many other industrialized nations, Sweden's energy system was never very dominated by coal, but rather went from dependence on firewood in 1850 (roughly 75 per cent of energy according to Gales et al., 2007) to becoming relatively dependent on primary electricity. Sweden is well endowed with hydropower resources and great advances were made in the electricity infrastructure from the 1910s to the 1950s. The national electrical grid was integrated in the 1930s and the technology of high voltage transmission made it possible to supply industries with electricity at lower prices and with great regularity. In 2000, primary electricity constituted around 30 per cent of the energy consumed in the country, half of which came from nuclear energy and half from hydropower (Gales et al., 2007). Reliance on oil came only after the Second World War.

During the interwar period, when the rest of Europe was torn by world wars and depression, the Swedish economy fared relatively well. The post-war period saw rapid economic growth in all of Europe, and Sweden was no exception. In addition, this period also saw the cementation of the “Swedish Model” that was said to be built on two main pillars: public responsibility for social security and the regulation of labor and capital markets. As long as the export sectors grew, the model worked well. However, in the 1970s and early 1980s Sweden was hit by the oil crisis and faced subsequent problems with structural adjustment of the economy. Industries such as steel, pulp and paper, shipbuilding, and mechanical engineering ran into crises and the Swedish Model began to disintegrate. The labor and capital markets were deregulated and the expansion of public sector services came to an end. During the last decades of the twentieth century, Swedish economic policy converged to European norms and this facilitated Swedish entry into the European Union in 1995 (Schön, 2010).

The data we use is identical to that used by Stern and Kander (2012), where a full description can be found. The energy data comes from Kander (2002) and the other data from the Swedish historical national accounts (Krantz and Schön, 2007). The variables considered in our models and tests are: Gross output (GRO), GDP, capital

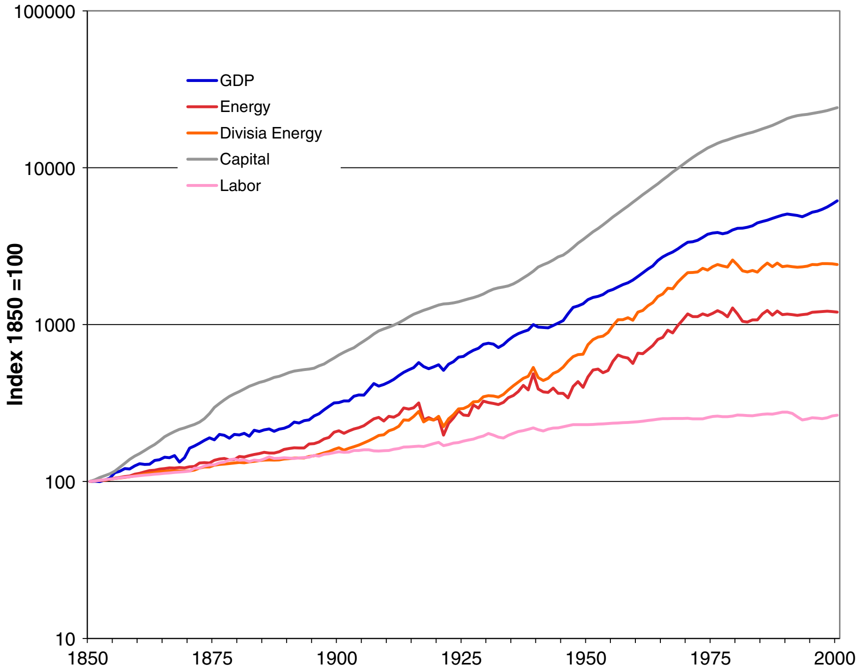

Fig. 1. Quantity variables: Sweden 1850–2000.

(K), labor (L), the heat content of primary energy (HE), a Divisia volume index of primary energy (DE),[2] the Divisia energy price index deflated by the GDP deflator (PE), and the oil price in Swedish Kroner deflated by the GDP deflator (PO). The reason for including the price of oil in our dataset is that it is more exogenous than the energy price index. However, the series only starts in 1885. We transform all these variables into logarithms. We use the Divisia energy volume index to take into account the increased productivity of energy over time due to the shift from coal and biomass to oil, natural gas, and primary electricity (Stern and Kander, 2012). Data quality is better for more recent decades, especially for the energy quantity data. Therefore, we carry out all tests for 1900–2000 and 1950–2000 sub-periods in addition to the full 150-year period.

Fig. 1 presents the time paths of the key quantity variables. As all the variables are strongly trending they are highly correlated. Fluctuations and changes in the trend slope also appear to be correlated. Fig. 2 compares the growth rates of the Divisia energy index (which is less volatile than the heat units of energy series) and GDP. The two series are strongly correlated in the mid-20th century. In the 19th century the energy series is much less volatile than the GDP series and the reverse is true in the late 20th century. The reason for this is that the 19th century data are dominated by renewable energy and the way that this data was constructed from the original sources did not focus on annual fluctuations (Stern and Kander, 2012). The decline in energy's cost share as the 20th century progressed might explain the change in relative volatilities over the course of the century. The simple correlation between the rates of change in Fig. 2 is 0.49, which is highly statistically significant ($t = 6.89$). The correlation between the rates of change is suggestive of a functional relationship but the direction of causation and the role of other variables are not indicated.

Fig. 3 shows the two price series – the real price of oil and the Divisia energy price index deflated by the GDP deflator. Oil is relatively expensive compared to the average energy carrier and its price is also much more volatile. In particular, the 1st and 2nd World Wars generated massive price spikes and a smaller spike follows the oil crisis of the 1970s. These two series are strongly correlated ($r = 0.56$). The direction of causation between the two series is pretty certain – oil prices are one component of the energy price index and are largely driven by global oil prices and exogenous disruptions such as the World Wars.

4. Econometric methods

4.1. Unit root tests

First we test for unit roots assuming that there are no structural breaks using the Phillips and Perron (1988) test (PP) and the Kwiatkowski et al. (1992) test (KPSS), which have null hypotheses of a unit root and stationarity, respectively. For the PP test we estimate the following regression:

$$y_t = \alpha + \beta t + \gamma y_{t-1} + u_t \tag{3}$$

where y is the log of the variable of interest. The null hypothesis is that y_t contains a unit root and so $\gamma = 1$. The PP test is a modified t-statistic for $\gamma = 1$. The alternative hypothesis is that y_t is trend stationary with slope β. We also estimate:

$$\Delta y_t = \alpha + \gamma \Delta y_{t-1} + v_t \tag{4}$$

where the alternative hypothesis is levels stationarity. We use the default four lags to compute the standard errors used in the PP test statistic by the RATS procedure unitroot.src.

We also test for unit roots assuming the presence of structural breaks. We assume both that the timing of the structural breaks is known – the breakpoints used by Stern and Kander (2012) – using Park and Sung's (1994) unit root test and that the timing is unknown,

[2] The heat content of primary energy is simply the total joules of the various forms of energy available to the economy before combustion of some fossil fuel and biomass to produce secondary electricity. Most electricity is primary (from nuclear, hydropower, etc.) in Sweden. Divisia indexation computes a "volume index" of energy input taking into account shifts between fuels with different productivities or "quality" as reflected in their prices (see Stern, 2010).

