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
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.