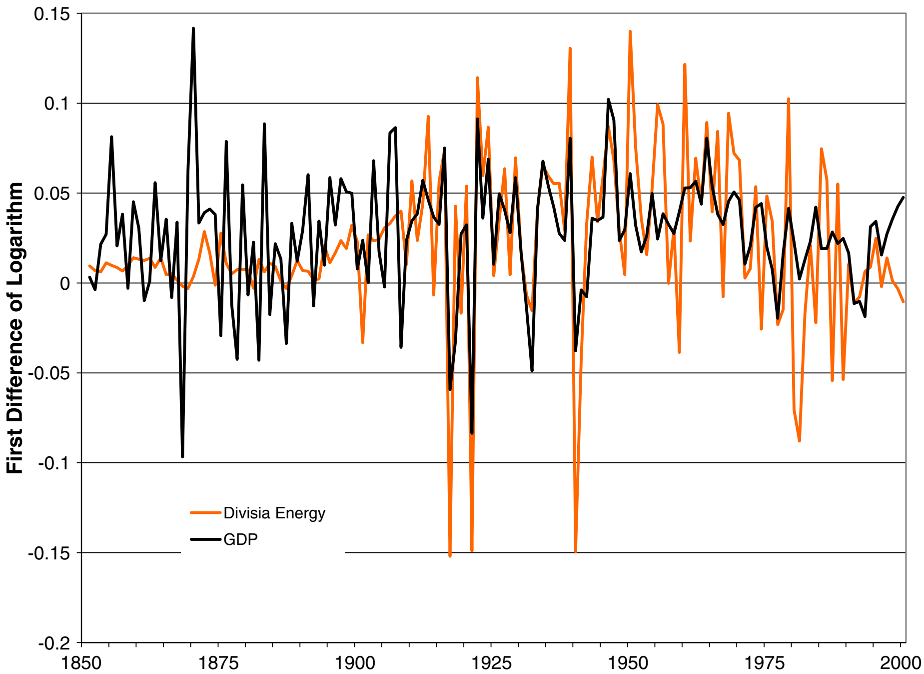

Fig. 2. Growth rates of GDP and Divisia index of energy use.

a priori, using the tests developed by Lee and Strazicich (2003, 2004). Park and Sung's (1994) tests modify Eqs. (3) and (4) to allow for breaks in the intercept and trend and to create test statistics that are invariant to the location of the breakpoints. Park and Sung (1994) provide the distribution of these test statistics for one or two breakpoints. Like Lee and Strazicich's test, Park and Sung's test allows for a structural break under the null hypothesis. For the log levels, the alternative hypothesis is trend stationarity with breaks – Lee and Strazicich's "break" model – while for the first differences of logs the alternative is levels stationarity with breaks – Lee and Strazicich's "crash" model. We used the RATS procedure lsunit.src to compute the Lee and Strazicich tests and we wrote the code for the Park and Sung test ourselves in RATS.

4.2. Granger causality tests

For all VAR models, including the cointegration models discussed in the next subsection, we select the optimal lag length, p, using the Akaike Information Criterion considering a maximum of four lags (Schwert, 1989). We use the Toda and Yamamoto (1995) procedure for testing for causality in the possible presence of unit roots and non-cointegration. We add additional lags of all variables to account

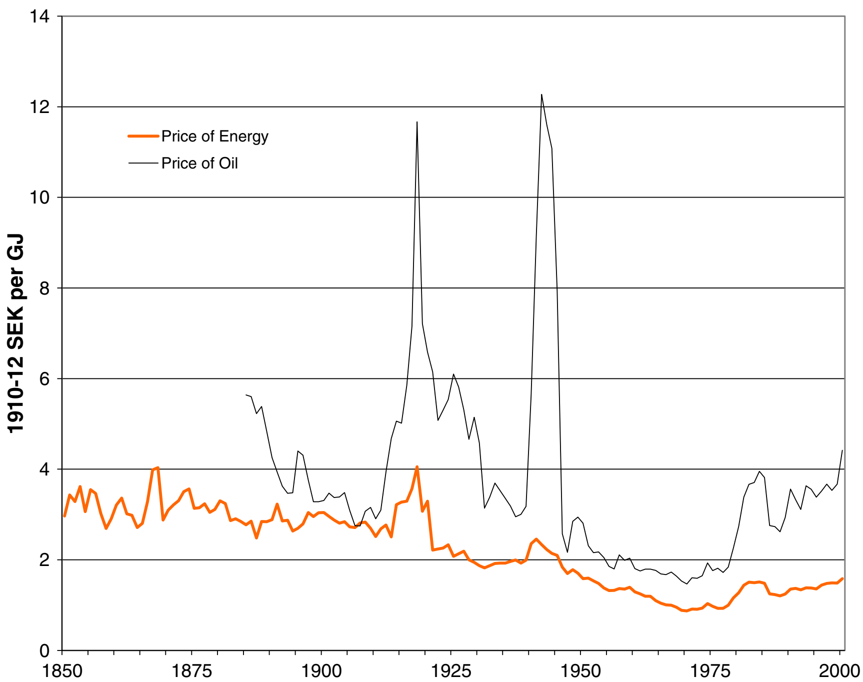

Fig. 3. Energy prices.

for possible unit roots in the time series and compute the Wald test statistic for excluding only the first p lags of the variable of interest from the relevant equation. Based on Joyeux (2007), the model we estimate is:

$$y_t = \sum_{j=0}^{n}\left(\alpha_j t_{jt-p} + \beta_j \Delta t_{jt-p} + \sum_{i=0}^{p-1}\gamma_{ji}\Delta^2 t_{jt-i}\right) + \sum_{i=1}^{p}\Pi_i y_{t-i} + \sum_{i=1}^{d}\Pi_{i+p} y_{t-p-i} + \varepsilon_t \qquad (5)$$

where d is the maximal order of integration in the data and n is the number of structural breaks in the data. y_t is the vector of m variables modeled in the VAR in year t and ε_t is the corresponding vector of disturbances. α_j, β_j, and γ_{ji} are $m \times 1$ vectors and the Π_i are $m \times m$ matrices of regression coefficients. The trend terms model a possible unobserved technology trend with n structural breaks. t_{0t} is a simple linear time trend and, therefore, its first difference, Δt_{0t}, is a constant. For $j > 0$, Δt_{jt} is equal to zero up to and including the year of the structural break and unity after it, so that t_{jt} is zero up to and including the year of the structural break and a linear time trend increasing by one unit per year thereafter. Therefore, the intercept term and the slope of the time trend in period t are $\sum_{k=0}^{j}\beta_j \Delta t_{jt}$ and $\sum_{k=0}^{j}\alpha_j \Delta t_{jt}$, respectively. The second differences of the time trend in Eq. (5) form intervention dummies that set the residuals of the first p observations in each subsample to zero, effectively excluding those observations from the estimation (Joyeux, 2007).

To test whether variable y^l causes variable y^k, where the superscripts indicate individual variables in the vector y, we need to test that $\Pi_1^{kl} = \Pi_2^{kl} = \ldots = \Pi_p^{kl} = 0$, where Π_i^{kl} is the element of the matrix Π_i in the kth row and lth column. We stack the matrices in Eq. (5) into a single matrix $\Pi = vec[\Pi_1, \Pi_2, \ldots, \Pi_{p+d}]$ and define R as a selector matrix so that $R\Pi = vec[\Pi_1^{kl}, \Pi_2^{kl}, \ldots, \Pi_p^{kl}]$. The null hypothesis can now be expressed as $R\Pi = 0$ and the Wald test statistic is:

$$W = \hat{\Pi}' R' \left[R\hat{V}R'\right]^{-1} R\hat{\Pi} \qquad (6)$$

where hats indicate estimated parameters and $\hat{V}$ is the estimated covariance matrix of Π. The test statistic is distributed asymptotically as chi-square with p degrees of freedom.

4.3. Cointegration modeling

Finally, we test for cointegration allowing for both deterministic trends and deterministic trends with structural breaks using the methodology of Johansen et al. (2000). We compute the critical values that correspond to the model that Johansen et al. (2000) call the $H_l(r)$ test, which means that we assume a structural change in both the unrestricted constant and the slope of the trend in the cointegration relation. The distribution of the critical values depends on the proportion of the way through the sample that the break occurs. We estimated the VECM models using E-Views implementing the structural breaks and associated cointegration tests using code provided by Giles and Godwin (2012). The purpose of this analysis is to see if there is a difference between linear cointegration analysis and Stern and Kander's (2012) nonlinear cointegration analysis as well as to test for the direction of causality in the cointegration framework.

5. Results

5.1. Unit root tests

Table 1 presents the results of the Phillips–Perron (PP) and Park and Sung (PS) unit root tests, Table 2 the Lee and Strazicich (LS) unit root tests, and Table 3 the KPSS unit root tests. Looking first at the PP tests on the log levels, the null of a unit root cannot be rejected for any series for either the complete period, 1850–2000, or either of

Table 1
Phillips and Perron and Park and Sung unit root tests.

Variable form	Log levels	First differences of logs	Log levels	First differences of logs
H0	Unit root	Unit root	Unit root	Unit root
H1	Trend stationary	Levels stationary	Trend stationary	Levels stationary
1850–2000 No structural breaks			1850–2000 Structural breaks in 1900 and 1950	
GRO	−2.03	**−13.15**	−3.18	**−13.86**
GDP	−2.15	**−12.26**	−2.72	**−12.36**
K	−1.19	**−3.52**	0.24	−3.08
L	−0.87	**−9.89**	−3.15	**−11.87**
HE	−2.50	**−15.71**	−4.08	**−15.96**
DE	−1.83	**−12.59**	−2.60	**−11.99**
PE	−2.52	**−13.23**	−4.00	**−13.85**
PO	−2.55	**−7.40**	−2.66	**−8.26**
1900–2000 No structural breaks			1900–2000 Structural break in 1950	
GRO	−1.93	**−9.16**	0.15	**−7.66**
GDP	−1.60	**−8.80**	0.10	**−7.08**
K	−0.73	−2.27	0.30	−1.70
L	−0.85	**−6.91**	−0.05	**−7.59**
HE	−2.17	**−12.92**	−1.81	**−11.59**
DE	−0.48	**−10.63**	−0.94	**−8.65**
PE	−1.45	**−10.18**	−2.02	**−7.79**
PO	−2.39	**−6.93**	−2.28	**−7.47**
1950–2000				
GRO	−0.99	**−4.48**		
GDP	−1.14	**−3.75**		
K	−0.02	−1.02		
L	−1.90	**−3.60**		
HE	−1.28	**−7.32**		
DE	−1.10	**−6.75**		
PE	−1.67	**−4.53**		
PO	−2.65	**−5.22**		

Notes: For definition of variables see the main text. For the price of oil the first observation is for 1885. Values significant at the 5% level are in bold. For the trend stationarity tests the critical value for the Phillips–Perron test at the 5% level is −3.45. For the Park and Sung test the critical values are −4.15 for one structural break and −4.75 for two structural breaks. For the levels stationarity tests the critical value for Phillips–Perron is −2.89 while for the Park and Sung tests they are −3.33 and −3.72.

Table 2
Lee and Strazicich unit root tests.

Variable form	Log levels			First differences of logs		
H0	Unit root			Unit root		
H1	Trend stationary, "Break"			Levels stationary, "Crash"		
	Test statistic	Breakpoint 1	Breakpoint 2	Test statistic	Breakpoint 1	Breakpoint 2
GRO	−3.97	1882	1961	**−6.26**	1890	1945
GDP	−4.16	1882	1961	**−5.44**	1904	1970
K	−4.12	1925	1962	−3.61	1934	1974
L	−3.55	1885	1936	**−5.61**	1890	1942
HE	−5.14	1948	1978	**−5.45**	1923	1980
DE	−3.60	1912	1959	**−6.22**	1945	1973
PE	−4.86	1912	1965	**−5.75**	1928	1973
PO	−4.76	1910	1950	**−5.90**	1921	1941

Notes: For definition of variables see the main text. For the price of oil the first observation is for 1885. Values significant at the 5% level are in bold. Exact critical values for the trend stationarity test depend on the location of the breakpoints and vary from −5.59 to −5.74. For the levels stationarity test the critical value is −3.84.

Table 3
KPSS unit root tests.

Variable	Log levels		Log first differences	
	H0: Levels stationary	H0: Trend stationary	H0: Levels stationary	H0: Trend stationary
1850–2000				
GRO	**3.10**	**0.58**	0.23	0.08
GDP	**3.11**	**0.50**	0.15	0.09
K	**3.09**	**0.37**	0.19	**0.18**
L	**3.08**	**0.46**	0.39	0.06
HE	**3.04**	**0.37**	0.08	0.08
DE	**3.02**	**0.59**	0.41	**0.26**
PE	**2.65**	**0.26**	0.09	0.09
PO	**0.73**	**0.17**	0.07	0.04
1900–2000				
GRO	**2.12**	**0.22**	0.09	0.09
GDP	**2.12**	**0.21**	0.12	0.11
K	**2.12**	**0.28**	0.31	**0.31**
L	**2.01**	**0.45**	0.31	0.06
HE	**2.06**	**0.21**	0.10	0.10
DE	**2.09**	**0.25**	0.30	**0.21**
PE	**1.64**	**0.26**	0.17	0.07
PO	**0.63**	**0.19**	0.06	0.06
1950–2000				
GRO	**1.09**	**0.27**	**0.53**	0.10
GDP	**1.08**	**0.27**	**0.48**	0.10
K	**1.09**	**0.29**	**0.97**	0.10
L	**0.83**	**0.20**	0.12	0.05
HE	**0.91**	**0.26**	**0.59**	0.07
DE	**0.93**	**0.28**	**0.82**	0.08
PE	0.26	**0.22**	0.35	0.07
PO	**0.75**	**0.19**	0.33	0.08

Notes: For definition of parameters and variables see the main text. Values significant at the 5% level are in bold. For the price of oil the first observation is for 1885.

the sub-periods.[3] However, we can reject the unit root null for all the differenced series except for capital in the two sub-periods. Allowing for structural breaks in 1900 and 1950 (PS tests) does not change this picture substantially. Neither does allowing endogenous selection of the breakpoints (LS tests, Table 2). The KPSS test (Table 3) easily rejects the null of trend stationarity for all the variables in log levels in all time periods. For the first differences of the variables, we cannot reject the null of levels stationarity for any variable for the full sample or the 1900–2000 subsample. However, levels stationarity can be rejected for several variables in the 1950–2000 period. We conclude that all series are I(1) apart from capital, which is possibly I(2) and, therefore, the Toda–Yamamoto test needs two extra lags.

The endogenous breakpoints (Table 2) differ across the variables and the levels and first difference specifications. We also found that the number of lags included in the procedure affected the choice of breakpoint. We also carried out LS tests with a single structural break (Lee and Strazicich, 2004) and three structural breaks. These resulted in a different selection of breakpoints that also differed across variables. Looking at Table 2, the obvious break in the energy quantity series following the oil price shock in the early 1970s only shows up in the first differences for DE and PE as well as for GDP and K. In levels, these series have breaks in the early- or mid-1960s, which are not at all visible in the data (Fig. 1). Given the disagreement across these tests we use exogenous breakpoints – the 1900 and 1950 breakpoints used by Stern and Kander (2012) and 1916 (First World War) and 1973 (Oil Crisis) breakpoints, which are apparent in the energy series.

[3] This contrasts with the finding of International Monetary Fund (2011, 91–92) that real oil prices from 1875 to 2010 are I(0). Our Swedish price series appears much less volatile in the short-run than their US dollar real oil price series. On the other hand, Hamilton (2009b) argues that real oil prices follow a random walk with no drift.

5.2. Toda–Yamamoto causality tests

We start by testing the simple bivariate model for GDP and the heat content of primary energy. We find (first two columns of Table 4) that GDP causes energy use but not *vice versa* in each sub-period. When we replace the heat content of energy with the Divisia index we find that there is causality from energy to GDP in the full sample ($p = 0.015$) and causality from GDP to energy in the 1950–2000 subsample but no causality in either direction for the 1900–2000 sub-period. This shows the sensitivity of bivariate tests to the definitions of variables.

Next, we estimate a multivariate VAR for GDP, capital, labor, and Divisia energy. This shows causality from energy to GDP for the full period ($p = 0.031$) and for 1900 to 2000 ($p = 0.037$). But GDP causes energy ($p = 0.000$) in the 1950–2000 period. When GDP is replaced with gross output, energy causes output in the full period and output causes energy in the 1950–2000 period at the 10% significance level, while there is no causation in either direction in the 1900–2000 subsample. So these multivariate results are also somewhat ambiguous. The final two columns of the table allow for a trend break in 1900 and 1950. This does not change the results much though it does increase the significance of the GDP causes energy test for the bivariate model using the Divisia index of energy use.

Table 5 shows the results of estimating VARs including GDP and the quantity and price of energy. For the first model with Divisia energy quantities and price index, the energy price index Granger causes energy use in all samples. GDP causes energy use in the 1950–2000 subsample but the significance level is much lower in the full sample and the 1900–2000 subsample. So there is only tentatively a demand function relationship in these data in the full sample. The Divisia price index Granger causes GDP in the full period and in the 1900–2000 sub-period but the quantity of energy has no significant effect on GDP. In the full sample, the Divisia price index is, however, endogenous with respect to the quantity of energy ($p = 0.019$) and GDP ($p = 0.053$) and GDP causes prices in the 1950–2000 subsample ($p = 0.072$).

Next, we replace the price of energy by the price of oil and the Divisia energy quantity index by the heat equivalent of energy. The price of oil is clearly exogenous as we would expect. The other main differences are that GDP causes energy use in all samples and the price of energy does not cause energy use at the 10% significance level in the 1950–2000 subsample. There is also more evidence of causation from energy to GDP.

When trends with structural breaks are used the results are very similar. The main difference is that now GDP causes Divisia energy volume in each sample so that there is stronger support for the demand function interpretation. We also added capital and labor to these latter models to produce a composite of the Tables 4 and 5 models. The results were very similar to the models in Table 5. The price of energy plays the dominant role in the models and capital and labor are mostly insignificant. We also estimated all the models with structural breaks in 1916 and 1973 instead of 1900 and 1950. The results were similar with generally lower significance levels.

5.3. Linear cointegration analysis

We estimate a VECM for capital, labor, energy, and output, measuring energy using either heat units or the Divisia index and output using GDP or gross output. We estimate models with and without linear trends in the cointegration space and with and without structural breaks. Finally, some versions include the Divisia energy price index and others do not. We assume that all variables are I(1). The results of the Johansen trace statistic tests for the number of cointegrating vectors are presented in Tables 6–11 and coefficient estimates for some of the models that passed the cointegration tests are presented in Tables 12 and 13.

Table 4
Causality tests: Production function models.

Model	Period	Simple time trend		Time trend with structural breaks	
		Energy- > GDP	GDP- > Energy	Energy- > GDP	GDP- > Energy
Bivariate GDP & HE	1850–2000	0.0123 (0.994)	8.8298 (0.012)	0.0392 (0.843)	8.0455 (0.005)
	1900–2000	1.7882 (0.181)	10.341 (0.001)	1.7923 (0.181)	9.1150 (0.003)
	1950–2000	0.3247 (0.850)	14.343 (0.001)		
Bivariate GDP & DE	1850–2000	5.8844 (0.015)	0.5195 (0.471)	5.3078 (0.021)	2.4836 (0.115)
	1900–2000	1.4505 (0.228)	0.2129 (0.644)	1.5294 (0.216)	2.7267 (0.099)
	1950–2000	0.2742 (0.872)	10.1098 (0.006)		
Multivariate GDP, DE, K, L	1850–2000	6.9394 (0.031)	1.4469 (0.485)	5.5187 (0.063)	3.2332 (0.198)
	1900–2000	6.5671 (0.037)	0.9707 (0.615)	4.6698 (0.097)	1.9436 (0.379)
	1950–2000	2.0888 (0.719)	19.5444 (0.000)		
Multivariate GRO, DE, K, L	1850–2000	6.6914 (0.035)	0.9950 (0.498)	3.4467 (0.032)	0.4250 (0.809)
	1900–2000	3.0552 (0.217)	2.2524 (0.325)	3.0378 (0.219)	1.4296 (0.489)
	1950–2000	9.7121 (0.045)	18.1519 (0.001)		

Notes: All variables are in log levels and all equations include a constant and a time trend. Statistics are chi-square statistics for excluding the first p lags of the variable listed first in the equation of the variable listed second. Significance levels in parentheses. Structural breaks are in 1900 and 1950.

Table 5
Causality tests: Demand function models.

Model	Period	Energy- > GDP	Price- > GDP	GDP- > Energy	Price- > Energy	GDP- > Price	Energy- > Price
GDP, DE, PE, simple trend	1850–2000	1.2242 (0.542)	29.131 (0.000)	4.5628 (0.102)	12.775 (0.002)	5.8744 (0.053)	7.9718 (0.019)
	1900–2000	6.5279 (0.163)	26.418 (0.000)	6.3804 (0.172)	31.402 (0.000)	3.699 (0.448)	6.2954 (0.178)
	1950–2000	0.5031 (0.777)	2.4693 (0.290)	11.458 (0.003)	13.787 (0.001)	5.2602 (0.072)	1.5143 (0.469)
GDP, HE, PO, simple trend	1850–2000	2.4696 (0.116)	10.465 (0.001)	5.8271 (0.016)	7.5631 (0.006)	0.0148 (0.903)	0.2500 (0.617)
	1900–2000	3.3500 (0.187)	7.5522 (0.023)	5.7254 (0.057)	7.4194 (0.024)	1.1222 (0.571)	2.5951 (0.273)
	1950–2000	3.3911 (0.066)	2.6685 (0.102)	11.187 (0.001)	1.5898 (0.207)	0.0008 (0.977)	0.0730 (0.787)
GDP, DE, PE, trend with 2 structural breaks (1900, 1950)	1850–2000	2.6006 (0.272)	28.761 (0.000)	8.2264 (0.016)	10.349 (0.006)	4.9551 (0.084)	6.4758 (0.039)
	1900–2000	1.4536 (0.483)	9.4172 (0.009)	9.6113 (0.008)	6.6202 (0.037)	1.1685 (0.558)	7.8743 (0.020)
GDP, HE, PO, trend with 2 structural breaks (1900, 1950)	1850–2000	2.0807 (0.149)	9.3287 (0.002)	6.5733 (0.010)	9.2369 (0.002)	0.0198 (0.888)	0.0474 (0.828)
	1900–2000	1.0787 (0.340)	3.1162 (0.044)	3.5564 (0.029)	5.4642 (0.004)	0.2222 (0.801)	0.3790 (0.685)

Notes: All variables are in log levels and all equations include a constant and a linear time trend. The test statistics are chi-square statistics with p-values given in parentheses.

The models in both Tables 6 and 7 are estimated for two specifications of the deterministic components: under the assumption of an unrestricted constant term in the VAR but no linear trends in the cointegrating relation (denoted as case 3 in Juselius, 2006, p.100) and under the assumption that there is a linear trend in the cointegration space (denoted as case 4 in Juselius, 2006). In Tables 8 to 11 we allow

Table 6
Johansen test for cointegration, no structural breaks.

Variables	# CI vectors under H0	Unrestricted constant				Unrestricted constant & linear trend in cointegration space			
		Trace stat	5% Crit. val	10% Crit. val.	Conclusion	Trace stat	5% Crit. val.	10% Crit. val.	Conclusion
1850–2000									
GRO, HE, K, L	0	37.43	47.21	44.49	H0	47.98	62.99	60.08	H0
	≤1	18.98	29.80	27.06	H0	26.09	42.92	39.75	H0
GDP, HE, K, L	0	33.02	47.21	44.49	H0	43.95	62.99	60.08	H0
	≤1	16.20	29.39	27.06	H0	24.73	42.92	39.75	H0
GRO, DE, K, L	0	41.04	47.21	44.49	H0	62.70	62.99	60.08	Reject H0 @ 10%
	≤1	18.69	29.79	27.06	H0	29.76	42.44	39.75	H0
GDP, DE, K, L	0	39.29	47.21	44.49	H0	55.93	62.99	60.08	H0
	≤1	20.03	29.79	27.06	H0	32.56	42.92	39.75	H0
1900–2000									
GRO, HE, K, L	0	34.73	47.21	44.49	H0	43.55	62.99	60.08	H0
	≤1	14.90	29.80	27.06	H0	20.89	42.92	39.75	H0
GDP, HE, K, L	0	32.37	47.21	44.49	H0	41.56	62.99	60.08	H0
	≤1	17.05	29.39	27.06	H0	21.38	42.92	39.75	H0
GRO, DE, K, L	0	38.79	47.21	44.49	H0	54.18	62.99	60.08	H0
	≤1	17.28	29.79	27.06	H0	29.38	42.44	39.75	H0
GDP, DE, K, L	0	31.21	47.21	44.49	H0	48.98	62.99	60.08	H0
	≤1	15.98	29.79	27.06	H0	25.99	42.92	39.75	H0
1950–2000									
GRO, HE, K, L	0	31.49	62.99	60.08	H0	47.44	62.99	60.08	H0
	≤1	17.60	42.92	39.75	H0	27.48	42.92	39.75	H0
GDP, HE, K, L	0	32.91	62.99	60.08	H0	53.54	62.99	60.08	H0
	≤1	18.79	42.92	39.75	H0	27.73	42.92	39.75	H0
GRO, DE, K, L	0	30.39	62.99	60.08	H0	48.03	62.99	60.08	H0
	≤1	16.24	42.44	39.75	H0	27.85	42.44	39.75	H0
GDP, DE, K, L	0	31.06	62.99	60.08	H0	50.06	62.99	60.08	H0
	≤1	17.16	42.92	39.75	H0	27.93	42.92	39.75	H0

Table 7
Johansen test for cointegration, no structural breaks, energy price included.

Variables	# CI vectors under H0	Unrestricted constant				Unrestricted constant & linear trend in cointegration space			
		Trace stat	5% Crit. val	10% Crit. val.	Conclusion	Trace stat	5% Crit. val.	10% Crit. val.	Conclusion
1850–2000									
GRO, HE, K, L, PE	0	68.34	69.81	65.81	Reject H0 @ 10%	83.7	88.8	84.38	H0
	≤1	39.46	47.85	44.49	H0	50.04	63.88	60.08	H0
GDP, HE, K, L, PE	0	63.90	69.81	65.81	H0	81.65	88.8	84.38	H0
	≤1	34.16	47.85	44.49	H0	50.75	63.88	60.08	H0
GRO, DE, K, L, PE	0	61.88	69.81	65.81	H0	81.65	88.8	84.38	H0
	≤1	29.90	47.85	44.49	H0	43.93	63.88	60.08	H0
GDP, DE, K, L, PE	0	66.77	69.81	65.81	Reject H0 @ 10%	83.72	88.8	84.38	H0
	≤1	31.94	47.85	44.49	H0	47.57	63.88	60.08	H0
1900–2000									
GRO, HE, K, L, PE	0	65.63	69.81	65.81	Reject H0 @ 10%	76.13	88.8	84.38	H0
	≤1	33.65	47.85	44.49	H0	43.65	63.88	60.08	H0
GDP, HE, K, L, PE	0	66.73	69.81	65.81	Reject H0 @ 10%	80.44	88.8	84.38	H0
	≤1	33.64	47.85	44.49	H0	46.68	63.88	60.08	H0
GRO, DE, K, L, PE	0	59.49	69.81	65.81	H0	80.96	88.8	84.38	H0
	≤1	31.27	47.85	44.49	H0	46.75	63.88	60.08	H0
GDP, DE, K, L, PE	0	61.08	69.81	65.81	H0	82.38	88.8	84.38	H0
	≤1	31.90	47.85	44.49	H0	48.64	63.88	60.08	H0
1950–2000									
GRO, HE, K, L, PE	0	65.21	69.81	65.81	H0	86.40	88.8	84.38	Reject H0 @ 10%
	≤1	38.30	47.85	44.49	H0	53.88	63.88	60.08	H0
GDP, HE, K, L, PE	0	63.82	69.81	65.81	H0	87.59	88.8	84.38	Reject H0 @ 10%
	≤1	37.17	47.85	44.49	H0	54.80	63.88	60.08	H0
GRO, DE, K, L, PE	0	67.20	69.81	65.81	H0	93.08	88.8	84.38	Reject H0 @ 5%
	≤1	33.91	47.85	44.49	H0	59.34	63.88	60.08	H0
GDP, DE, K, L, PE	0	65.30	69.81	65.81	Reject H0 @ 10%	91.26	88.8	84.38	Reject H0 @ 5%
	≤1	32.90	47.85	44.49	H0	58.81	63.88	60.08	H0

for structural breaks in the intercept and in the linear trend of the cointegration relation and so we only estimate the models that allow for linear trends (the case 4 model). Each model uses 2 lags in the levels specification as suggested by the Akaike information criterion.

In Table 6, the null hypothesis of no cointegration cannot be rejected at the 5% significance level in any model. However, for the model with the Divisia index of energy and gross output allowing for a trend in the cointegration relation, we come very close to rejection of the null of non-cointegration at the 5% level in the full sample (the trace statistic of 62.7 is very close to the critical value of 62.99). At the 10% significance level, the null of non-cointegration can be rejected while the hypothesis of at most one cointegrating relation cannot be rejected. Of course, we would expect more than one false rejection at the 10% level when 24 tests are carried out.

The models in Table 7 are the same as in Table 6 except that the price of energy is included alongside the other four variables. In the full sample, we can formally reject the null of zero cointegrating vectors for two of the eight models at the 10% level, while several of the models are very close to being significant at the 10% level. In the 1950–2000 subsample, we can reject the null of non-cointegration at either the 5% or 10% level for all models with a trend in the cointegration space despite the expected low power of the test in a sample of this size.

Table 8 presents the trace test statistics and the critical values for the model with structural breaks in 1900 and 1950 but without energy prices. Adding structural breaks every 50th year does not increase the rate of rejection of the null hypothesis of non-cointegration. On the contrary, the null hypothesis of no cointegration cannot be rejected in any model. In Table 9 we add energy prices to this model but the results stay the same: the null hypothesis of zero cointegration relations cannot be rejected.

Still, allowing for a structural break in the trend every 50th year is arbitrary. In Table 10 we conduct a similar analysis with the four variables from the production function framework but allowing for structural breaks in 1916 and 1973 instead of 1900 and 1950. The null hypothesis of no cointegration can now be rejected at the 10% level in the model with Divisia energy and gross output. This finding indicates that both the definition of variables and the choice of structural breaks in the cointegration relation can have an important effect on the results. In Table 11 we add energy prices to the analysis in

Table 8
Johansen test for cointegration, structural breaks in 1900 and 1950.

Model	# CI vectors under H0	Trace statistic	5% Crit. val	10% Crit. val.	Conclusion
GRO, HE, K, L	0	75.52	105.44	100.64	H0
	≤1	42.17	75.26	71.17	H0
GDP, HE, K, L	0	71.07	105.44	100.64	H0
	≤1	42.22	75.26	71.17	H0
GRO, DE, K, L	0	82.71	105.44	100.64	H0
	≤1	44.7	75.26	71.17	H0
GDP, DE, K, L	0	76.99	105.44	100.64	H0
	≤1	45.12	75.26	71.17	H0

Note: Models with changing intercept and broken linear trend in the cointegration space (HLr) and structural breaks in 1900 and 1950 (v1 = 0.33, v2 = 0.66).

Table 9
Johansen test for cointegration, structural breaks in 1900 and 1950, energy prices included.

Model	# CI vectors under H0	Trace statistic	5% critical value	10% critical value	Conclusion
GRO, HE,	0	111.37	139.37	133.9	H0
K, L, PE	≤1	76.26	105.44	100.64	H0
GDP, HE,	0	115.53	139.37	133.9	H0
K, L, PE	≤1	79.44	105.44	100.64	H0
GRO, DE,	0	108.1	139.37	133.9	H0
K, L, PE	≤1	71.71	105.44	100.64	H0
GDP, DE,	0	116.52	139.37	133.9	H0
K, L, PE	≤1	77.3	105.44	100.64	H0

Note: Models with changing intercept and broken linear trend in the cointegration space (HLr) and structural breaks in 1900 and 1950 (v1 = 0.33, v2 = 0.66).

Table 10
Johansen test for cointegration, structural breaks in 1916 and 1973.

Model	# CI vectors under H0	Trace statistic	5% critical value	10% critical value	Conclusion
GRO, HE, K, L	0	89.92	102.6	97.8	H0
	≤1	54.05	72.96	68.7	H0
	≤2	24.84	48.64	43.7	H0
GDP, HE, K, L	0	95.44	102.6	97.8	H0
	≤1	53.63	72.96	68.7	H0
	≤2	20.19	48.64	43.7	H0
GRO, DE, K, L	0	101.49	102.6	97.8	Reject H0 @ 10%
	≤1	61.23	72.96	68.7	H0
	≤2	25.5	48.64	43.7	H0
GDP, DE, K, L	0	93.81	102.6	97.8	H0
	≤1	54.27	72.96	68.7	H0
	≤2	20.02	48.64	43.7	H0

Note: Models with changing intercept and broken linear trend in the cointegration space (HLr) and structural breaks in 1916, 1973 (v1 = 0.44, v2 = 0.82).

Table 11
Johansen test for cointegration, structural breaks in 1916 and 1973, energy prices included.

Model	# CI vectors under H0	Trace statistic	5% critical value	10% critical value	Conclusion
GRO, HE,	0	143.96	136.2	130.74	Reject H0 @ 5%
K, L, PE	≤1	92.09	102.6	97.8	H0
	≤2	54.79	72.96	68.7	H0
GDP, HE,	0	145.13	136.2	130.74	Reject H0 @ 5%
K, L, PE	≤1	96.81	102.6	97.8	H0
	≤2	58.21	72.96	68.7	H0
GRO, DE,	0	133.22	136.2	130.74	Reject H0 @ 10%
K, L, PE	≤1	83.04	102.6	97.8	H0
	≤2	46.98	72.96	68.7	H0
GDP, DE,	0	140.75	136.2	130.74	Reject H0 @ 5%
K, L, PE	≤1	92.01	102.6	97.8	H0
	≤2	55.16	72.96	68.7	H0

Note: Models with changing intercept and broken linear trend in the cointegration space (HLr) and structural breaks in 1916, 1973 (v1 = 0.44, v2 = 0.82).

Table 10. We are now able to reject the null of no cointegration in all the models, but the test still only suggests at most one cointegrating vector.

In Tables 12 and 13 we report estimates of the cointegrating vector, β, and the adjustment parameters, α, for the models where we find cointegration in Tables 7 and 11. In Table 7 we found one cointegrating vector at the 10% significance level or higher for each model we tested for the 1950 to 2000 period. We found cointegration at the 10% significance level or higher for all models in Table 11, which cover the entire 1850–2000 period allowing for two structural breaks. We rejected non-cointegration in further isolated cases but do not report further results for those models, which might simply be cases of Type 1 error. We normalized the estimates of the cointegrating vectors on the energy variable and do not report the constant term or any of the trend terms for the models with structural breaks.

Capital is not significant in any of the long-run relationships in Table 12, but energy prices are highly significant in each, and labor, output, and the trend term are highly significant for the first two models. Output, energy quantity and price, and the time trend have the expected signs. If we interpret the labor variable as a proxy for population then we can interpret the output variable as the effect of income per capita, while the labor variable is the effect of increasing population while reducing income per capita. Therefore, the effect of population alone is the sum of these two elasticities. For the first two models income per capita has a greater than unit elasticity while the implied elasticity of population is rather small but positive. The elasticity of demand with respect to the price of energy is inelastic (0.28 to 0.38). Energy use declines autonomously at a rate of 1.4% to 1.8% per annum. The models for Divisia energy show a much lower but less precisely estimated income elasticity and an implied population elasticity of close to unity. The elasticity of energy demand with respect to prices is less inelastic (0.64 to 0.73) and the autonomous rate of reduction of energy use is lower too.

The adjustment parameter, α, for energy is, as expected, highly significant and negative in each model in Table 12, implying that energy is endogenous. The adjustment parameter for output is significant at the 5 or 10% level in the first three models but not in the fourth. Only one other variable in one model – the energy price in the third model – has an adjustment parameter that is significant at the 10% level. These causality test results conform well to our findings using the Toda–Yamamoto test in Table 5.[4] The results in Table 13 are similar but harder to interpret, probably because most variables are now endogenous.

[4] Testing for causality using a t-test on the adjustment parameters is not a formally correct test (Clarke and Mirza, 2006) but is intuitive and widely used in the literature.

6. Discussion and conclusions

Our review of the literature showed that multivariate models that include capital and perhaps labor inputs and/or improved measures of the energy input tend to find causality from energy to GDP. Results are more mixed for bivariate models. Models with oil prices, energy, and output tend to find that in the long-run GDP growth drives energy use while energy prices are exogenous in at least the short run.

For our Swedish data, a simple bivariate energy and GDP VAR shows causation from GDP to energy but this was reversed in the full sample period when we used a Divisia index of energy. But we still found causality from GDP to energy in the 1950–2000 subsample. A multivariate model including capital and labor inputs also showed causality from energy to GDP in the full 1850–2000 sample but from GDP to energy in the 1950–2000 sub-sample. These results for the most recent period are intriguing because Stern and Kander (2012) find that the contribution of energy to economic growth was much greater in the 19th and early 20th centuries than in the late 20th century. As the cost share of energy fell, its relative contribution to production fell too. The only other long-term study of energy-growth causality (Vaona, 2012) found mutual causation between non-renewable energy and GDP and from one measure of renewable energy to GDP using

Table 12
Parameter estimates for models in Table 7, 1950–2000.

Variables	HE	GRO	K	L	PE	Trend
β	1.000	−1.12	−0.07	0.83	0.38	0.014
		(−5.37)	(−0.38)	(2.98)	(6.16)	(4.01)
α	−1.34	−0.20	0.01	−0.03	0.41	
	(−5.93)	(−2.23)	(0.52)	(−0.55)	(1.42)	
Variables	**HE**	**GDP**	**K**	**L**	**PE**	**Trend**
β	1.000	−1.09	−0.17	0.88	0.28	0.018
		(−5.00)	(−0.93)	(2.87)	(3.97)	(4.60)
α	−1.25	−0.14	0.002	−0.035	0.42	
	(−6.25)	(−1.77)	(0.13)	(−0.63)	(1.61)	
Variables	**DE**	**GRO**	**K**	**L**	**PE**	**Trend**
β	1.000	−0.45	−0.23	−0.57	0.73	0.007
		(−1.47)	(−0.83)	(−1.39)	(7.81)	(1.28)
α	−0.44	−0.11	−0.02	−0.02	−0.31	
	(−3.28)	(−1.85)	(−1.34)	(−0.61)	(−1.74)	
Variables	**DE**	**GDP**	**K**	**L**	**PE**	**Trend**
β	1.00	−0.41	−0.37	−0.43	0.64	0.010
		(−1.45)	(−1.54)	(−1.08)	(7.12)	(1.98)
α	−0.51	−0.04	−0.02	−0.03	−0.28	
	(−3.7)	(−0.59)	(−1.23)	(−0.62)	(−1.42)	

t-statistics in parentheses.

Table 13
Parameter estimates for models in Table 11, 1850–2000.

Variables	HE	GRO	K	L	PE
β	1.000	−0.59	0.16	1.62	0.18
		(−3.71)	(0.10)	(4.72)	(2.34)
α	−0.39	−0.18	−0.00	−0.07	0.29
	(−4.88)	(−3.32)	(−0.44)	(−2.94)	(2.26)
Variables	HE	GDP	K	L	PE
β	1.000	−0.76	0.22	1.97	−0.06
		(−3.45)	(1.70)	(5.00)	(−0.63)
α	−0.35	−0.16	−0.01	−0.07	0.29
	(−4.73)	(−4.06)	(−1.31)	(−3.89)	(2.54)
Variables	DE	GRO	K	L	PE
β	1.000	−0.78	−0.15	1.16	−0.01
		(−6.89)	(−2.06)	(4.80)	(−0.13)
α	−0.33	0.06	−0.01	−0.10	0.86
	(−4.12)	(−0.79)	(−0.86)	(−3.27)	(5.03)
Variables	DE	GDP	K	L	PE
β	1.00	−0.81	−0.13	1.65	−0.00
		(−4.76)	(−1.30)	(5.44)	(−0.20)
α	−0.26	−0.07	−0.01	−0.09	0.60
	(−4.00)	(−1.40)	(−0.86)	(−3.83)	(4.36)

t-statistics in parentheses.

bivariate models. Our VAR models of GDP, energy quantity, and energy prices mostly find that energy prices, and particularly oil prices, are exogenous, that prices have a more significant impact on GDP than energy quantities, and that GDP and energy prices drive energy use. But the significance of the effect of energy prices on GDP was also attenuated in the 1950–2000 period. We find that the Granger causality technique is very sensitive to variable definition, choice of additional variables in the model, and sample periods.

We also estimated VECMs, using the Johansen procedure, allowing for both simple linear trends and time trends with structural breaks in the long-run relations. We found that a model that includes the price of energy in addition to output and the three factors of production and has structural breaks in 1916 and 1973 has at least one cointegrating vector. We also found cointegration in the 1950–2000 subsample for models that include the price of energy and a simple linear trend. Directions of causality in the long-run relations of the VECM models quite closely matched those found with the Granger causality tests. The long-run relationship seems to identify an energy demand model. However, VECMs that do not include the price of energy and have no structural breaks, or structural breaks at other times, only find cointegration for a few specifications, which could simply be explained by type 1 error. Vaona (2012) also found that non-cointegration could only be rejected when he allowed for a structural break.

These findings are in contrast to Stern and Kander (2012) who estimate a static nonlinear production function model. They found that when they introduce arbitrary structural breaks in the time trend every fifty years that represent a varying rate of technological change, they could not reject the null of cointegration using the Choi and Saikkonen (2010) nonlinear cointegration test. This suggests that the long-run relationship between energy and output is in fact nonlinear due to the low elasticity of substitution between energy and other inputs.

Acknowledgments

We thank Stephan Bruns, Paul Burke, Zsuzsanna Csereklyei, David Giles, Md. Shahiduzzaman and two anonymous referees for their very useful comments and advice. We thank Astrid Kander and Lennart Schön for providing the data. David Stern acknowledges funding from the Australian Research Council under Discovery Project DP120101088 and Kerstin Enflo acknowledges funding from the Wallander Research Foundation, grant number W2008:0357-1.

References

Bauer, D., Maynard, A., 2012. Persistence-robust surplus-lag Granger causality testing. J. Econ. 169 (2), 293–300.
Blanchard, O.J., Galí, J., 2008. The macroeconomic effects of oil price shocks: why are the 2000s so different from the 1970s? In: Galí, J., Gertler, M.J. (Eds.), International Dimensions of Monetary Policy. University of Chicago Press, Chicago.
Burbridge, J., Harrison, A., 1984. Testing for the effects of oil price rises using vector autoregressions. Int. Econ. Rev. 25, 459–484.
Chen, B., Pearl, J., 2012. Regression and causation: a critical examination of econometrics textbooks. Technical Report R-395.Computer Science Department, University of California, Los Angeles.
Chiou-Wei, S.Z., Chen, C.-F., Zhu, Z., 2008. Economic growth and energy consumption revisited — Evidence from linear and nonlinear Granger causality. Energy Econ. 30, 3063–3076.
Choi, I., Saikkonen, P., 2010. Tests for nonlinear cointegration. Econ. Theory 26 (2010), 682–709.
Chontanawat, J., Hunt, L.C., Pierse, R., 2008. Does energy consumption cause economic growth? Evidence from a systematic study of over 100 countries. J. Policy Model. 30, 209–220.
Clarke, J.A., Mirza, S., 2006. A comparison of some common methods for detecting Granger noncausality. J. Stat. Comput. Simul. 76, 207–231.
Costantini, V., Martini, C., 2010. The causality between energy consumption and economic growth: a multi-sectoral analysis using non-stationary cointegrated panel data. Energy Econ. 32, 591–603.
Edelstein, P., Kilian, L., 2009. How sensitive are consumer expenditures to retail energy prices? J. Monet. Econ. 56 (6), 766–779.
Ferraro, P.J., Hanauer, K.M., 2011. Causal Mechanisms of Protected Areas on Poverty. Paper presented at the AIJOS Capstone Conference, University of Michigan, Ann Arbor MI, September.
Gales, B., Kander, A., Malanima, P., Rubio, M., 2007. North versus South: energy transition and energy intensity in Europe over 200 years. Eur. Rev. Econ. Hist. 11 (2), 219–253.
Gerring, J., 2010. Causal mechanisms: yes, but…. Comp. Polit. Stud. 43 (11), 1499–1526.
Ghali, K.H., El-Sakka, M.I.T., 2004. Energy use and output growth in Canada: a multivariate cointegration analysis. Energy Econ. 26, 225–238.
Giles, D.E.A., Godwin, R.T., 2012. Testing for multivariate cointegration in the presence of structural breaks: p values and critical values. Appl. Econ. Lett. 19, 1561–1565.
Glasure, Y.U., 2002. Energy and national income in Korea: further evidence on the role of omitted variables. Energy Econ. 24, 355–365.
Granger, C.W.J., 1988. Some recent developments in a concept of causality. J. Econ. 39, 199–211.
Greasley, D., Madsen, J.B., Wohar, M.E., 2013. Long-run growth empirics and new challenges for unified theory. Appl. Econ. 45, 3973–3987.
Hamilton, J.D., 1983. Oil and the macroeconomy since World War II. J. Polit. Econ. 91, 228–248.
Hamilton, J.D., 2009. Causes and consequences of the oil shock of 2007–08. Brook. Pap. Econ. Act. (1), 215–261.
Hamilton, J.D., 2009. Understanding crude oil prices. Energy J. 30 (2), 179–206.
Hassani, H., Zhigljavsky, A., Patterson, K., Soofi, A., 2010. A comprehensive causality test based on the singular spectrum analysis. In: Illari, P.M., Russo, F., Williamson, J. (Eds.), Causality in Science. Oxford University Press.
Heckman, J.J., 2008. Econometric causality. Int. Stat. Rev. 76 (1), 1–27.
Hiemstra, C., Jones, J.D., 1994. Testing for linear and nonlinear Granger causality in the stock price-volume relation. J. Finance 49 (5), 1639–1664.
Hondroyiannis, G., Lolos, S., Papapetrou, E., 2002. Energy consumption and economic growth: assessing the evidence from Greece. Energy Econ. 24, 319–336.
International Monetary Fund, 2011. World Economic Outlook April 2011: Tensions from the Two-Speed Recovery Unemployment, Commodities, and Capital Flows. International Monetary Fund Publication Services, Washington DC.
Johansen, S., Mosconi, R., Nielsen, B., 2000. Cointegration analysis in the presence of structural breaks in the deterministic trend. Econ. J. 3, 216–249.
Joyeux, R., 2007. Structural breaks in cointegration analysis: an applied analysis, In: Bhaskara Rao, B. (Ed.), Cointegration for the Applied Economist, 2nd ed. Palgrave Macmillan.
Joyeux, R., Ripple, R.D., 2011. Energy consumption and real income: a panel cointegration multi-country study. Energy J. 32 (2), 107–141.
Juselius, K., 2006. The Cointegrated VAR Model: Methodology and Applications. Oxford University Press.
Kander, A., 2002. Economic growth, energy consumption and CO_2 emissions in Sweden 1800–2000. Lund Studies in Economic History 19.
Kraft, J., Kraft, A., 1978. On the relationship between energy and GNP. J. Energy Dev. 3, 401–403.
Krantz, O., Schön, L., 2007. Swedish historical national accounts 1800–2000. Lund Studies in Economic History 41.
Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against an alternative of a unit root: how sure are we that economic time series have a unit root? J. Econ. 54, 159–178.
Lee, C.-C., Chang, C.-P., 2008. Energy consumption and economic growth in Asian economies: a more comprehensive analysis using panel data. Resour. Energy Econ. 30 (1), 50–65.

Lee, J., Strazicich, M.C., 2003. Minimum LM unit root test with two structural breaks. Rev. Econ. Stat. 85 (4), 1048–1062.

Lee, J., Strazicich, M.C., 2004. Minimum LM unit root test with one structural break. *Appalachian State University Working Paper* 0417.

Lee, C.-C., Chang, C.-P., Chen, P.-F., 2008. Energy-income causality in OECD countries revisited: the key role of capital stock. Energy Econ. 30, 2359–2373.

Lütkepohl, H., 1982. Non-causality due to omitted variables. J. Econ. 19, 367–378.

Lütkepohl, H., Saikkonen, P., Trenkler, C., 2004. Testing for the cointegrating rank of a VAR with level shift at unknown time. Econometrica 72, 647–662.

Malanima, P., 2006. Energy consumption in Italy in the 19th and 20th century. Consiglio Nazionale delle Ricerche.Instituto di Studi sulle Società del Mediterraneo.

Masih, A.M.M., Masih, R., 1997. On the temporal causal relationship between energy consumption, real income, and prices: some new evidence from Asian-energy dependent NICs based on a multivariate cointegration/vector error-correction approach. J. Policy Model. 19, 417–440.

Oh, W., Lee, K., 2004. Causal relationship between energy consumption and GDP revisited: the case of Korea 1970–1999. Energy Econ. 26, 51–59.

Ohanian, L.E., 1988. The spurious effects of unit roots on vector autoregressions: a Monte Carlo study. J. Econ. 39, 251–266.

Ozturk, I., 2010. A literature survey on energy-growth nexus. Energy Policy 38, 340–349.

Park, J.V., Sung, J., 1994. Testing for unit roots in models with structural-change. Econ. Theory 10 (5), 917–936.

Phillips, P., Perron, P., 1988. Testing for a unit root in time series regression. J. Econ. 33, 335–346.

Qin, D., 2011. Rise of VAR modeling approach. J. Econ. Surv. 25 (1), 156–174.

Sargent, T., 1979. Estimating vector autoregressions using methods not based on explicit economic theories, *Federal Reserve Bank of Minneapolis*. Q. Rev. 3 (3), 8–15.

Schön, L., 2008. Sweden — Economic Growth and Structural Change, 1800–2000. In: Whaples, Robert (Ed.), EH.Net Encyclopedia (February 10, 2008. URL: http://eh.net/encyclopedia/article/schon.sweden).

Schön, L., 2010. Sweden's Road to Modernity: An Economic History. SNS Förlag, Stockholm.

Schwert, G.W., 1989. Tests for unit roots: a Monte Carlo investigation. J. Bus. Econ. Stat. 7, 147–159.

Sims, C.A., 1972. Money, income and causality. Am. Econ. Rev. 62, 540–552.

Sims, C.A., 1980. Macroeconomics and reality. Econometrica 48, 1–48.

Stanley, T.D., Jarrell, S.B., Doucouliagos, H., 2010. Could it be better to discard 90% of the data? A statistical paradox. Am. Stat. 64 (1), 70–77.

Stern, D.I., 1993. Energy use and economic growth in the USA: a multivariate approach. Energy Econ. 15, 137–150.

Stern, D.I., 2000. A multivariate cointegration analysis of the role of energy in the U.S. macroeconomy. Energy Econ. 22, 267–283.

Stern, D.I., 2010. Energy quality. Ecol. Econ. 69 (7), 1471–1478.

Stern, D.I., 2011. The role of energy in economic growth. Ann. N. Y. Acad. Sci. 1219, 26–51.

Stern, D.I., Common, M.S., 2001. Is there an environmental Kuznets curve for sulfur? J. Environ. Econ. Manag. 41, 162–178.

Stern, D.I., Kander, A., 2012. The role of energy in the industrial revolution and modern economic growth. Energy J. 33 (3), 125–152.

Sugihara, G., May, R., Ye, H., Hsieh, C.-H., Deyle, E., Fogarty, M., Munch, S., 2012. Detecting causality in complex ecosystems. Science 338, 496–500.

Sun, X., 2008. Assessing nonlinear Granger causality from multivariate time series. In: Daelemans, W., Goethals, B., Morik, K. (Eds.), Machine Learning and Knowledge Discovery in Databases, European Conference, ECML PKDD 2008, Antwerp, Belgium, September 15–19, 2008, Proceedings, Part II. Lecture Notes in Computer Science, 5212. Springer, Heidelberg, pp. 440–455.

Toda, H.Y., Phillips, P.C.B., 1993. The spurious effect of unit roots on vector autoregressions: an analytical study. J. Econ. 59, 229–255.

Toda, H.Y., Yamamoto, T., 1995. Statistical inference in vector autoregressions with possibly integrated processes. J. Econ. 66, 225–250.

Vaona, A., 2012. Granger non-causality between (non)renewable energy consumption and output in Italy since 1861: the (ir)relevance of structural breaks. Energy Policy 45, 226–236.

Warr, B., Ayres, R.U., 2010. Evidence of causality between the quantity and quality of energy consumption and economic growth. Energy 35, 1688–1693.

Wilde, J., 2012. Effects of simultaneity on testing Granger-causality — a cautionary note about statistical problems and economic misinterpretations. Working Paper, Institute for Empirical Economic Research, Univeristy of Osnabrück, p. 93.

Yu, E.S.H., Jin, J.C., 1992. Cointegration tests of energy consumption, income, and employment. Resour. Energy 14, 259–266.